Activity Report 2011

Section New Results

Edition: 2012-03-23
<table>
<thead>
<tr>
<th>No.</th>
<th>Project-Team</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>BYMOORE Exploratory Action</td>
<td>9</td>
</tr>
<tr>
<td>2.</td>
<td>ABSTRACTION Project-Team</td>
<td>11</td>
</tr>
<tr>
<td>3.</td>
<td>ALF Project-Team</td>
<td>17</td>
</tr>
<tr>
<td>4.</td>
<td>ALGORITHMS Project-Team</td>
<td>26</td>
</tr>
<tr>
<td>5.</td>
<td>AOSTE Project-Team</td>
<td>27</td>
</tr>
<tr>
<td>6.</td>
<td>ARENAIRE Project-Team</td>
<td>32</td>
</tr>
<tr>
<td>7.</td>
<td>ATEAMS Project-Team</td>
<td>37</td>
</tr>
<tr>
<td>8.</td>
<td>CAIRN Project-Team</td>
<td>39</td>
</tr>
<tr>
<td>9.</td>
<td>CAMUS Team</td>
<td>49</td>
</tr>
<tr>
<td>10.</td>
<td>CARAMEL Project-Team</td>
<td>53</td>
</tr>
<tr>
<td>11.</td>
<td>CARTE Project-Team</td>
<td>56</td>
</tr>
<tr>
<td>12.</td>
<td>CASCADE Project-Team (section vide)</td>
<td>59</td>
</tr>
<tr>
<td>13.</td>
<td>CASSIS Project-Team</td>
<td>60</td>
</tr>
<tr>
<td>14.</td>
<td>CELTIQUE Project-Team</td>
<td>68</td>
</tr>
<tr>
<td>15.</td>
<td>COMETE Project-Team</td>
<td>71</td>
</tr>
<tr>
<td>16.</td>
<td>COMPSYS Project-Team</td>
<td>74</td>
</tr>
<tr>
<td>17.</td>
<td>CONTRAINTES Project-Team</td>
<td>80</td>
</tr>
<tr>
<td>18.</td>
<td>DART Project-Team</td>
<td>84</td>
</tr>
<tr>
<td>19.</td>
<td>ESPRESSO Project-Team</td>
<td>92</td>
</tr>
<tr>
<td>20.</td>
<td>FORMES Team</td>
<td>101</td>
</tr>
<tr>
<td>21.</td>
<td>GALAAD Project-Team</td>
<td>108</td>
</tr>
<tr>
<td>22.</td>
<td>GALLIUM Project-Team</td>
<td>117</td>
</tr>
<tr>
<td>23.</td>
<td>GEOMETRICA Project-Team</td>
<td>124</td>
</tr>
<tr>
<td>24.</td>
<td>LFANT Project-Team</td>
<td>133</td>
</tr>
<tr>
<td>25.</td>
<td>LICIT Exploratory Action</td>
<td>136</td>
</tr>
<tr>
<td>26.</td>
<td>MARELLE Project-Team</td>
<td>139</td>
</tr>
<tr>
<td>27.</td>
<td>MEXICO Project-Team</td>
<td>142</td>
</tr>
<tr>
<td>28.</td>
<td>MOSCOVA Project-Team</td>
<td>145</td>
</tr>
<tr>
<td>29.</td>
<td>PAREO Project-Team</td>
<td>150</td>
</tr>
<tr>
<td>30.</td>
<td>PARKAS Team</td>
<td>154</td>
</tr>
<tr>
<td>31.</td>
<td>PARSIFAL Project-Team</td>
<td>158</td>
</tr>
<tr>
<td>32.</td>
<td>PLR2 Project-Team</td>
<td>161</td>
</tr>
<tr>
<td>33.</td>
<td>POP ART Project-Team</td>
<td>167</td>
</tr>
<tr>
<td>34.</td>
<td>PROVAL Project-Team</td>
<td>179</td>
</tr>
<tr>
<td>35.</td>
<td>S4 Project-Team</td>
<td>182</td>
</tr>
<tr>
<td>36.</td>
<td>SALSA Project-Team</td>
<td>186</td>
</tr>
<tr>
<td>37.</td>
<td>SECRET Project-Team</td>
<td>190</td>
</tr>
<tr>
<td>38.</td>
<td>SECSI Project-Team</td>
<td>194</td>
</tr>
<tr>
<td>39.</td>
<td>TANC Project-Team</td>
<td>199</td>
</tr>
<tr>
<td>40.</td>
<td>TASC Project-Team</td>
<td>201</td>
</tr>
</tbody>
</table>
41. TRIO Project-Team ................................................................. 205
42. TYPICAL Project-Team ...................................................... 213
43. VASY Project-Team ............................................................. 218
44. VEGAS Project-Team .......................................................... 228
45. VERIDIS Team ................................................................. 231
46. VERTECS Project-Team ...................................................... 236

APPLIED MATHEMATICS, COMPUTATION AND SIMULATION

47. ALEA Project-Team ............................................................. 241
48. APICS Project-Team ........................................................... 246
49. ASPI Project-Team ............................................................. 255
50. BACCHUS Team ............................................................... 257
51. BIPOP Project-Team .......................................................... 263
52. CAD Team ................................................................. 269
53. CAGIRE Team ................................................................. 286
54. CALVI Project-Team ........................................................... 287
55. CLASSIC Project-Team ....................................................... 293
56. COMMANDS Project-Team ................................................ 296
57. CONCHA Project-Team ...................................................... 300
58. CORIDA Project-Team ....................................................... 309
59. CQFD Project-Team .......................................................... 313
60. DEFI Project-Team ............................................................ 321
61. DISCO Team ................................................................. 328
62. DOLPHIN Project-Team .................................................... 335
63. GAMMA3 Project-Team ...................................................... 341
64. GECO Team ................................................................. 346
65. GEOSTAT Project-Team ..................................................... 349
66. I4S Team ................................................................. 353
67. IPSO Project-Team ........................................................... 355
68. MATHFI Project-Team ...................................................... 363
69. MAXPLUS Project-Team .................................................... 367
70. MC2 Project-Team ........................................................... 384
71. MICMAC Project-Team ...................................................... 388
72. MISTIS Project-Team ......................................................... 396
73. MODAL Team ............................................................... 404
74. NACHOS Project-Team ...................................................... 407
75. NANO-D Team ............................................................... 412
76. NECS Project-Team .......................................................... 421
77. NON-A Team ................................................................. 432
78. OPALE Project-Team ....................................................... 437
79. POEMS Project-Team ....................................................... 446
80. PUMAS Team ............................................................... 464
81. REALOPT Project-Team ................................................................. 469
82. REGULARITY Team ................................................................. 476
83. SELECT Project-Team ............................................................... 487
84. SEQUEL Project-Team ............................................................... 490
85. SIERRA Project-Team ............................................................... 499
86. SIMPAF Project-Team ............................................................... 506
87. SMASH Project-Team ............................................................... 509
88. TAO Project-Team ................................................................. 512
89. TOSCA Project-Team ............................................................... 519
90. TROPICS Project-Team ............................................................ 523

COMPUTATIONAL SCIENCES FOR BIOLOGY, MEDICINE AND THE ENVIRONMENT
91. ABS Project-Team ................................................................. 526
92. AMIB Project-Team ................................................................. 529
93. ANUBIS Project-Team .............................................................. 534
94. ASCLEPIOS Project-Team .......................................................... 536
95. ATHENA Project-Team ............................................................ 551
96. BAMBOO Team ................................................................. 563
97. BANG Project-Team ............................................................... 570
98. Beagle Team ................................................................. 582
99. BIGS Project-Team ............................................................... 584
100. BIOCORE Project-Team ........................................................ 590
101. BONSAI Project-Team ........................................................... 598
102. CARMEN Team ................................................................. 600
103. CLIME Project-Team ............................................................ 601
104. CORTEX Project-Team .......................................................... 615
105. DEMAR Project-Team ............................................................ 621
106. DIGIPLANTE Team .............................................................. 638
107. DRACULA Project-Team ......................................................... 643
108. ESTIME Project-Team (section vide) ........................................ 644
109. FLUMINANCE Project-Team ................................................... 645
110. GALEN Team ................................................................. 654
111. IBIS Project-Team ............................................................... 657
112. MACS Project-Team ............................................................. 662
113. MAGIQUE-3D Project-Team .................................................... 665
114. MAGNOME Project-Team ....................................................... 673
115. MASAIE Project-Team ........................................................... 677
116. MODEMIC Team ................................................................. 680
117. MOISE Project-Team ........................................................... 688
118. NEUROMATHCOMP Project-Team ............................................ 704
119. NUMED Project-Team (section vide) ........................................ 711
120. PARIETAL Project-Team ......................................................... 712
<table>
<thead>
<tr>
<th>Project-Team</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>REO Project-Team</td>
<td>721</td>
</tr>
<tr>
<td>S.H.A.M.A.N Team</td>
<td>725</td>
</tr>
<tr>
<td>SAGE Project-Team</td>
<td>730</td>
</tr>
<tr>
<td>SERPICO Team</td>
<td>738</td>
</tr>
<tr>
<td>SISYPHE Project-Team</td>
<td>742</td>
</tr>
<tr>
<td>STEEP Exploratory Action</td>
<td>747</td>
</tr>
<tr>
<td>SYMBIOSE Project-Team</td>
<td>749</td>
</tr>
<tr>
<td>VIRTUAL PLANTS Project-Team</td>
<td>753</td>
</tr>
<tr>
<td>VISAGES Project-Team</td>
<td>765</td>
</tr>
<tr>
<td>ACES Project-Team</td>
<td>775</td>
</tr>
<tr>
<td>ADAM Project-Team</td>
<td>779</td>
</tr>
<tr>
<td>ALGORILLE Project-Team</td>
<td>781</td>
</tr>
<tr>
<td>AMAZONES Team</td>
<td>789</td>
</tr>
<tr>
<td>ARLES Project-Team</td>
<td>793</td>
</tr>
<tr>
<td>ASAP Project-Team</td>
<td>802</td>
</tr>
<tr>
<td>ASCOLA Project-Team</td>
<td>812</td>
</tr>
<tr>
<td>ATLANMOD Project-Team</td>
<td>819</td>
</tr>
<tr>
<td>CEPAGE Project-Team</td>
<td>821</td>
</tr>
<tr>
<td>CIDRE Project-Team</td>
<td>828</td>
</tr>
<tr>
<td>DIONYSOS Project-Team</td>
<td>833</td>
</tr>
<tr>
<td>DISTRIBCOM Project-Team</td>
<td>840</td>
</tr>
<tr>
<td>DNET Team</td>
<td>849</td>
</tr>
<tr>
<td>FOCUS Project-Team</td>
<td>851</td>
</tr>
<tr>
<td>GANG Project-Team</td>
<td>856</td>
</tr>
<tr>
<td>GRAAL Project-Team</td>
<td>864</td>
</tr>
<tr>
<td>GRAND-LARGE Project-Team</td>
<td>878</td>
</tr>
<tr>
<td>HIEPACS Project-Team</td>
<td>888</td>
</tr>
<tr>
<td>HIPERCOM Project-Team</td>
<td>892</td>
</tr>
<tr>
<td>INDES Project-Team</td>
<td>902</td>
</tr>
<tr>
<td>KERDATA Team</td>
<td>907</td>
</tr>
<tr>
<td>MADYNES Project-Team</td>
<td>914</td>
</tr>
<tr>
<td>MAESTRO Project-Team</td>
<td>922</td>
</tr>
<tr>
<td>MASCOTTE Project-Team</td>
<td>933</td>
</tr>
<tr>
<td>MESCAL Project-Team</td>
<td>946</td>
</tr>
<tr>
<td>MOAIS Project-Team</td>
<td>950</td>
</tr>
<tr>
<td>MYRIADS Team</td>
<td>952</td>
</tr>
<tr>
<td>OASIS Project-Team</td>
<td>957</td>
</tr>
<tr>
<td>PHOENIX Project-Team</td>
<td>961</td>
</tr>
<tr>
<td>PLANETE Project-Team</td>
<td>963</td>
</tr>
<tr>
<td>POPS Project-Team</td>
<td>978</td>
</tr>
</tbody>
</table>
161. RAP Project-Team ................................................................. 986
162. REGAL Project-Team .............................................................. 991
163. RESO Project-Team ............................................................... 999
164. RMOD Project-Team .............................................................. 1004
165. RUNTIME Project-Team .......................................................... 1009
166. SARDES Project-Team ........................................................... 1015
167. SCORE Team ........................................................................ 1022
168. SWING Team ....................................................................... 1026
169. TREC Project-Team ................................................................. 1034
170. TRISKELL Project-Team ............................................................ 1048

PERCEPTION, COGNITION, INTERACTION

171. ALICE Project-Team ............................................................... 1051
172. ALPAGE Project-Team ............................................................. 1055
173. ARIANA Project-Team ............................................................. 1065
174. AROBAS Project-Team ............................................................. 1082
175. ARTIS Project-Team ................................................................. 1094
176. AVIZ Project-Team ................................................................. 1104
177. AXIS Project-Team ................................................................. 1107
178. COPRIN Project-Team ............................................................. 1115
179. DAHU Project-Team ............................................................... 1121
180. DREAM Project-Team .............................................................. 1123
181. E-MOTION Project-Team .......................................................... 1127
182. EDELWEISS Project-Team ....................................................... 1148
183. EVASION Project-Team ........................................................ 1158
184. EXMO Project-Team ............................................................ 1166
185. FLOWERS Project-Team ......................................................... 1168
186. GRAPHIK Project-Team ........................................................ 1176
187. IMARA Project-Team ............................................................. 1180
188. IMEDIA Project-Team ............................................................. 1186
189. IN-SITU Project-Team ........................................................... 1194
190. IPARLA Project-Team ............................................................. 1198
191. LAGADIC Project-Team ........................................................... 1203
192. LEAR Project-Team ................................................................. 1206
193. LEO Team ............................................................................ 1212
194. MAGRIT Project-Team ............................................................ 1220
195. MAIA Project-Team ................................................................. 1224
196. METISS Project-Team ............................................................ 1236
197. MIMETIC Team .................................................................... 1245
198. MINT Team ......................................................................... 1253
199. Morpheo Team ..................................................................... 1258
200. MOSTRARE Project-Team ....................................................... 1264
4. New Results

4.1. Iterative Optimization for the Data Center (Alchemy-related research)

This result corresponds to research started within Alchemy, and it is less related to the objectives of ByMoore itself.

Iterative optimization is a simple but powerful approach that searches for the best possible combination of compiler optimizations for a given workload. However, each program, if not each data set, potentially favors a different combination. As a result, iterative optimization is plagued by several practical issues that prevent it from being widely used in practice: a large number of runs are required for finding the best combination; the process is inherently data set sensitive; and the exploration process incurs significant overhead that needs to be compensated for by performance benefits. Therefore, while iterative optimization has been shown to have significant performance potential, it is seldomly used in production compilers.

We propose [4] Iterative Optimization for the Data Center (IODC): we show that servers and data centers offer a context in which all of the above hurdles can be overcome. The basic idea is to spawn different combinations across workers and recollect performance statistics at the master, which then evolves to the optimum combination of compiler optimizations. IODC carefully manages costs and benefits, and is transparent to the end user.

We evaluate IODC using both MapReduce and throughput server applications. In order to reflect the large number of users interacting with the system, we gather a very large collection of data sets (at least 1000 and up to several million unique data sets per program), for a total storage of 10.7TB, and 568 days of CPU time. We report an average performance improvement of $1.48 \times$, and up to $2.08 \times$, for the MapReduce applications, and $1.14 \times$, and up to $1.39 \times$, for the throughput server applications.

4.2. Statistical Performance Comparisons of Computers (Alchemy-related research)

This result corresponds to research started within Alchemy, and it is less related to the objectives of ByMoore itself.

As a fundamental task in computer architecture research, performance comparison has been continuously hampered by the variability of computer performance. In traditional performance comparisons, the impact of performance variability is usually ignored (i.e., the means of performance measurements are compared regardless of the variability), or in the few cases where it is factored in using parametric confidence techniques, the confidence is either erroneously computed based on the distribution of performance measurements, instead of the distribution of sample mean of performance measurements, or too few measurements are considered for the distribution of sample mean to be normal. We first illustrate how such erroneous practices can lead to incorrect comparisons.

Then, we propose [3] a non-parametric Hierarchical Performance Testing (HPT) framework for performance comparison, which is significantly more practical than standard parametric confidence tests because it does not require to collect a large number of measurements in order to achieve a normal distribution of the sample mean. This HPT framework has been implemented as an open-source software.
4.3. Implementation of Signal Processing Tasks on Neuromorphic Hardware

Because of power and reliability issues, computer architects are forced to explore new types of architectures, such as heterogeneous systems embedding hardware accelerators. Neuromorphic systems are good candidate accelerators that can perform efficient and robust computing for certain classes of applications. We propose a spiking neurons based accelerator, with its hardware and software, that can be easily programmed to execute a wide range of signal processing applications. A library of operators is built to facilitate implementation of various types of applications. Automated placement and routing software tools are used to map these applications onto the hardware. Altogether, this system aims at providing to the user a simple way to implement signal processing tasks on neuromorphic hardware.

4.4. Automatic Abstraction and Fault Tolerance in Cortical Microarchitectures

Recent advances in the neuroscientific understanding of the brain are bringing about a tantalizing opportunity for building synthetic machines that perform computation in ways that differ radically from traditional Von Neumann machines. These brain-like architectures, which are premised on our understanding of how the human neocortex computes, are highly fault-tolerant, averaging results over large numbers of potentially faulty components, yet manage to solve very difficult problems more reliably than traditional algorithms. A key principle of operation for these architectures is that of automatic abstraction: independent features are extracted from highly disordered inputs and are used to create abstract invariant representations of the external entities. This feature extraction is applied hierarchically, leading to increasing levels of abstraction at higher levels in the hierarchy. This work describes and evaluates a biologically plausible computational model for this process, and highlights the inherent fault tolerance of the biologically-inspired algorithm. We introduce a stuck-at fault model for such cortical networks, and describe how this model maps to hardware faults that can occur on commodity GPGPU cores that used to realize the model in software. We show experimentally that the model software implementation can intrinsically preserve its functionality in the presence of faulty hardware, without requiring any reprogramming or recompilation. This model is a first step towards developing a comprehensive and biologically-plausible understanding of the computational algorithms and microarchitecture of computing systems that mimic the human cortex, and to applying them to the robust implementation of computational tasks on future computing systems built of faulty components.
6. New Results

6.1. Abstractions of Functions

Participants: Patrick Cousot, Radhia Cousot.

The idea of domain segmentation for arrays [18] has been extended to the abstraction of functions [41] by combination of a partitioning of their domain of definition and a functional or relational abstraction of blocks into their codomain [17].

6.2. Analysis of Biological Pathways

We have improved our framework to design and analyze biological networks. This framework focused on protein-protein interaction networks described as graph rewriting systems. Such networks can be used to model some signaling pathways that control the cell cycle. The task is made difficult due to the combinatorial blow up in the number of reachable species (i.e., non-isomorphic connected components of proteins).

6.2.1. Automatic Reduction of Differential Semantics

Participants: Ferdinanda Camporesi, Vincent Danos [University of Edinburgh], Jérôme Feret, Walter Fontana [Harvard Medical School], Russ Harmer [Harvard Medical School], Jean Krivine [Paris VII].

We have developed an abstract interpretation-based framework that enables the reduction of the differential semantics for protein-protein interaction networks. Results are sound since trajectories in the abstract system are projections of the trajectories in the concrete system.

The flow of information is a key element in our model reduction framework because it enables the identification of the correlations which are useless when computing observables of interest. Thus there is a need of providing good trade-off in the description of the flow of information throughout the biochemical structure of chemical species.

The notion of symmetries between sites is also important, since knowing that two sites have exactly the same capabilities of interaction enable exact quotienting (or lumping) of the set of reachable species.

In [13], [14], we have proposed a heterogeneous over-approximation of the flow of information where the flow that is attached to an agent can depend on its relative position in a chemical species. Moreover, we have showed how to use symmetries between sites so as to define another model reduction and we have proposed an algebraic product to combine model reductions, the product of two reduced models being the least abstract model which is at least as abstract as both model.

6.2.2. Automatic Reduction of Stochastic Semantics

Participants: Ferdinanda Camporesi, Jérôme Feret, Thomas Henzinger [Institute of Science and Technology, Austria], Heinz Koeppl [ETH Zürich], Tatjana Petrov [ETH Zürich].

We have proposed an abstract interpretation-based framework for reducing the state-space of stochastic semantics for protein-protein interaction networks. Our framework ensures that the trace distribution of the reduced system is the exact projection of the trace distribution of the concrete system. Moreover, when the abstraction is complete, if any pair of concrete states that have the same abstraction are equiprobable at initial state, any pair of concrete states that share the same abstraction are equiprobable at any time $t$.

In [12], we have formalized the model reduction framework for the stochastic semantics and we have established the relationships with the notions of lumpability, and bisimulation is established.
6.3. Automatic Array Content Analysis by Segmentation

Participants: Patrick Cousot, Radhia Cousot, Francesco Logozzo [Microsoft Research (Redmond, USA)].

In [18], we introduce FunArray, a parametric segmentation abstract domain functor for the fully automatic and scalable analysis of array content properties. The functor enables a natural, painless and efficient lifting of existing abstract domains for scalar variables to the analysis of uniform compound data-structures such as arrays and collections (as well as matrices when instantiating the functor on itself). The analysis automatically and semantically divides arrays into consecutive non-overlapping possibly empty segments. Segments are delimited by sets of bound symbolic expressions and abstracted uniformly. All bound expressions appearing in a set are equal in the concrete. The FunArray can be naturally combined via reduced product with any existing analysis for scalar variables. The bound expressions, the segment abstractions and the reduction operator are the three parameters of the analysis. Once the functor has been instantiated with fixed parameters, the analysis is fully automatic.

We first prototyped FunArray in Arrayal to adjust and experiment with the abstractions and the algorithms to obtain the appropriate precision/ratio cost. Then it was implemented into CCCHECK (formerly CLOUSOT), an abstract interpretation-based static contract checker for .NET by Francesco Logozzo. The precision and the performance of the analysis has been empirically validated by running it on the main libraries of .NET and on its own code. It was able to infer thousands of invariants and to verify the implementation with a modest overhead (circa 1%). To the best of our knowledge this is the first analysis of this kind applied to such a large code base, and proven to scale.

6.4. Extrapolation operators for combinations of abstract domains

Participants: Agostino Cortesi [Università Ca’ Foscari di Venizia], Matteo Zanioli.

Extrapolation operators are crucial to ensure the scalability of the analysis to large software systems. In [10], we set the ground for a systematic design of widening and narrowing operators, by comparing the different definitions introduced in the literature and by discussing how to tune them in case of domain abstraction and domains’ combination through Cartesian and reduced products.

6.5. Grammar Semantics, Analysis and Parsing

Participants: Patrick Cousot, Radhia Cousot.

In [11], we study the abstract interpretations of a fixpoint protoderivation semantics defining the maximal derivations of a transitional semantics of context-free grammars akin to pushdown automata. The result is a hierarchy of bottom-up or top-down semantics refining the classical equational and derivational language semantics and including Knuth grammar problems, classical grammar flow analysis algorithms, and parsing algorithms.

6.6. Information Flow

The analysis of the flow of information in a program consists in detecting the propagation of sensitive information through the program points of this program thanks to a dependency analysis.

6.6.1. Dependency Analysis and Numerical Invariants

Participants: Agostino Cortesi [Università Ca’ Foscari di Venizia], Matteo Zanioli.

A new framework has been proposed in [16], that combines variable dependency analysis, based on propositional formulas, and variables’ value analysis, based on generic numerical domains.

6.6.2. Leakage Analysis

Participants: Matteo Zanioli [Correspondent], Pietro Ferrara [ETH, Zurich], Agostino Cortesi [Università Ca’ Foscari].
In [24], we present SAILS, a new tool that combines SAMPLe, a generic static analyzer, and a sophisticated domain for leakage analysis. This tool does not require to modify the original language, since it works with mainstream languages like JAVA™, and it does not require any manual annotation. SAILS can combine the information leakage analysis with different heap abstractions, inferring information leakage over programs with complex data structures. SAILS has been applied to the analysis of the SecuriBench-micro suite. The experimental results underline the effectiveness of the analysis, since SAILS is in position to analyze several benchmarks in about 1 second without producing false alarms in more than 90% of the programs.

6.7. Linear Absolute Value Relation Analysis

**Participants:** Liqian Chen [National Laboratory for Parallel and Distributed Processing, Changsha, P. R. China], Antoine Miné, Ji Wang [National Laboratory for Parallel and Distributed Processing, Changsha, P. R. China], Patrick Cousot.

We present in [15] an abstract domain dealing with linear inequalities involving variables together with their absolute values. It is an extension of the classical linear relation analysis, which permits to deal with some non convex numerical sets. A first nice result states the equivalence between these “linear absolute value inequalities” (AVI) and “interval linear inequalities”, and “extended linear complementary inequalities” (XLCP, pairs of positive solutions whose pairwise components are not both not zero). The key contribution is the extension of the double-description of polyhedra to XLCP solutions, which is then used to define the standard operations on AVI. The method has been implemented, and experiments show interesting results, with reasonable performances with respect to linear relation analysis.

6.8. Probabilistic Analysis

**Participants:** Patrick Cousot, Michaël Monerau.

The abstract interpretation theory has been widely used in the past decades for verifying properties of computer systems. We have introduced a new extension of this well-known framework to the case of probabilistic systems [21].

The probabilistic abstraction framework we propose allows to systematically lift any classical analysis or verification method to the probabilistic setting by separating in the program semantics the probabilistic behavior from the (non-)deterministic behavior. This separation provides new insights for designing novel probabilistic static analyses and verification methods.

We have defined concrete probabilistic semantics and proposed different ways to abstract them. The approach is expressive and effective. The previous techniques for probabilistic analysis are actually abstractions expressible in our framework.

6.9. Safety

**Participants:** Patrick Cousot, Radhia Cousot.

The abstract interpretation design principle has been applied to the design of new forward and backward proof, verification and analysis methods for safety [17]. The safety collecting semantics defining the strongest safety property of programs is first expressed in a constructive fixpoint form. Safety proof and checking/verification methods then immediately follow by fixpoint induction. Static analysis of abstract safety properties such as invariance are constructively designed by fixpoint abstraction (or approximation) to (automatically) infer safety properties.

6.10. Security

**Participants:** Patrick Cousot, Radhia Cousot.
We have developed, episodically since 2007, an abstract interpretation framework for security and program securization that is the transformation of a program into a secured program satisfying security criteria defined by a human or artificial supervisor (this is verification when no transformation is needed). The securization is based on the notion of responsibility analysis determining which choices in the program (inputs, random draws, interrupts, schedules, etc.) can definitely cause or avoid desired or menacing events, or have no control at all on the occurrence of these events. Various securization policies (eager, early or late lazy, etc.) have been identified to prevent or enforce the occurrence of events.

6.11. Shape Analysis

We have extended the XISA (eXtensible Inductive Shape Analysis) framework, in order to better deal with low level coding styles and programming languages, and in order to analyze recursive programs in a context dependent way. We also introduced a classification for semantic memory models.

6.11.1. Abstracting Calling-Context with Shapes

Participants: Bor-Yuh Evan Chang [University of Colorado at Boulder (USA)], Xavier Rival.

Interprocedural program analysis is often performed by computing procedure summaries. While possible, computing adequate summaries is difficult, particularly in the presence of recursive procedures. In [23], we propose a complementary framework for interprocedural analysis based on a direct abstraction of the calling context. Specifically, our approach exploits the inductive structure of a calling context by treating it directly as a stack of activation records. We built an abstraction based on separation logic with inductive definitions. A key element of this abstract domain is the use of parameters to refine the meaning of such call stack summaries and thus express relations across activation records and with the heap. In essence, we define an abstract interpretation-based analysis framework for recursive programs that permits a fluid per call site abstraction of the call stack—much like how shape analyzers enable a fluid per program point abstraction of the heap.

6.11.2. Abstract domains for the analysis of programs manipulating complex data-structures

Participant: Xavier Rival.

We proposed a framework for building abstract domains for the static analysis of programs which manipulate complex data-structures [8]. Our abstract domain is parametric in the choice of a numerical abstract domain to represent properties of numeric memory cells and in the choice of a set of inductive definitions to be used in order to summarize unbounded heap regions. It features standard primitives for the computation of transfer functions, for the inclusion checking and for the computation of widening iterates. We also proposed an extension to handle programs that make use of low-level memory addressing, and proposed an extension of the widening to infer inductive definitions.

6.11.3. Composite abstract domain for the analysis of dynamic structures

Participants: Xavier Rival, Antoine Toubhans.

Reduced product is a general operation to combine abstract domains into more powerful abstract domains, which has been especially used to construct numerical abstract domains. However, until now, it has not been applied to memory structures. We proposed an instance of a reduced product operation, which can be applied on shape abstract domains based on separation logic and on inductive definitions. The advantage of this construction is that it allows to describe more complex heap dynamic data structures without making the design of all abstract operation more complex. In the other hand, it incurs a reduction cost, whenever we need to transport some information from one domain to the other. We showed that optimal reduction cannot be achieved, and identified the main source of complexity of this operation. A prototype implementation was also carried out. This work was done as part of Antoine Toubhans Master internship.

6.12. Static Analysis of Parallel Software

Participant: Antoine Miné.
We present in [22] a static analysis by abstract interpretation to check for run-time errors in parallel C programs. Following our work on ASTRÉE, we focus on embedded critical programs without recursion nor dynamic memory allocation, but extend the analysis to a static set of threads. Our method iterates a slightly modified non-parallel analysis over each thread in turn, until thread interferences stabilize. We prove the soundness of the method with respect to a sequential consistent semantics and a reasonable weakly consistent memory semantics. We then show how to take into account mutual exclusion and thread priorities through partitioning over the scheduler state. We present preliminary experimental results analyzing a real program with our prototype ASTRÉEA (see 5.3) and demonstrate the scalability of our approach.

### 6.13. Termination

**Participants:** Patrick Cousot, Radhia Cousot.

In [17], we have introduced an abstract interpretation for termination. Proof, verification and analysis methods for termination all rely on two induction principles: (1) a variant function or induction on data ensuring progress towards the end and (2) some form of induction on the program structure.

So far, no clear design principle did exist for termination as is the case for safety so that the existing approaches are scattered and largely not comparable with each other.

For (1), we show that this design principle applies equally well to potential and definite termination. The trace-based termination collecting semantics is given a fixpoint definition. Its abstraction yields a fixpoint definition of the best variant function. By further abstraction of this best variant function, we derive the Floyd/Turing termination proof method as well as new static analysis methods to effectively compute approximations of this best variant function.

For (2), we introduce a generalization of the syntactic notion of structural induction (as found in Hoare logic) into a semantic structural induction based on the new semantic concept of inductive trace cover covering execution traces by segments, a new basis for formulating program properties. Its abstractions allow for generalized recursive proof, verification and static analysis methods by induction on both program structure, control, and data. Examples of particular instances include Floyd’s handling of loop cut-points as well as nested loops, Burstall’s intermittent assertion total correctness proof method, and Podelski-Rybalchenko transition invariants.

### 6.14. Theories, Solvers and Static Analysis

**Participants:** Patrick Cousot, Radhia Cousot, Laurent Mauborgne [IMDEA Software (Madrid, Spain)].

In [20], we have introduced a reduced product combining algebraic and logical abstractions to design program correctness verifiers and static analyzers by abstract interpretation. The key new idea is to show that the Nelson-Oppen procedure for combining theories in SMT-solvers computes a reduced product in an observational semantics, so that algebraic and logical abstract interpretations can naturally be combined in a classical way using a reduced product on this observational semantics. The main practical benefit is that reductions can be performed within the logical abstract domains, within the algebraic abstract domains, and also between the logical and the algebraic abstract domains, including the case of abstractions evolving during the analysis.

### 6.15. Underapproximation for Precondition Inference

**Participants:** Patrick Cousot, Radhia Cousot, Francesco Logozzo [Microsoft Research (Redmond, USA)], Manuel Fähndrichh [Microsoft Research (Redmond, USA)].
In the context of program design by contracts, programmers often insert assertions in their code to be optionally checked at runtime, at least during the debugging phase. These assertions would better be given as a precondition of the method/procedure in which they appear. Potential errors would be discovered earlier and, more importantly, the precondition could be used in the context of separate static program analysis as part of the abstract semantics of the code. However in the case of collections (data structures such as arrays, lists, etc) checking both the precondition and the assertions at runtime appears superfluous and costly. So the precondition is often omitted since it is checked anyway at runtime by the assertions. It follows that the static analysis can be much less precise, a fact that can be difficult to understand since “the precondition and assertions are equivalent” (i.e. at runtime, up to the time at which warnings are produced, but not statically) e.g. for separate static analysis. Moreover preconditions are often understood as overapproximations and thus may exclude good runs which is counter-intuitive for programmers. On the contrary, with considering underapproximations [37], [28] which exclude no good run, ensures that if the precondition is violated then a runtime error must definitely be raised later, and if the precondition is not strong enough to catch all errors they will definitely be captured by a later runtime check.

In [19], we define precisely and formally the contract inference problem from intermittent assertions on scalar variables and elements of collections inserted in the code by the programmer. Our definition excludes no good run even when a non-deterministic choice (e.g. an interactive input) could lead to a bad one. We then introduce new abstract interpretation-based methods to automatically infer both the static contract precondition of a method/procedure and the code to check it at runtime on scalar and collection variables. It has been implemented in CCCHECK (formerly CLOUSOT) by Francesco Logozzo and Manuel Fähndrich.

6.16. Verification of spreadsheet programs by abstract interpretation

Participants: Tie Cheng, Xavier Rival.

Spreadsheet tools (Excel, Openoffice) come with powerful languages which can manipulate sheets in various ways. However, no type discipline is enforced, so that the programs may corrupt spreadsheet contents in many ways. We proposed an abstraction to describe sets of valid spreadsheet states, and designed a verifier for invariants expressed in this abstract domain. Our verifier assumes invariants are defined at the head of loops in the programs (as widening operators for the inference of loop invariants). This work was done as part of Tie Cheng Master internship.
6. New Results

6.1. Processor Architecture

Participants: Damien Hardy, Pierre Michaud, Nathanaël Prémillieu, Ricardo Andrés Velasquéz, Luis-Germán García Morales, Bharath Narasimha Swamy, André Seznec.

Our research in computer architecture covers memory hierarchy, branch prediction, superscalar implementation, as well as SMT and multicore issues.

This year, we have also initiated new research directions within the context of the ERC DAL project.

6.1.1. Null block management on the memory hierarchy

Participant: André Seznec.

It has been observed that some applications manipulate large amounts of null data. Moreover these zero data often exhibit high spatial locality. On some applications more than 20% of the data accesses concern null data blocks. To reduce the pressure on main memory, we have proposed a hardware compressed memory that only targets null data blocks, the decoupled zero-compressed memory [27]. Borrowing some ideas from the decoupled sectored cache [20], the decoupled zero-compressed memory, or DZC memory, manages the main memory as a decoupled sectored set-associative cache where null blocks are only represented by a validity bit. Our experiments show that for many applications, the DZC memory allows to artificially enlarge the main memory, i.e. it reduces the effective physical memory size needed to accommodate the working set of an application without excessive page swapping. Moreover, the DZC memory can be associated with a ZCA cache [5] to manage null blocks across the whole memory hierarchy. For some applications, such a management significantly decreases the memory traffic and therefore can significantly improve performance.

This work corresponds to the PhD of Julien Dusser defended in december 2010.

6.1.2. Emerging memory technologies

Participant: André Seznec.

Phase change memory (PCM) technology appears more scalable than DRAM technology. As PCM exhibits access time slightly longer but in the same range as DRAMs, several recent studies have proposed to use PCMs for designing main memory systems. Unfortunately PCM technology suffers from a limited write endurance; typically each memory cell can only be written a large but still limited number of times (10 millions to 1 billion writes are reported for current technology). Research proposals have essentially focused their attention on designing memory systems that will survive the average behavior of conventional applications. However PCM memory systems should be designed to survive worst-case applications, i.e., malicious attacks targeting the physical destruction of the memory through overwriting a limited number of memory cells.

In 2010, we have proposed the first design of a secure PCM-based main memory that would by construction survive overwrite attacks [19]. This secure PCM-based main memory requires a significant read and write extra memory traffic (an extra memory write per 8 demand memory writes) on all applications. Concurrent proposals require even higher extra read and write memory traffic. In collaboration with a research group from IBM, we have proposed a hardware method to detect malicious overwrite attacks on the main memory, thus limiting the memory traffic overhead on non-malicious applications [32].

6.1.3. Microarchitecture exploration of control flow reconvergence

Participants: Nathanaël Prémillieu, André Seznec.
After continuous progress over the past 15 years [18], [17], the accuracy of branch predictors seems to be reaching a plateau. Other techniques to limit control dependency impact are needed. Control flow reconvergence is an interesting property of programs. After a multi-option control-flow instruction (i.e., either a conditional branch or an indirect jump including returns), all the possible paths merge at a given program point: the reconvergence point.

Superscalar processors rely on aggressive branch prediction, out-of-order execution and instruction level parallelism for achieving high performance. Therefore, on a superscalar core, the overall speculative execution after the mispredicted branch is cancelled, leading to a substantial waste of potential performance. However, deep pipelines and out-of-order execution induce that, when a branch misprediction is resolved, instructions following the reconvergence point have already been fetched, decoded and sometimes executed. While some of this executed work has to be cancelled since data dependencies exist, cancelling the control independent work is a waste of resources and performance. We have proposed a new hardware mechanism called SYRANT, SYmmetric Resource Allocation on Not-taken and Taken paths, addressing control flow reconvergence at a reasonable cost. Moreover, as a side contribution of this research we have shown that, for a modest hardware cost, the outcomes of the branches executed on the wrong paths can be used to guide branch prediction on the correct path.

6.1.4. Confidence estimation for the TAGE predictor
Participant: André Seznec.

For the past 15 years, it has been shown that confidence estimation of branch prediction (i.e., estimating the probability of correct or incorrect prediction) can be used for various usages such as fetch gating or throttling for power saving or for controlling resource allocation policies in an SMT processor. In many proposals, using extra hardware and particularly storage tables for branch confidence estimators has been considered as a worthwhile silicon investment.

The TAGE predictor, presented in 2006 [18], is so far considered as the state-of-the-art conditional branch predictor. We have shown that very accurate confidence estimations can be done for the branch predictions realized by the TAGE predictor by simply observing the outputs of the predictor tables. Many confidence estimators proposed in the literature only discriminate between high confidence predictions and low confidence estimations. It has been recently pointed out that a more selective confidence discrimination could be useful. The observation of the outputs of the predictor tables is sufficient to grade the confidence in the branch predictions with a very good granularity. Moreover a slight modification of the predictor automaton allows to discriminate the prediction in three classes, low-confidence (with a misprediction rate in the 30 % range), medium confidence (with a misprediction rate in 8-12% range) and high confidence (with a misprediction rate lower than 1 %) [37].

6.1.5. Improving branch prediction accuracy
Participant: André Seznec.

The TAGE predictor [18] is often considered as state-of-the-art in conditional branch predictors proposed by academy. For the 3rd championship branch prediction, we have further improved its accuracy by augmenting it with small side predictors (Loop predictor, Statistical Corrector Predictor, Immediate Update Mimicker) [34]. This predictor won the conditional branch track of the 3rd championship branch prediction. In order to further argue for real hardware implementation of the TAGE predictor, we have presented several propositions to reduce the complexity of its hardware design, to reduce its energy consumption [36] and further improve branch accuracy. On a hardware implementation of a conditional branch predictor, the predictor tables are updated at retire time. A retired branch normally induces three accesses to the branch predictor tables: read at prediction time, read at retire time and write for the update. We show that in practice, the TAGE predictor accuracy would not be significantly impaired by avoiding a systematic second read of the prediction tables at retire time for correct prediction. Combined with the elimination of silent updates, this significantly reduces the number of accesses to the predictor. Furthermore, we present a technique allowing to implement the TAGE predictor tables as bank-interleaved structures using single-port memory components. This significantly
reduces the silicon footprint of the predictor as well as its energy consumption without significantly impairing its accuracy.

Correctly predicting the indirect branches has become critical with the introduction of object-oriented programming, Java programming as well as with the renewed importance of interpreters. The ITTAGE indirect branch predictor was introduced in [15]. Three versions of the ITTAGE predictor were presented at the indirect branch track at the championship branch prediction by three different teams, and secured the three first places. Our proposition [35] won the championship.

6.1.6. Hardware acceleration of sequential loops

Participant: Pierre Michaud.

In a decade it will be possible to put on a single chip several hundreds of superscalar cores. A simple application of Amdahl’s law shows that it will make sense to dedicate to sequential performance the silicon area and power budget corresponding to that of several tens, or perhaps several hundreds of conventional superscalar cores. This will lead to a sequential accelerator which will be used to accelerate sequential programs and sequential code sections in parallel programs. The question is, what will this sequential accelerator look like? In a previous work, we have proposed a possible solution for implementing a sequential accelerator, which is to implement a superscalar core with a very “aggressive” microarchitecture and design, and to replicate this core and migrate the execution periodically on the replicas to keep the temperature resulting from the high power density under control [11]. However, future sequential accelerators will probably rely on a combination of several techniques, some already known, some yet to be invented.

We have started exploring a new solution for sequential acceleration, the hardware acceleration of dynamic loops, which are periodic sequences of dynamic instructions. A loop accelerator sits beside a conventional superscalar core and is specialized in executing dynamic loops [40]. Dynamic loops are detected and accelerated automatically, without help from the programmer or the compiler. The execution is migrated from the superscalar core to the loop accelerator when a dynamic loop is detected, and back to the superscalar core when a loop exit condition is encountered. Our simulations show that about one third of all the instructions executed by the SPEC CPU2006 benchmark suite belong to dynamic loops with a length of several thousands dynamic instructions, or more. The loop body size is quite diverse, ranging from a few instructions to several hundreds.

We have described a possible loop accelerator microarchitecture that exploits loop properties and avoids the main bottlenecks of conventional superscalar microarchitectures. Our preliminary study demonstrates significant global speedup on some benchmarks, with a local acceleration for loops typically around 2. Our future research on loop acceleration will explore the solution space for obtaining greater performance speedups.

6.1.7. Exploiting confidence in SMT processors

Participants: Pierre Michaud, André Seznec.

Simultaneous multithreading (SMT) [57] processors dynamically share processor resources between multiple threads. The hardware allocates resources to different threads. The resources are either managed explicitly through setting resource limits to each thread or implicitly through placing the desired instruction mix in the resources. In this case, the main resource management tool is the instruction fetch policy which must predict the behavior of each thread (branch mispredictions, long-latency loads, etc.) as it fetches instructions.

We propose the use of Speculative Instruction Window Weighting (SIWW) [25] to bridge the gap between implicit and explicit SMT fetch policies. SIWW estimates for each thread the amount of outstanding work in the processor pipeline. Fetch proceeds for the thread with the least amount of work left. SIWW policies are implicit as fetch proceeds for the thread with the least amount of work left. They are also explicit as maximum resource allocation can also be set. SIWW can use and combine virtually any of the indicators that were previously proposed for guiding the instruction fetch policy (number of in-flight instructions, number of low confidence branches, number of predicted cache misses, etc.). Therefore, SIWW is an approach to design SMT fetch policies, rather than a particular fetch policy.
Targeting fairness and throughput is often contradictory and a SMT scheduling policy often optimizes only one performance metric with the sacrifice of the other metric. Our simulations show that the SIWW fetch policy can achieve at the same time state-of-the-art throughput, state-of-the-art fairness and state-of-the-art harmonic performance mean.

As a side contribution of this study, we have published a study on fairness metrics for SMT processors and multicore [24].

This study was done in collaboration with Hans Vandierendonck from University of Ghent.

### 6.1.8. Analytical model to estimate the interaction between hardware faults on caches and predictors

**Participant:** Damien Hardy.

*This research was undertaken during Damien Hardy’s stay in the Computer Architecture group of the University of Cyprus (June-August 2011).*

Technology trends suggest that in tomorrow’s computing systems, failures will become a commonplace due to many factors, and the expected probability of failure will increase with scaling. Faults can result in execution errors (e.g., on caches) or simply in performance loss (e.g., predictors). Although faults can occur anywhere in the processor, the performance implications of a faulty cell vary depending on how the array is used in a processor.

A direction to determine the impact on performance due to permanent faulty cells is to predict the performance vulnerability by using analytical models. Such models, studied at the University of Cyprus, are representative for the average performance and its probability distribution. So far, analytical models have been defined to determine the impact on performance of faulty mechanisms such as caches and predictors in isolation without any interactions between them.

On the other side, in the real-time systems community, caches and predictors have been intensively studied to estimate the worst-case execution time of application by using static analysis. The ongoing research aims at defining an analytical model of performance that captures the effects of faults on both caches and predictors.

This analytical model will be useful to predict future processors performance vulnerability to faults and to determine the benefits in terms of performance of reliability mechanisms.

### 6.1.9. Hardware support for transactional memory

**Participants:** Mridha-Mohammad Waliullah, André Seznec.

Parallel programming has become immensely important to harness the power of today’s many-core CPU. Over several years, a lot of efforts has been laid out to make parallel programming easier. Transactional memory (TM) has come out as an infrastructure that promises to simplify parallel programming. Implementation of TM in hardware is carried out to get higher performance, which is referred to as hardware transactional memory (HTM). We have focused mainly into two issues related to HTM: (1) exploring TM benchmarks to better understand the performance bottlenecks and (2) exploring innovative techniques that can streamline common case transitional execution to achieve higher performance [38].

This work was done in the framework of the Ercim postdoc stay (01/04/11 to 30/11/11) of Mridha-Mohammad Waliullah.

### 6.1.10. Microarchitecture research initiated in the DAL project

**Participants:** Pierre Michaud, Luis-Germán García Morales, Bharath Narasimha Swamy, André Seznec.
Multicore processors have now become mainstream for both general-purpose and embedded computing. Instead of working on improving the architecture of the next generation multicore, with the DAL project, we deliberately anticipate the next few generations of multicores. While multicores featuring 1000s of cores might become feasible around 2020, there are strong indications that sequential programming style will continue to be dominant. Even future mainstream parallel applications will exhibit large sequential sections. Amdahl’s law indicates that high performance on these sequential sections is needed to enable overall high performance on the whole application. On many (most) applications, the effective performance of future computer systems using a 1000-core processor chip will significantly depend on their performance on both sequential code sections and single threads.

We envision that, around 2020, the processor chips will feature a few complex cores and many (may be 1000’s) simpler, more silicon and power effective cores.

In the DAL research project, we will explore the microarchitecture techniques that will be needed to enable high performance on such heterogeneous processor chips. Very high performance will be required on both sequential sections, -legacy sequential codes, sequential sections of parallel applications-, and critical threads on parallel applications, -e.g. the main thread controlling the application. Our research will focus on enhancing single processes performance.

On the microarchitecture side, we will explore both a radically new approach, the sequential accelerator [11], and more conventional processor architectures. We will also study how to exploit heterogeneous multicore architectures to enhance sequential thread performance. Two PhD thesis have been initiated on these topics at fall 2011.

6.2. Compiler, vectorization, interpretation

Participants: Erven Rohou, David Yuste, André Seznec.

The usage of the bytecote-based languages such as Java has been generalized in the past few years. Applications are now very large and are deployed on many different platforms, since they are highly portable. With the new diversity of multicore platforms, functional, but also performance portability will become the major issue in the next 10 years. Therefore our research effort focuses on efficiently compiling towards bytecodes and on efficiently executing the bytecodes through JIT compilation or through direct interpretations.

6.2.1. Iterative and JIT compilation

Participants: Erven Rohou, David Yuste.

Over the last decade, iterative compilation has been an attempt to overcome the difficulty to generate extremely optimized code by letting the compilers explore many alternatives to select the best one. In this research, we extend previous work in the direction of portability. Future processors will be increasingly diverse and heterogenous, and portability is likely to be attained thanks to a bytecode format and JIT compilers. We explore how iterative compilation performed offline can generate useful information to allow the online JIT compiler to generate efficient code at very limited cost.

Part of this research is done in collaboration with STMicroelectronics, in the context of the Nano2012 Mediacom project.

6.2.2. Split vectorization

Participants: Erven Rohou, David Yuste, André Seznec.

We attempt to reconcile two apparently contradictory trends of computing systems. On the one hand, hardware heterogeneity favors the adoption of bytecode format and late, just-in-time code generation. On the other hand, exploitation of hardware features, in particular SIMD extensions through vectorization, is key to obtaining the required performance.
We showed in [33] that speculatively vectorized bytecode is: (1) robust — the approach is general enough to allow execution, both when using SIMD capabilities and also in the absence of SIMD extensions, or when using an unmodified, non-vectorizing JIT compiler; (2) risk-free — the penalty of running vectorized bytecode without SIMD support is kept at a minimum; (3) efficient — the improvement of running vectorized bytecode with SIMD support is maximized.

In [31], we focused on providing an infrastructure capable of supporting diverse SIMD targets (SSE, AltiVec, NEON), across a wide range of vectorizable kernels, with performance comparable to monolithic compiler vectorization.

This research is done within the framework of the HIPEAC2 network in collaboration with Albert Cohen (INRIA Alchemy), Ayal Zaks and Dorit Nuzman (IBM Research Labs, Haifa, Israel).

6.2.3. Vectorization Technology To Improve Interpreter Performance

Participants: Erven Rohou, David Yuste.

Recent trends in consumer electronics have created a new category of portable, lightweight software applications. Typically, these applications have fast development cycles and short life spans. They run on a wide range of systems and are deployed in a target independent bytecode format over Internet and cellular networks. Their authors are untrusted third-party vendors, and they are executed in secure managed runtimes or virtual machines. Furthermore, due to security policies, these virtual machines are often lacking just-in-time compilers and are reliant on interpreter execution.

The main performance penalty in interpreters arises from instruction dispatch. Each bytecode requires a minimum number of machine instructions to be executed. In this work we introduce a powerful and portable representation that reduces instruction dispatch thanks to vectorization technology. It takes advantage of the vast research in vectorization and its presence in modern compilers. Thanks to a split compilation strategy, our approach exhibits almost no overhead. Complex compiler analyses are performed ahead of time. Their results are encoded on top of the bytecode language, becoming new SIMD IR (i.e., intermediate representation) instructions. The bytecode language remains unmodified, thus this representation is compatible with legacy interpreters.

This approach drastically reduces the number of instructions to interpret and improves execution time. SIMD IR instructions are mapped to hardware SIMD instructions when available, with a substantial improvement.

6.2.4. Tiptop

Participant: Erven Rohou.

Hardware performance monitoring counters have recently received a lot of attention. They have been used by diverse communities to understand and improve the quality of computing systems: for example, architects use them to extract application characteristics and propose new hardware mechanisms; compiler writers study how generated code behaves on particular hardware; software developers identify critical regions of their applications and evaluate design choices to select the best performing implementation.

We propose [41] that counters be used by all categories of users, in particular non-experts, and we advocate that a few simple metrics derived from these counters are relevant and useful. For example, a low IPC (number of executed instructions per cycle) indicates that the hardware is not performing at its best; a high cache miss ratio can suggest several causes, such as conflicts between processes in a multicore environment.

We propose tiptop: a new tool, similar to the UNIX top utility, that requires no special privilege and no modification of applications. Tiptop provides more informative estimates of the actual performance than existing UNIX utilities, and better ease of use than current tools based on performance monitoring counters. With several use cases, we have illustrated possible usages of such a tool.

6.3. Understanding performance issues

Participants: Junjie Lai, Ricardo Andrés Velasquéz, Pierre Michaud, André Seznec.
6.3.1. Behavioral application-dependent superscalar core modeling

**Participants:** Ricardo Andrés Velasquéz, Pierre Michaud, André Seznec.

In recent years, research in microarchitecture has shifted from single-core to multi-core processors. Cycle-accurate models for many-core processors featuring hundreds or even thousands of cores are out of reach for realistic workloads. Approximate simulation methodologies which trade accuracy for simulation speed are necessary for conducting certain research, in particular for studying the impact of resource sharing between cores, where the shared resource can be caches, on-chip network, memory bus, power, temperature, etc.

Behavioral superscalar core modeling is a possible way to trade accuracy for simulation speed in situations where the focus of the study is not the core itself but what is outside the core, i.e., the uncore. In this modeling approach, a superscalar core is viewed as a black box emitting requests to the uncore at certain times. A behavioral core model can be connected to a cycle-accurate uncore model. Behavioral core models are built from detailed simulations. Once the time to build the model is amortized, significant simulation speedups are achieved.

We have proposed a new method for defining behavioral models for modern superscalar cores. Our method, *behavioral application-dependent superscalar core* (BADCO) modeling, requires two traces generated with cycle-accurate simulations. For the first trace, all the requests from the core (which includes the level-1 caches) to the uncore are forced with a null latency, i.e., we simulate a perfect uncore. For the second trace, all the requests are forced with a fixed and very long latency. Then we build a BADCO model from the timing information recorded in these two traces. A BADCO model is basically a directed graph where each node represents a group of micro-ops that may carry some requests to the uncore. Edges in this graph represent dependencies between requests. After the model is built, it can be used for simulations. During simulation, the BADCO model emulates the processor’s reorder buffer, the level-1 miss status holding registers, and honors dependencies between nodes. We have compared BADCO with Lee et al.’s PDCM behavioral core model. BADCO is more accurate than PDCM on average and is more reliable [42]. BADCO predicts the execution time of a thread running on a modern superscalar core with an error typically under 5%. From our experiments, we found that BADCO is qualitatively accurate, being able to predict how performance changes when we change the uncore. The simulation speedups obtained with BADCO are typically greater than 10.

6.3.2. Architecture for Lattice QCD

**Participants:** Junjie Lai, André Seznec.

Simulation of Lattice QCD is a challenging computational problem that requires very high performance exceeding sustained Petaflops/s. In the framework of the ANR Cosinus PetaQCD project, we are modeling the demands of this application on the memory system and synchronization mechanisms.

In particular, GPUs have become popular to execute computing intensive scientific applications. In [39], we have introduced a GPU timing model to provide more insights into the applications’ performance on GPU. A GPU CUDA program timing estimation tool (TEG) is developed based on the GPU timing model. Especially, TEG illustrates how performance scales from one warp (CUDA thread group) to multiple concurrent warps on SM (Streaming Multiprocessor). Because TEG takes the native GPU assembly code as input, it allows to estimate the execution time with only a small error. TEG can help programmers to better understand the performance results. It also allows to identify and quantify performance bottlenecks.

6.4. WCET estimation

**Participants:** Damien Hardy, Benjamin Lesage, Isabelle Puaut, Erven Rohou, André Seznec.

Predicting the amount of resources required by embedded software is of prime importance for verifying that the system will fulfill its real-time and resource constraints. A particularly important point in hard real-time embedded systems is to predict the Worst-Case Execution Times (WCETs) of tasks, so that it can be proven that tasks temporal constraints (typically, deadlines) will be met. Our research concerns methods for obtaining automatically upper bounds of the execution times of applications on a given hardware. Our focus this year is on (i) multicore architectures, (ii) applications compiled using just-in-time (JIT) compilation, and (iii) definition of both predictable and efficient hardware.
6.4.1. Timing analysis for multicore platforms with shared caches

Participants: Benjamin Lesage, Damien Hardy, Isabelle Puaut.

WCET estimation for multicore platforms is challenging task because of the possible interferences between cores due to shared hardware resources such as shared caches, memory bus, etc. Moreover, multi-core platforms use a hierarchy of caches, whose worst-case behavior has to be predicted safely and as tightly as possible.

6.4.1.1. Scalable fixed-point free instruction cache analysis

Estimating worst-case execution times (WCETs) for architectures with caches requires the worst-case number of cache misses to be upper bounded. Most existing static cache analysis methods, a large majority of those applying to unicore and a single cache level, use fixed-point computation and do not scale well with large code sizes.

Estimating WCETs for multicores requires the base cache analysis used to analyze every cache level in the memory hierarchy to be fast. To address this issue, we have proposed in [28] a new fast and scalable instruction cache analysis technique. In contrast to existing work, neither fixed-point computation nor heavyweight interprocedural analysis are required. Thus, code sizes that are too long to analyze with existing techniques becomes then analyzable with lower analysis time and memory consumption, and with only a slight degradation of the analysis precision. Experimental results show a reduction of the analysis execution time of a factor 5 in average (with a peak near 30 for the largest and most complex code) with a limited waste of tightness of the analysis.

6.4.1.2. Static analysis of cache hierarchies

Participants: Benjamin Lesage, Damien Hardy, Isabelle Puaut.

Determining the worst-case number of cache misses in multicore architectures requires the definition of analysis techniques applicable to cache hierarchies. In our previous work [6] we have defined the first technique that safely analyzes such hierarchies, first considering non-inclusive caches and LRU cache replacement.

We have recently generalized our previous work to support different cache hierarchy management policies between cache levels: non-inclusive, inclusive and exclusive cache hierarchies. Moreover, our analysis now supports cache hierarchies with different replacement policies: Least Recently Used (LRU), Pseudo-LRU, First-In First-Out (FIFO), Most Recently Used (MRU) and Random. Experimental results, detailed in [21] show that the method is precise in many cases (non-inclusive and exclusive cache hierarchies with the LRU replacement policy) and has a reasonable computation time. Nevertheless, considering inclusion enforcement mechanisms and non-LRU replacement policies leads to an increase of the analysis pessimism. Moreover, these two sources of pessimism are cumulative, thus resulting, in some cases, in a significant overestimation. Although inclusive cache hierarchies with non-LRU replacement policies can be analyzed statically, the cache hierarchies to be favored to obtain the tighter WCET estimates are hierarchies of non-inclusive or exclusive caches with the LRU replacement policy.

6.4.1.3. Reconciling Predictability and Just-In-Time Compilation

Participants: Isabelle Puaut, Erven Rohou.

Virtualization and just-in-time (JIT) compilation have become important tools to address application portability issues without deteriorating average-case performance. Unfortunately, JIT compilation raises predictability issues, which currently hinders its dissemination in real-time applications. Our work aims at reconciling the two domains, i.e. taking advantage of the portability and performance provided by JIT compilation, while providing predictability guarantees.
As a first step towards this ambitious goal, we have proposed two structures of code caches and have demonstrated their predictability [26]. On the one hand, the binary code caches we propose avoid too frequent function recompilations, providing good average-case performance. On the other hand, and more importantly for the system determinism, we show that the behavior of the code cache is predictable: a safe upper bound of the number of function recompilations can be computed, enabling the verification of timing constraints. Experimental results show that fixing function addresses in the binary cache ahead of time results in tighter Worst Case Execution Times (WCETs) than organizing the binary code cache in fixed-size blocks replaced using a Least Recently Used (LRU) policy.

6.4.1.4. Predictable shared caches for mixed-criticality real-time systems

Participants: Benjamin Lesage, Isabelle Puaut, André Seznec.

The general adoption of multi-core architectures has raised new opportunities as well as new issues in all application domains. In the context of real-time applications, it has created one major opportunity and one major difficulty. On the one hand, the availability of multiple high-performance cores has created the opportunity to run the execution of a complex critical real-time workload and the execution of non-critical applications. On the other hand, for real-time tasks timing deadlines must be met and enforced. Hardware resource sharing inherent to multi-cores hinders the timing analysis of concurrent tasks. Two different objectives are then pursued: enforcing timing deadlines for real-time tasks and achieving highest possible performance for the non-critical workload.

In this work, we suggest a hybrid hardware-based cache partitioning scheme that aims at achieving these two objectives at the same time. Plainly considering inter-task conflicts on shared cache for real-time tasks yields very pessimistic timing estimates. We remove this pessimism by reserving private cache space for real-time tasks. Upon the creation of a real-time task, our scheme reserves a fixed number of cache lines per set for the task. Therefore uniprocessor worst case execution time (WCET) estimation techniques can be used, resulting in tight WCET estimates. Upon the termination of the real-time task, this private cache space is released and made available for all the executed threads including non-critical ones. That is, apart the private spaces reserved for the real-time tasks currently running, the cache space is shared by all tasks running on the processor, i.e. non-critical tasks but also the real-time tasks for their least recently used blocks. Experiments show that the proposed cache scheme allows to both guarantee the schedulability of a set of real-time tasks with tight timing constraints and enable high performance on the non-critical tasks.

6.4.2. Preemption delay analysis for floating non-preemptive region scheduling

Participant: Isabelle Puaut.

This is joint work with Stefan M. Petters, Vincent Nélis and José Marinho, ISEP Porto, Portugal.

In real-time systems, there are two distinct trends for scheduling task sets on unicore systems: non-preemptive and preemptive scheduling. Non-preemptive scheduling is obviously not subject to any preemption delays but its schedulability may be quite poor, whereas fully preemptive scheduling is subject to preemption delays, but benefits from a higher flexibility in the scheduling decisions.

The time-delay involved by task preemptions is a major source of pessimism in the analysis of the task Worst-Case Execution Time (WCET) in real-time systems. Cache related preemption delays (CRPD) are the most important ones, and are caused by the preemptions that modify the cache; the preempted task then suffers an indirect delay after the preemption to reload the cache with useful information.

Preemptive scheduling policies including non-preemptive regions are a hybrid solution between non-preemptive and fully preemptive scheduling paradigms, which enables to conjugate both worlds benefits. In this work [29], we exploit the connection between the progression of a task in its operations, and the knowledge of the preemption delays as a function of its progression. Thus the pessimism in the preemption delay estimation is reduced, in comparison to state of the art methods, due to the increase in information available in the analysis.
5. New Results

5.1. Analysis of Algorithms

The following articles, conference communications and reports summarize new results in analysis of algorithms over the period: [20], [18], [17], [19], [16], [13], [4], [10], [11], [3], [5].

5.2. Computer Algebra

The following articles, conference communications and reports summarize new results in computer algebra over the period: [9], [21], [2], [8], [22], [7], [15], [1], [12], [6], [14].
AOSTE Project-Team

6. New Results

6.1. Logical time in Model-Based embedded design

Participants: Charles André, Frédéric Mallet, Julien Deantoni, Robert de Simone, Marie-Agnès Peraldi Frati, Régis Gascon, Calin Glitia, Kelly Garces Pernett, Benoît Ferrero, Nicolas Chleq, Arda Goknil.

The foundational basis of our approach to modeling and analysis of embedded systems using logical time and logical clock specification contraints (CCSL) is recalled in 3.2, and was surveyed in [2]. This year we conducted a number of works exploiting this approach and promoting its introduction to various application domains.

Charles André presented the general approach in an invited lecture at the French Summer School on Real-Time, in Brest [21].

The HDR manuscript of Frédéric Mallet, where the MARTE Time Model is deeply considered, also in relation with other standards such as AADL, was published in book format [39].

In the article [19] we showed how CCSL observers could be encoded in the synchronous language Esterel, using crucial features of simultaneity, and how otherwise simultaneity could be obtained in simulation. This work was also presented internally as deliverable of the FUI Lambda project (see 8.2.3).

We drew a definite link with our activities on Process Network analysis (see 6.3), by showing how the CCSL primitives could be used to provide the loose timing semantic constraints of exiting PN models such as SDF (Synchronous Data-Flow domain of UC Berkeley’s Ptolemy), and its Multi-Dimensional extension (MD-SDF). This resulted in a journal publication [38]. Existing static schedules can then be obtained by analysis with K-Passa 5.2, or simulated using TimeSquare 5.1 (with an ASAP strategy).

In a collaboration with researchers at East China Normal University (ECNU Shanghai), we showed how CCSL constraints could be translated towards the PROMELA language implemented in the SPIN model-checker, which once again raises the issues of faithfully modeling simultaneity. This work resulted in a communication at the ICECCS conference [33]. Following this work one of our co-author, Yin Ling, earned a one-year scholarship from the Chinese government to visit us as part of her PhD.

The usage of CCSL expressions in the role of predicate property formulas, and their comparison with the more classical temporal formalisms such as PSL (Property Specification Language), was investigated in [24]. A longer internal report version can be found at [42].

In [23] we tackle the issue of recovering global information from multiple execution traces living in distinct logical time bases, with polychronous constraints relating them. The use for efficient debug of embedded systems from distributed traces is exemplified on a case study of terrestrial robot. This work was conducted in the framework of the ANR RT-Simex project, see 8.2.1.

A case study in modeling with logical time and CCSL, from requirements to implementation, based on an automotive spark ignition system, is provided in [31]. We worked more generally on the introduction of our approach to existing formalism in the automotive domain, such as EAST-ADL2 and AutoSar, as part of our contribution to the new ITEA2 Timmo2U project. Premises of this effort are described in [32].

The use of CCSL constraints in general requirement engineering was also studied and demonstrated in a conference article, jointly with colleagues at ECNU Shanghai, presented at APSEC’2011 [22].

The use and modeling of priorities amongst timed events (i.e., logical clock ticks), which has strong impacts on efficient logical clock based simulations and scheduling (as the choice of next event), is still a topic of ongoing work. Several advanced considerations are to be found as part of jean-François Le Tallec PhD thesis, to be defended in January 2012 [16].
6.2. Model-Based approaches to SoC design

Participants: Charles André, Robert de Simone, Benoît Ferrero, Carlos Gomez Cardenas, Jean-François Le Tallec.

The main practical goal of this work was to combine in a sensible way the various formalisms SystemC, IP-XACT, UML MARTE, and UPF (for power consumption representation) (see 4.1 for further descriptions). There were true motivations for this: SystemC is a de-facto standard for SoC simulation at various levels, but currently lacks any sort of formal description so that systems can be analysed, reasoned about for correctness and optimized (and it becomes even more so with newer draft standard evolutions). IP-XACT was introduced as an ADL to ease composition and assembly of IP components (written in SystemC or not), but again it currently fails short of its goal, and in particular does not allow standard decoration of model attributes in prominent non-functional domains such as timing/performance and low-power/energy consumption. These could be provided with the help of dedicated features in UML MARTE, and aligned on the UPF standard for power management modeling.

While the intended design flow would take the UML MARTE and UPF to IP-XACT to SystemC direction, it was important to extract IP-XACT and MARTE structural representation from existing SystemC programs, both to populate the flow with existing legacy models, and to explore better the requirements for complete and consistent modeling towards IP block assembly. This work was conducted in Jean-François Le Tallec PhD, to be defended in January 2012 [16]. Together with Benoît Ferrero he defined and realized a software tool named SCiPX (SystemC to IP-XACT translator), originally based on the PinaVM tool by VERIMAG and the DoxyGen syntactic analyzer.

SCiPX is available in prototype version from our site http://www-sop.inria.fr/aoste/index.php?page=software/scipx. It can be combined with the former software transformation modules IPXACT2Marte and Marte2IPXACT developed previously. These results were partly supported by the ID/TL-M contract with STMicroelectronics (see 7.1), and the ANR HeLP project (see 8.2.2), and were presented in [26], [25]

As part of his PhD thesis, Carlos Gomez Cardenas described a subset of UPF standard as a metamodel inside UML MARTE. He also considered compatibility and interconnections with the industrial environments AcePlorer (by Docea Power), and Synopsys Virtualizer (formerly CoWare), provided to us in the context of the CIM PACA tool farm 8.1.1. Preliminary results were presented in [36]. This work was also presented during internal meetings of the ANR HeLP project, and coordinated with work conducted in the team of Michel Auguin at CNRS UMR LEAT (also in Sophia-Antipolis).

6.3. Process Network analysis


This year we comforted the type of analysis on regular static scheduling and routing in dedicated process network models such as studied in the successive PhD thesis of Julien Boucaron, Jean-Vivien Millo, and Anthony Coadou, and recently surveyed in [5]. This resulted mainly in further implementation upgrades of our K-Passa tool (see 5.2), performed first by Anthony Coadou (before he left on a postdoc internship), then continued by Jean-Vivien Millo (on a return postdoc position with us).

In a work mostly conducted while member of the Alchemy EPI in Saclay, but which draws a clear link to our past and present activities in the subject, Sid-Ahmed-Ali Touati studied efficient heuristics to the general problem of one-dimensional periodic task scheduling under storage requirements, using a modeling framework akin to Process Networks. This resulted in a journal article accepted once the author had become attached to the Aoste EPI [17].

6.4. Correct and efficient implementation of polychronous formalisms

Participants: Thomas Carle, Manel Djemal, Virginia Papailiopoulos, Dumitru Potop Butucaru, Robert de Simone, Yves Sorel.
Existing analysis techniques for synchronous and polychronous languages, such as clock calculi, are meant to extract relations of simultaneity (time inclusion) and exclusiveness (time exclusion) between the various computations and communications. This approach is well-suited when targeting sequential processors. For distributed or multi-threaded implementations, further independence relations are needed to express potential concurrency. This resulted in a general theory of endochronous systems, meant to support this additional analysis [11].

Last year we completed a first prototype tool implementation for weak endochrony checking. This was completed this year in two directions:

- connecting our tool with Signal as input language, and interface it in practice to the Polychrony/SME environment developed by the Espresso EPI;
- improving algorithmic complexity and internal data representation, so that our tool can now handle reasonable size Signal programs.

This work was of course conducted in collaboration with Espresso members. Experimental results were presented at the ESLsyn 2011 conference [30]. We are currently expanding the framework in order to take modes/states into account in the program specifications. Effective generation of multi-threaded GALS wrappers for Signal programs is also under way.

We worked at extending the AAA methodology for polychronous processes by providing a better integration of clock analysis in the various phases of the implementation process (allocation, scheduling, pipelining, etc.). We also considered a wider range of implementation targets (time-triggered, MPSoC). We defined a dedicated software pipelining algorithm to match conditional scheduling/reservation tables such as used in SynDex, with the goal of improving throughput with the same duration of individual computation cycles (as is the goal of any pipelining techniques). The originality here is to make logical clocks of polychronous systems act as triggers for the predicated executions as used in traditional software pipelining. First results have been presented during the Synchron 2011 workshop and in a research report [41].

Further work on time-triggered systems was submitted inside the FUI Parsec 8.2.4 and P 8.2.5 projects, including real-time implementation methods for the IMA/ARINC 653 avionics platforms. In particular we conducted experiments to replace the scheduling policy of the second-level scheduler (L1 in the standard) from dynamic priority-driven to dynamic Time Division time triggered (TT-IMA). Preliminary results are under way, and were informally presented at the yearly Synchron seminar.

An important emerging trend in target MPSoC platforms is that On-chip networks are progressively introduced to cope with the bottleneck of inter processor communications. Correct implementation of polychronous systems in this context thus relies on efficient routing of data in such networks, and ultimately may assume that on-chip NoC routers may be programmed in one way or another to behave predictably according to the global application distributed on the cores. We started a collaboration on this topic with the "Embedded Systems-on-Chips" department of the LIP6 laboratory, one of the main site of expertise for SoC/NoC design and Hardware/software codesign. This collaboration first materialized with the co-supervision of M. Djemal’s PhD thesis. A generic MPSoC architecture is being defined, which includes a 2D mesh network-on-chip with programmable routers, on which static routing schedules such as synthesized by our tools may be implemented and run.

### 6.5. Uniprocessor Real-time Scheduling

**Participants:** Laurent George, Mohamed Marouf, Daniel de Rauglaudre, Yves Sorel.

#### 6.5.1. Strict periodic harmonic tasks

This year, we focused our work on scheduling of strict periodic tasks to the particular case of harmonic tasks [28]. After transforming the scheduling problem into a bin-packing problem, we performed a schedulability analysis and proposed schedulability conditions in each sub-case of harmonic tasks: we proposed a necessary and sufficient condition in the case where all tasks periods are distinct, and we proposed a sufficient condition in the case where some tasks have the same period. Finally, we proposed a scheduling algorithm based on the bin-packing problem resolution.
6.5.2. Combination of strict periodic and sporadic tasks

Non-preemptive strict periodic tasks are harder to schedule than preemptive ones. One can hope to extend schedulability results when combining non-preemptive strict periodic tasks with preemptive sporadic ones.

We proposed in [27] a schedulability analysis for a combination of strict periodic and sporadic tasks. We considered all tasks with fixed priorities, where the highest priorities are given to strict periodic tasks and the lower priorities are given to sporadic tasks. First, we scheduled strict periodic tasks using our former scheduling algorithm. Then, we computed the critical instants which maximize the response time of a sporadic task. We proved that the critical instants are contained in the permanent phase of strict periodic tasks, and are given by the start times of strict periodic jobs in a hyper-period. We also proved that we can reduce critical instants by eliminating some of them. Then, we gave the analytic expression of the computing time $W_i(t)$ at any time $t$ necessary for the execution of a task $\tau_i$ taking in consideration all the tasks with higher priorities. That allows the computation of the response time $r_i$ by solving the equation $W_i(t) = r_i$. Therefore, for a sporadic task, if its response time $r_i$ is less of equal to its deadline for all critical instants, then this latter task is schedulable, else it is not schedulable. We proceed similarly for all sporadic tasks to prove that a tasks set is schedulable or not.

6.5.3. Exact cost of RTOS

It is important to determine the exact cost of the real-time operating system (RTOS) when preemptive scheduling is used for better processor utilization compared to non preemptive scheduling [43]. Indeed, in this case it is possible to trust the schedulability conditions when they include that cost, and also to avoid waste resources. This year we developed a generic RTOS modelled with Petri nets and we determined its exact cost on an ARM9 processor. We used Petri nets on the one hand to choose through simulations the best structure of that scheduler, and on the other hand to verify non blocking properties. In order to obtain its exact cost the scheduler was programmed in assembly language, and coded such as it is deterministic, i.e. its cost does not depend on alternative statements but only on the number of tasks which is known a priori. Using this RTOS we experimented simple task sets on the ARM9 processor for which we were able to include the exact RTOS cost in the schedulability conditions.

6.5.4. Formal proofs of real-time scheduling theorems

Scheduling involves numerous models and theorems, sometimes dated of several decades, but never formally proved. We made a formal proof in Coq (proof assistant developed at Inria) to check a classical theorem giving a schedulability condition for a set of real-time strictly periodic tasks (about 1500 lines of Coq). This work was published in a paper accepted for publication in the conference JFLA 2012.

A second proof is actually being carried now, dealing with response time of a set of fixed priority real-time preemptive tasks. The theorem states that the worst case of this response time occurs when all tasks start simultaneously. A step in the original argument by Jane W. S. Liu [55] involves the proof of a function whose fixpoint computes the response time of the first instance of the least priority task. This specific step is now formally proved in Coq (3500 lines of Coq), and we are now working on the completion of the full theorem.

6.6. Multiprocessor Real-time Scheduling

Participants: Laurent George, Maxence Guesdon, Mohamed Marouf, Falou Ndoye, Simon Nivault, Yves Sorel, Cécile Stentzel.

6.6.1. Partitioned scheduling with exact RTOS cost

In the case of partitioned scheduling we propose a greedy heuristic to solve the real-time scheduling problem of periodic preemptive tasks on a multiprocessor architecture while taking into account the exact RTOS cost. This is achieved by combining an allocation heuristic, of “best fit” type, and a schedulability condition based on the operation $\oplus$ which takes into account the exact RTOS cost [43]. The allocation heuristic minimizes the makespan (total execution time of the tasks executed on the multiprocessor taking into account inter-processor communication costs). A first version of that work was presented in [29].
6.6.2. Semi-partitioned scheduling

In [18] we study two cases of semi-partitioned scheduling of sporadic tasks with constrained deadlines on homogeneous multi-processor: (i) the case where the Worst Case Execution Time (WCET) of a job can be portioned, each portion being executed on a dedicated processor, according to a static pattern of migration and using for solving the local assignment problem a linear programming approach; (ii) the case where the jobs of a task are released on a processor, $1$ time out of $p$, where $p$ is an integer less than or equal to the number of processors, according to a Round-Robin migration pattern. The first approach has been investigated in the state-of-the-art by migrating a job at its local deadline, computed from the deadline of the task it belongs to.

6.6.3. Fault tolerance on electric vehicles

We consider applications composed of a real-time task set running on the distributed heterogeneous architecture of the CyCab (electric vehicle developed in the IMARA team-project) based on dsPICs processors, MPC555 micro-controllers, and an embedded PC all together connected through CAN (Controller Area Network) buses. For hardware reasons we suppose that only dsPICs and CAN buses can fail. Our goal is to find a fault-tolerant software solution to tolerate such failures while the applications satisfy the real-time constraints. Because extra hardware for error detection is expensive in such electric vehicle, we proposed a software error detection based on watchdogs. We solved separately two different problems: buses and dsPICs fault-tolerance. In both cases we use active redundancy policies. For buses fault-tolerance, we assume that all processors are reliable, and all but one bus can fail. The same data is sent through all the CAN buses. If a CAN bus fails then the data is sent by the other CAN buses. For processors fault-tolerance, we assume that all communication media are reliable and at least one processor can fail. The first step consists in performing active redundancy for all the tasks of the application. A task and their redundant tasks are assigned to different processors. If processor fails then the data which are not sent by tasks running on that faulty processor, are actually sent by the redundant tasks. All the tasks with their redundant counterparts are scheduled according to the schedulability analysis proposed in [28].

6.6.4. Scicos/SynDEx gateway and code generation for multi-core

This work was carried out in the Openprod project (see 8.3.2.2). The gateway between Scicos and SynDEx has been updated to deal with the last Scicos data structures and the last version of SynDEx. Besides, this gateway has been improved and partially rewritten to support as much Scicos blocks as possible. We use the gateway to automatically produce from a control model specified and simulated in Scicos a real-time executable running on a multi-core platform. The latter platform is described according to the shared memory model defined last year. In order to generate real-time executable code we had to develop a new SynDEx executive kernel based on Windows-RTX which supports shared memory communications and multi-core parallel execution. That executive kernel is used with the macro-code generated by SynDEx to produce the real-time executable code.

6.6.5. SynDEx updates

We continued the developments of future version 8 of SynDEx which will feature a new software architecture to allow better functionality evolutions and maintenance. On the other hand in the COTROS ADT (“Génération de code temps réel distribué optimisé et sûr”), we achieved the new automatic code generator for the current version 7 of SynDEx. This generator creates intra and inter-processor synchronizations according to well defined rules, checks deadlock absence in inter-processor synchronizations, manages efficiently buffers and semaphores (declaration, naming, etc.).
6. New Results

6.1. Hardware Arithmetic and Architecture

Participants: Florent de Dinechin, Hong Diep Nguyen, Bogdan Pasca, Honoré Takeugming, Álvaro Vázquez Álvarez, Nicolas Brunie, Sylvain Collange.

6.1.1. FPGA-specific arithmetic

Reconfigurable computing has the opportunity of using exotic operators that would not make sense in a general-purpose microprocessor [43], for instance the constant dividers studied in 6.1.2. Such operators must be also be matched to the precision and performance needed by applications. F. de Dinechin and B. Pasca described the FloPoCo framework that assists the construction of correct pipelines and the automatic testing of such operators [28]. For this context, B. Pasca, with H. D. Nguyen, now at U.C. Berkeley, and T. Preusser, from T. U. Darmstadt, described improved architectures for short-latency adders on modern FPGAs [39]. With Ch. Alias and A. Plesco (Compsys project-team), he studied the integration in of deeply pipelined arithmetic datapath in high-level synthesis tools [51].

6.1.2. Multiplication by Rational Constants versus Division by a Constant

Motivated by the division by 3 or by 9 appearing in some stencil kernels, F. de Dinechin investigated how the periodicity of the binary representation of a rational constant could be exploited to design an architecture multiplying by this constant [26]. With L. S. Didier, this approach was then compared to a specialisation of divider architectures to the division by small integer constants, which is shown to match well the fine structure of FPGAs [44].

6.1.3. Elementary Functions

A. Vázquez worked with J. Bruguera, from U. Santiago de Compostella, on hardware architectures for evaluating $q$-th roots [66]. Their solution composes digit-recurrence operators for reciprocal, logarithm, multiplication and exponential.

6.1.4. Extensions of the fused-multiply-and-add operator

With B. de Dinechin, from Kalray, N. Brunie and F. de Dinechin proposed to extend the classical fused-multiply-and-add operator with a larger addend and result. This enables higher-precision computation of sums of products at a cost that remains close to that of the classical FMA [56].

6.1.5. Emerging throughput-oriented architecture

On massively multi-threaded processors like GPUs, neighbor threads are likely to operate on similar data. S. Collange showed with A. Kouyoumdjian how it is possible to take advantage of this inter-thread value correlation at the hardware level with a hardware cache-compression technique on GPUs [59]. With D. Sampaio, R. Martins, and F. Magno Quintão Pereira (U. Minas Gerais), he then addressed this question also at the compiler level using a compiler stage to identify statically data patterns in GPGPU programs [65].

Current GPU architectures require specific instruction sets with control-flow reconvergence annotations, and only support a limited number of control-flow constructs. S. Collange and N. Brunie, with G. Diamos (NVIDIA) generalized dynamic vectorization to arbitrary control flow on standard instruction sets with no compiler involvement [46], [57], [54]. In addition, this technique allows divergent branches to be executed in parallel, as a way to increase the throughput of parallel architectures [55].
6.2. Efficient Floating-Point Arithmetic and Applications


6.2.1. Correctly Rounded Sums

P. Kornerup (Odense Univ., Denmark), V. Lefèvre, N. Louvet and J.-M. Muller have given a study of some basic blocks needed in the design of floating-point summation algorithms. In particular, in radix-2 floating-point arithmetic, they have shown that among the set of the algorithms with no comparisons performing only floating-point additions/subtractions, the 2Sum algorithm introduced by Knuth is minimal, both in terms of number of operations and depth of the dependency graph. They have investigated the possible use of another algorithm, Dekker’s Fast2Sum algorithm, in radix-10 arithmetic. Under reasonable conditions, they have also proven that no algorithms performing only round-to-nearest additions/subtractions exist to compute the round-to-nearest sum of at least three floating-point numbers. Starting from an algorithm due to Boldo and Melquiond, they have also presented new results about the computation of the correctly-rounded sum of three floating-point numbers [21].

6.2.2. Error of an FMA

The fused multiply-add (FMA) instruction, specified by the IEEE 754-2008 Standard for Floating-Point Arithmetic, eases some calculations, and is already available on some current processors such as the Power PC or the Itanium. S. Boldo (EPI Proval) and J.-M. Muller first extended an earlier work on the computation of the exact error of an FMA (by giving more general conditions and providing a formal proof). Then, they presented a new algorithm that computes an approximation to the error of an FMA, and provide error bounds and a formal proof for that algorithm [16].

6.2.3. Accurate computation of $ad - bc$ with an FMA

C.-P. Jeannerod, N. Louvet, and J.-M. Muller have provided in [60] a detailed rounding error analysis of Kahan’s FMA-based algorithm for the computation of expressions of the form $ad - bc$. They showed that Kahan’s algorithm is always highly accurate, and under mild assumptions on the radix and the precision gave an optimal bound on the absolute error and an asymptotically optimal bound on the relative error. They also studied how the relative error varies as a function of the relative order of magnitude of the two products $ad$ and $bc$. Finally, they investigated whether the error bounds can be improved in special cases like sums of squares and discriminants.

6.2.4. Performing Arithmetic Operations on Round-to-Nearest Operations

During any composite computation, there is a constant need for rounding intermediate results before they can participate in further processing. Recently, a class of number representations denoted RN-Codings were introduced, allowing an unbiased rounding-to-nearest to take place by a simple truncation, with the property that problems with double-roundings are avoided. P. Kornerup (Odense Univ., Denmark), J.-M. Muller and A. Panhaleux first investigate a particular encoding of the binary representation. This encoding is generalized to any radix and digit set; however, radix complement representations for even values of the radix turn out to be particularly feasible. The encoding is essentially an ordinary radix complement representation with an appended round-bit, but still allowing rounding-to-nearest by truncation, and thus avoiding problems with double-roundings. Conversions from radix complement to these round-to-nearest representations can be performed in constant time, whereas conversion the other way, in general, takes at least logarithmic time. Not only is rounding-to-nearest a constant time operation, but so is also sign inversion, both of which are at best log-time operations on ordinary two’s complement representations. Addition and multiplication on such fixed-point representations are first analyzed and defined in such a way that rounding information can be carried along in a meaningful way, at minimal cost. The analysis is carried through for a compact (canonical) encoding using two’s complement representation, supplied with a round-bit. Based on the fixed-point encoding, it is shown possible to define floating-point representations, and a sketch of the implementation of an FPU is presented [22].
6.2.5. Augmented Precision Square Roots, 2-D Norms, and Discussion on Correctly Rounding
\[ \sqrt{x^2 + y^2} \]

Define an “augmented precision” algorithm as an algorithm that returns, in precision-\( p \) floating-point arithmetic, its result as the unevaluated sum of two floating-point numbers, with a relative error of the order of \( 2^{-2p} \). Assuming an FMA instruction is available, N. Brisebarre, M. Joldeş, P. Kornerup (Odense University, Denmark), E. Martin-Dorel and J.-M. Muller perform a tight error analysis of an augmented precision algorithm for the square root, and introduce two slightly different augmented precision algorithms for the 2D-norm \( \sqrt{x^2 + y^2} \). Then they give tight lower bounds on the minimum distance (in ulps) between \( \sqrt{x^2 + y^2} \) and a midpoint when \( \sqrt{x^2 + y^2} \) is not itself a midpoint. This allows them to determine cases when their algorithms make it possible to return correctly-rounded 2D-norms [30].

6.2.6. Midpoints and exact points of some algebraic functions in floating-point arithmetic

When implementing a function \( f \) in floating-point arithmetic, if we wish correct rounding and good performance, it is important to know if there are input floating-point values \( x \) such that \( f(x) \) is either the middle of two consecutive floating-point numbers (assuming rounded-to-nearest arithmetic), or a floating-point number (assuming rounded toward \( \pm \infty \) or toward 0 arithmetic). In the first case \( f(x) \) is a midpoint, and in the second case it is an exact point. In [20] C.-P. Jeannerod, N. Louvet, J.-M. Muller, and A. Panhaleux have studied whether such midpoints and exact points exist for some usual algebraic functions and various floating-point formats. When midpoints or exact points exist, they have been characterized or, when possible, listed exhaustively. The results and the techniques presented in this paper can be used in particular to deal with both the binary and the decimal formats defined in the IEEE 754-2008 standard for floating-point arithmetic.

6.3. Correct Rounding of Elementary Functions

Participants: Florent de Dinechin, Vincent Lefèvre, Jean-Michel Muller, Bogdan Pasca, Serge Torres.

6.3.1. FPGA Acceleration of the Search For Hardest-to-Round Cases

The IEEE 754-2008 standard for floating-point arithmetic recommends (yet does not dictate) that some elementary functions should be correctly rounded. That is, given a rounding function \( \circ \), (e.g., round to nearest even, or round to \( \pm \infty \)), when evaluating function \( f \) at the floating-point number \( x \), the system should always return \( \circ(f(x)) \).

Building a fast correctly rounded library for some target floating-point (FP) format requires preliminarily solving a problem called the table maker’s dilemma. This requires very large computations which may use environments and formats totally different from the target environment and format. F. de Dinechin, V. Lefèvre, J.-M. Muller, B. Pasca and A. Plesco suggest performing these computations on an FPGA. Their paper [45] won the best paper award at the ASAP2011 conference.

6.3.2. Hierarchical Polynomial Approximation of a Function by Polynomials

Algorithms used to search for the hardest-to-round cases of a function requires the approximation of the function by small-degree polynomials on small intervals. This can be done efficiently by a hierarchical polynomial approximation. Work is being done to improve this method by replacing interval arithmetic (as partly used in the current tools) by static error bounds. This will allow us to better control the precision needed to compute the coefficient of the polynomials. The implementation will also be simpler.

6.4. Validation and Automation

6.4.1. Efficient Implementation of Algorithms for Interval Linear Algebra

H.-D. Nguyen and N. Revol proposed an algorithm to solve linear systems with interval coefficients. The same approach can be used to verify the solution of a linear system with floating-point coefficients, i.e. to compute an interval enclosing the error between the exact solution and an approximate solution. The goal is twofold: on the one hand the accuracy of the solution is desired up to the last bit of the floating-point solution, on the other hand the efficiency of the implementation is obtained through the use of optimized BLAS3 routines [48]. The PhD thesis of H.-D. Nguyen [13] contains in particular the algorithm [24] and its properties. Its complexity has been established [47] and its potential use for symbolic-numeric computations has been discussed [50].

6.4.2. Standardization of Interval Arithmetic

We contributed to the creation, and now chair, the IEEE 1788 working group on the standardization of interval arithmetic http://grouper.ieee.org/groups/1788/. The main discussion topics of this working group [49], for the year 2011, were exception handling (via decorations). An emerging topic is the repetitability and reproducibility of interval computations, on the same platform or across different platforms.

6.4.3. Formal Proofs of the Arithmetic on Polynomial Models

Using as starting point [9], the calculus with polynomial models, such as Taylor models, based on floating-point coefficients and floating-point operations, has been formalized and checked in Coq [18]. This calculus is at the core of Ariadne, an environment for the study of hybrid systems: the idea is to prove the environment itself, instead of using model-checking on the systems.

6.4.4. Formal Proof Generation for Elementary functions

The proof of the correct rounding property for an elementary function requires tight bounds on the error involved in the function code. F. de Dinechin, with Ch. Lauter (LIP6) and G. Melquiond (INRIA Proval) have described the use of the Gappa proof assistant to compute such tight bounds rigorously [27].

6.4.5. Code Generation for Polynomial Evaluation

A given arithmetic expression may be evaluated on a computer in several ways, depending on the parenthesization and the ordering of terms in use. Among all the possible evaluations, one may want to choose one that is as fast and accurate as possible. In [12] Ch. Mouilleron introduced a set of algorithms in order to generate all these possible evaluations, to count them, and to find an optimal or nearly optimal one according to a given criteria. Thanks to this work, several sequences related to numbers of evaluations have been discovered and added to Sloane’s on-line encyclopedia of integer sequences (OEIS). Moreover, this allowed to show experimentally that an algorithm by Paterson and Stockmeyer for the evaluation of a polynomial $p$ at a matrix point is optimal for small degrees of $p$. Finally, this work has led to the revamping of the software tool CGPE presented in [38] (see also §5.5).

6.5. Arithmetic and Algorithms

Participants: Guillaume Hanrot, Claude-Pierre Jeannerod, Adeline Langlois, Ivan Morel, Christophe Mouilleron, Andrew Novocin, Xavier Pujol, Damien Stehlé, Gilles Villard.

6.5.1. Faster Lattice Reduction

Andrew Novocin, Damien Stehlé and Gilles Villard [40] designed an algorithm, $\tilde{L}^1$, with the following specifications: It takes as input an arbitrary basis $B$ in $\mathbb{Z}^{d\times d}$ of a lattice $L$; it computes a basis of $L$ which is reduced for a mild modification of the Lenstra-Lenstra-Lovász reduction; it terminates in time $O(d^2\beta + d^{\omega+1}\beta)$ where $\beta = \log \|B\|$ (and $\omega$ is a valid exponent for matrix multiplication). This is the first LLL-reducing algorithm with a time complexity that is quasi-linear in the bit-length beta of the entries and polynomial in the dimension $d$. A critical ingredient for achieving this result was the study of the effect of small perturbations on the LLL-reducedness of a lattice basis [17].
6.5.2. Computing Short Lattice Vectors

Among all known lattice reduction algorithms, BKZ provides the best trade-off between run-time and smallness of the computed lattice basis. Guillaume Hanrot, Xavier Pujol and Damien Stehlé [32] showed that BKZ can be terminated long before its completion, while still providing bases of excellent quality. More precisely, if it is terminated within a polynomial number of calls to a lower-dimensional Shortest Vector Problem solver, then the bounds on the output quality are as close as desired to the bounds that can be obtained by letting BKZ run until completion.

Guillaume Hanrot, Xavier Pujol and Damien Stehlé also surveyed the known algorithms for solving the Shortest Vector Problem [31].

6.5.3. Lattice-Based Cryptography

NTRUEncrypt is the fastest known lattice-based encryption scheme. Its moderate key-sizes, excellent asymptotic performance and conjectured resistance to quantum computers could make it a desirable alternative to factorisation and discrete-log based encryption schemes. Damien Stehlé and Ron Steinfeld [41] showed how to modify NTRUEncrypt to make it provably resistance to Chosen Plaintext Attacks, under the assumed quantum hardness of standard worst-case lattice problems restricted to a family of lattices related to some cyclotomic fields.

6.5.4. Lattices and Communication Theory

Cong Ling, Shuiyin Liu, Laura Luzzi and Damien Stehlé studied and optimized lattice algorithms that are relevant for MIMO communications [23], [37]. These algorithms tackle the Bounded Distance Decoding Problem: Given a point within a small prescribed distance to a given lattice, find the lattice vector closest to it.

6.5.5. Other Applications of Lattice Reduction Algorithms

In [35] Jürgen Klüners, Mark van Hoeij, and Andrew Novocin showed how to use the LLL lattice reduction algorithm for computing a compact representation of the set of all subfields of any given number field. William Hart (Warwick Mathematics Institute, UK), Mark van Hoeij (Florida State University, USA) and Andrew Novocin exploited the very latest progress in lattice reduction to propose a fine-tuned cutting-edge implementation of a polynomial factorization algorithm.

6.5.6. Polynomial Arithmetic

With William Hart and Mark van Hoeij, A. Novocin proposed in [33] a state of the art algorithm for factoring polynomials in $\mathbb{Z}[x]$. The algorithm is fast in practice, saving in a large class of common examples, without sacrificing performance on worst-case polynomials. The presented algorithm is structured along the lines of algorithms with the best theoretical complexity. In [34] William Hart and A. Novocin proposed an efficient algorithm for computing the composition of two univariate polynomials. Their work builds upon the Brent-Kung algorithm.

6.5.7. Exact Linear Algebra

Transforming a matrix over a field to echelon form, or decomposing the matrix as a product of structured matrices that reveal the rank profile, is a fundamental building block of computational exact linear algebra. For such tasks the best algorithms available so far were either rank sensitive (i.e., of complexity expressed in terms of the exponent of matrix multiplication and the rank of the input matrix) or in place (i.e., using essentially no more memory that what is needed for matrix multiplication). In [61] C.-P. Jeannerod, Clément Pernet (U. Joseph Fourier, Grenoble), and Arne Storjohann (U. Waterloo, Canada) have proposed algorithms that are both rank sensitive and in place. These algorithms are based on a new matrix factorization, namely $A = CUP$ with $C$ a column echelon form revealing the row rank profile of $A$, $U$ a unit upper triangular matrix, and $P$ a permutation matrix.
6. New Results

6.1. Ambiguity Detection in Context-free Grammars

The work on static detection of ambiguity in context-free grammars continued in 2011. Bas Basten has worked on scaling previous results to so-called character level grammars. This includes the application of declarative disambiguation filters that increase the efficiency as well as the accuracy of the analysis [14].

6.2. Automated Diagnostics for Ambiguity in Context-free Grammars

When an ambiguity is found this is reported by a complex trace (usually a set of parse trees). It is difficult for a human to spot the cause of the ambiguity and devise a fitting solution. The Dr Ambiguity algorithm is an expert tool, by Bas Basten and Jurgen Vinju, that compares different parse trees for the same sentence on essential attributes that can be distinguished by declarative disambiguation methods [15].

6.3. A general library for software visualization

The Rascal standard library was extended with a very flexible component for the rapid construction of new/experimental (inter-active) visualizations. The current library is fully working and forms the inspiration for a possible domain-specific extension of the Rascal meta programming language to be integrated at a later stage [23]. This is work by Paul Klint and Atze van der Ploeg.

6.4. Comparing Design Patterns - the case of Visitor versus Interpreter

In this research application of Rascal, Mark Hills, Tijs van der Storm, Paul Klint and Jurgen Vinju focused on analyzing the emergent differences in quality when choosing between two different source code design patterns. We constructed a refactoring tool that translates instances of the Visitor design pattern to the Interpreter design pattern in a semi-automated fashion. This then allowed us to study two versions of an otherwise equivalent system in terms of efficiency and maintainability [24].

6.5. Entry in the language workbench competition

Tijs van der Storm and Jurgen Vinju participated on the Language Workbench Competition 2011 (LWC’11), showcasing the DSL construction and capabilities of Rascal. The objective was the development of a number of DSLs for entity-relation modeling. The modular implementation of these languages was documented in a technical report [35] The DSL implementation featured modular context-free grammars for parsing, modular type checkers, modular code generators and full-fledged IDEs (syntax highlighting, outlining, error marking, etc.). The complete implementation only takes around 700 lines of Rascal code.

6.6. A compiler for OBERON-0

Tijs van der Storm led the participation of ATEAMS in the LDTA Tool Challenge 2011. This was a collaborative effort together with Atze van der Ploeg, Mark Hills, Bas Basten, Paul Klint, Bert Lisser, Jeroen van den Bos, Jurgen Vinju and Arnold Lankamp. The objective of the challenge was to implement all aspects of a simple, imperative language called Oberon-0. The components that had to be implemented included: parsing, pretty printing, name analysis, type checking, and compilation to C. Additionally, the components should be developed in a modular fashion according to four language levels: each language level added more language features and required the modular extension of the components. Although not required for the challenge, we also developed a compiler targeting Java, a compiler targeting JVM byte code, a control-flow graph visualizer, and IDE support. The implementation required only around 4200 lines of Rascal code. The result was presented at the international workshop on Language Descriptions Tools and Applications.
6.7. Generalized Grammar Recovery

Vadim Zaytsev managed to generalize the algorithm for recovering context-free grammars from legacy language documentation. This facilitates the recovery of a lot more grammars to be used in the study of grammarware and software language engineering.

6.8. Comparing Context-free Grammars through test data

Equivalence of the languages generated by two different context-free grammars is undecidable, yet language equivalence is one of the most important quality aspects of context-free grammars for programming languages. Vadim Zaytsev introduced and experimentally validated a method based on differential analysis: generating sentences from both grammars and cross-testing them.
6. New Results

6.1. Dynamically and Heterogeneous Reconfigurable Platforms

6.1.1. New Reconfigurable Architectures

6.1.1.1. Power models of reconfigurable architectures

Participants: Robin Bonamy, Daniel Chillet, Olivier Sentieys.

Including a reconfigurable area in complex systems-on-chip is now considered as an interesting solution to reduce the area of the global system and to support high performances. But the key challenge in the context of embedded systems is currently the power budget of the system, and the designer needs some early estimations of the power consumption of its system. Power estimation for reconfigurable systems is a difficult problem because several parameters need to be taken into account to define an accurate model.

Hardware implementation of an algorithm provides different choices to the designer compared to software implementation. It is possible to vary the parallelism level or loop unrolling index, which has a direct impact on area and execution time and therefore on power and energy consumption. First we evaluated delay, area, power and energy impacts of loop transformations using High Level Synthesis tools. We have made several power measurements on a real FPGA platform and for different task implementations in order to build a model of energy consumption versus execution time. Work is in progress to also characterize energy consumption of tasks through extracting the number of elementary operators used in the hardware implemented task.

Furthermore, we also consider the opportunity of the dynamic reconfiguration, which makes possible to partially reconfigure a specific part of the circuit while the rest of the system is running. This opportunity has two main effects on power consumption. First, thanks to the area sharing ability, the global size of the device can be reduced and the static (leakage) power consumption can thus be reduced. Secondly, it is possible to delete the configuration of a part of the device which reduces the dynamic power consumption when a task is no longer used. Although the cost of the reconfiguration is still important, in some cases this technique can be interesting to reduce the power of the system. To evaluate the potential gain of the dynamic reconfiguration, we have made some measurements on a Virtex 5 board. We have defined a first model of the power consumption of the reconfiguration. This model shows that the power consumption not only depends on the bitstream file size but also on the content of the reconfiguration region [41], [42].

These experiments allow us to define energy and delay models that will be used by the operating system including a power management strategy to decide on-line which task instances must be executed to efficiently manage the available power using dynamic partial reconfiguration [82].

6.1.1.2. High-level modeling of reconfigurable architectures

Participants: Robin Bonamy, Daniel Chillet, Sébastien Pillement.

To help System-on-Chip designers to explore the large design space, high-level methodologies and tools are more and more often required. The exploration phase is particularly difficult when the system must satisfy a large number of constraints, like performance, real time and power consumption. If the classical multiprocessor system-on-chips can be modeled without any difficulty, dynamically reconfigurable embedded accelerators are not correctly covered by the usual modeling languages.

In this context, we have extended the AADL (Architecture Analysis and Design Language) language to include the reconfiguration aspect included in nowadays’ MPSoC [19], [40]. This work is part of a more general project, Open-People, which proposes complete methodology for power and energy consumption analysis. The proposal is based on AADL property extensions which are applied on component models. A three-level model has been defined for every targeted FPGA. The first level defines a generic FPGA which allows to model every possible FPGA. The second level allows the specialization of the FPGA for a specific family. Finally, the third level provides the support to describe the deployment of an application on a specific FPGA circuit.
To complete these levels of description, we started the development of techniques for constraint verifications. These developments are based on the OCL language, which allows to extract characteristics on the AADL model, compute mathematical expressions and finally verify mathematical constraints. These verifications have been developed for power and energy consumption, they include static and dynamic power estimation and soon the power consumption during the dynamic reconfiguration process.

6.1.1.3. Reconfiguration controller

Participants: Manh Pham, Daniel Chillet, Sébastien Pillement.

Dynamically reconfigurable architectures, which can offer high performance, are increasingly used in different domains. Unfortunately, lots of applications cannot benefit from this new paradigm due to large timing overhead. Even for partial reconfiguration, modifying a small region of an FPGA takes few ms. To cope with this problem we have developed an ultra-fast power-aware reconfiguration controller (UPaRC) to boost the reconfiguration throughput up to 1.433 GB/s. UPaRC can not only enhance the system performance, but also auto-adapt to various performance and consumption conditions. This could enlarge the range of supported applications and can optimize power-timing trade-off of reconfiguration phase for each selected application during run-time. The energy-efficiency of UPaRC over state-of-the-art reconfiguration controllers is up to 45 times more efficient. This work has been accepted for publication in DATE’2012 [56].

6.1.2. Management of Dynamically Reconfigurable Systems

6.1.2.1. Spatio-Temporal Scheduling based on Artificial Neural Networks

Participants: Antoine Eiche, Daniel Chillet, Sébastien Pillement, Olivier Sentieys.

Management of task execution on dynamically reconfigurable accelerators is known to be a difficult problem due to the large number of possibilities of task instantiations. The problem to solve can be defined as the “spatio-temporal task scheduling”. The problem becomes even more difficult to solve when the solutions must be produced during the execution of the application, i.e. on-line. In this context, new algorithms must be defined and, to solve this problem, we propose to define a neural network based on the Hopfield model. We are therefore able to address heterogeneous multiprocessor systems and to manage the reconfigurable resources embedded within MPSoC [23]. Our latest works on this topic focused on two different issues. First we demonstrated that neural network structures used for task scheduling can continue to produce valid solutions even if one or several neurons are in fault [70]. This characteristic is very important for present and future technologies for which the fabrication process variability can lead to increase the number of defaults in the circuit. The second focus concerned the optimization of the neural network convergence by using parallel evaluation of neurons. We have shown how to define several neuron packets (from the neural network) that can be evaluated in parallel without modifying the convergence property [44], [76].

6.1.2.2. Flexible Communication OS Service

Participants: Daniel Chillet, Sébastien Pillement, Ludovic Devaux.

In a multiprocessor system, to gain the advantages of parallelism, efficient communication and memory management are highly required. Recent developments in the partial and dynamic reconfigurable computing domain demand better ways to manage simultaneous task execution. But, the requirements are slightly different from the traditional software based systems. In this context, Operating System (OS) services like scheduling, placement, inter-task communication have been developed to make such platforms flexible and self-sufficient. For task communications within flexible architectures, we defined a specific network-on-chip adapted to dynamically and partially reconfigurable resources included into modern SoC. The characterization of the DRAFT network was completed and its integration inside reconfigurable systems on chip was realized [14]. We then focused on the run-time communication service [50] and dynamic memory management [49] in reconfigurable System-on-Chips (RSoCs). We first developed a hardware communication block and the communication schemes supported by this new OS service. The originality relies on the implementation of this services directly inside the FPGA. We then demonstrated the requirements and advantages of having a local memory task or a dynamically configurable memory task, in order to improve effectiveness and efficiency of the proposed schemes.
6.1.3. Fault-Tolerant Reconfigurable Architectures

Participants: Sébastien Pillement, Manh Pham, Stanislaw Piestrak.

In terms of complex systems implementation, reconfigurable FPGA circuits are now part of the mainstream thanks to their flexibility, performances and high number of integrated resources. FPGAs enter new fields of applications such as aeronautics, military, automotive or confined control thanks to their ability to be remotely updated. However, these fields of applications correspond to harsh environments (cosmic radiation, ionizing, electromagnetic noise) and with high fault-tolerance requirements. We then propose a complete framework to design reconfigurable architecture supporting fault-tolerance mitigation scheme [58]. The proposed framework allows simulation, validation of mitigation operations, but also to size architecture resources. The physical implementation of the fault-tolerant reconfigurable platform permits to validate the proposed model and the effectiveness of the framework. This implementation shows the potential of dynamically reconfigurable architectures for supporting fault-tolerance in embedded systems. We also worked on new approach in order to include dependability in the DRAFT coarse-grained reconfigurable architecture [37].

6.1.4. Low-Power Architectures

6.1.4.1. Ultra Low-Power Architecture for Control-Oriented Applications in Wireless Sensor Nodes

Participants: Olivier Sentieys, Steven Derrien, Vivek D. Tovinakere, Philippe Quémerais, Romain Fontaine.

This research work aims at developing ultra low-power SoC for wireless sensor nodes, as an alternative to existing approaches based low-power micro-controllers. The proposed approach reduces the power consumption by using a combination of hardware specialization and power gating techniques. In particular, we use the fact that typical WSN applications are generally modeled as a set of small to medium grain tasks that are implemented on low power microcontroller using light weight thread-like OS constructs.

Rather than implementing these tasks in software, we instead propose to map each of these tasks to their own specialized hardware structures that we call a hardware micro-task. Such hardware task consists of a minimalistic (and customized) data-path controlled by a finite state machine (FSM). By customizing each of these hardware implementations to their corresponding task, we expect to significantly reduce the dynamic power dissipated by the whole system. Besides, to circumvent the increase in static power caused by the possibly numerous hardware tasks implemented in the chip, we also propose to combine our approach with power gating, so as to supply power to a hardware task only when it needs to be executed [28].

As a proof of concept, a chip has been designed and fabricated in a 65nm CMOS from STMicroelectronics using the CMP facilities. The area is about 1mm² in a QFN52 package. The circuit is a controller part of a wireless sensor network node. It embeds an OpenMSP microcontroller core with SRAM memories for data and programs and some dedicated hardware tasks to control an external radio transceiver such as the TI CC2420 commonly used in the industry.

To reduce energy consumption, low power design techniques such as power gating were used. Two power domains are implemented: one is dedicated to microcontroller and memories, while the goal of the second is to measure the efficiency of our hardware micro-task concept.

The input-output ring around the core is divided into three parts: two parts are digital I/O pads corresponding to a power domain and the third contains analog pads to control the power gating for monitoring. Our goal is to analyze the power benefits of our approach and to compare it with classical microprocessor architectures.

6.1.4.2. Wakeup Time and Wakeup Energy Estimation in Power-Gated Logic Clusters

Participants: Olivier Sentieys, Vivek D. Tovinakere.

Run-time power gating for aggressive leakage reduction has brought into focus the cost of mode transition overheads due to frequent switching between sleep and active modes of circuit operation. In order to design circuits for effective power gating, logic circuits must be characterized for overheads they present during mode transitions. We have proposed a method to determine steady-state virtual-supply voltage in active mode and hence present a model for virtual-supply voltage in terms of basic circuit parameters. Further, we derived
expressions for estimation of two mode transition overheads: wakeup time and wakeup energy for a power-gated logic cluster using the proposed model. Experimental results of application of the model to ISCAS85 benchmark circuits show that wakeup time may be estimated within an average error of 16.3% across 22× variation in sleep transistor sizes and 13× variation in circuit sizes with significant speedup in computation time compared to SPICE level circuit simulations [30], [63].

6.1.5. Arithmetic Operators for Cryptography

Participants: Arnaud Tisserand, Thomas Chabrier, Danuta Pamula.

6.1.5.1. ECC Processor with Protections Against SCA

A dedicated processor for elliptic curve cryptography (ECC) is under development. Functional units for arithmetic operations in $\mathbb{F}_{2^m}$ and $\mathbb{F}_p$ finite fields and 160–600-bit operands have been developed for FPGA implementation. Several protection methods against side channel attacks (SCA) have been studied. The use of some number systems, especially very redundant ones, allows to change the way some computations are performed and then their effects on side channel traces.

We propose in [83] hardware implementation of the double base number system (DBNS) random recoding of secret keys. This recoding is performed on-the-fly during the elliptic curve cryptography (ECC) scalar multiplication $[k]P$. This leads random behavior of the point operations at the side channel level.

We started a collaboration with the University of Sfax in Tunisia on the use of ECC processor for secure communications in low-cost wireless applications. A first FPGA implementation is under development and we expect to submit our first results in 2012.

6.1.5.2. Arithmetic Operators for High-Performance Cryptography

In [32], we published an extended version of the work started in 2010 on fast algorithms and implementations of $\mathbb{F}_{2^m}$ finite field multiplication units in FPGA. The proposed and compared methods are based on separated multiplication and reduction steps and analyzed various area and time dependency/efficiency/complexity tradeoffs.

With Mark Hamilton, PhD student in the Code and Crypto group from the University College Cork (UCC), we worked on fast algorithms and implementations of $\mathbb{F}_p$ finite field multipliers for some specific values of $p$. The corresponding results have been published in [46].

6.1.6. SoC Modeling and Prototyping on FPGA-based Systems

Participants: François Charot, Kevin Martin, Laurent Perraudieu, Charles Wagner.
SoCLib and MutekH are two software development projects to which we contribute. SoCLib (http://www.soclib.fr) is an open platform for modeling and simulation of multiprocessors system-on-chip (MP-SoC). MutekH (http://www.mutekh.org) is a free and portable operating system for embedded platforms, ranging from micro-controller to multiprocessor systems. The use of the configurable and extensible simulation model of the Altera NIOSII processor of the SoCLib library and of the MutekH operating system allows us to easily deploy software applications such as codes from MediaBench, MiBench and Cryptographic Library benchmark sets or multithreaded applications on monoprocessor and multiprocessor simulation platforms. These platforms are used on the one hand for the validation of the processor extensions automatically generated by our compilation tools and on the other hand for the measurement of the speedup achieved using these new extensions.

6.2. Compilation and Synthesis for Reconfigurable Platform

Participants: Steven Derrien, Emmanuel Casseau, Daniel Menard, François Charot, Christophe Wolinski, Olivier Sentieys, Patrice Quinton.

6.2.1. Polyhedral based loop transformations for High-Level synthesis

Participants: Steven Derrien, Antoine Morvan, Patrice Quinton.

After almost two decades of research effort, there now exists a large choice of robust and mature C to hardware tools that are used as production tools by world-class chip vendor companies. Although these tools dramatically slash design time, their ability to generate efficient accelerators is still limited, and they rely on the designer to expose parallelism and to use appropriate data layout in the source program. We believe this can be overcome by tackling the problem directly at the source level, using source-to-source optimizing compilers. More precisely, our aim is to study how polyhedral based program analysis and transformation can be used to address this problem.

In the context of the PhD of Antoine Morvan, we have studied how it was possible to improve the efficiency and applicability of nested loop pipelining (also known as nested software pipelining) in C to hardware tools. Loop pipelining is a key transformation in high-level synthesis tools as it helps maximizing both computational throughput and hardware utilization. Nevertheless, it somewhat losess its efficiency when dealing with small trip-count inner loops, as the pipeline latency overhead quickly limits its efficiency.

Even if it is possible to overcome this limitation by pipelining the execution of a whole loop nest, the applicability of nested loop pipelining has so far been limited to a very narrow subset of loops, namely perfectly nested loops with constant bounds. In this work, we have extended the applicability of nested-loop pipelining to imperfectly nested loops with affine dependencies. We have shown how such loop nest can be analyzed and, under certain conditions, how one can modify the source code in order to allow nested loop pipeline to be applied using a method called polyhedral bubble insertion. Our approach shown encouraging results and led to a publication to the IEEE International Conference on Field Programmable Technology [48] in December 2011.

6.2.2. Reconfigurable Processor Extensions Generation

Participants: Christophe Wolinski, François Charot, Erwan Raffin, Kevin Martin, Antoine Floch.

During this year, we have continued our work on the generation of reconfigurable processor extension using the constraint programming approach. Previously, we showed how all the problems ranging from instruction identification, scheduling and binding to optimized architecture synthesis can be defined and solved using the constraint programming approach. This year, a new pattern scheduling approach has been defined. It enables concurrent match selection and parallel match scheduling on the processor and extension assuming that the execution on an extension is not atomic. It means that the data produced by an extension must not necessarily be sent to the processor just after the end of processing. Thanks to that, a better scheduling can be obtained [71]. The efficient FPGA implementation of processing units require optimization of hardware resources, such as registers and multiplexers. The extension synthesis defined previously has been revisited. For applications from MediaBench, MiBench and MiCrypt benchmark sets, an improvement of 35%, after placement and routing on the Stratix2 Altera FPGA, is observed.
6.2.3. Run-time Reconfigurable Architecture Modeling

Participants: Christophe Wolinski, François Charot, Emmanuel Casseau, Daniel Menard, Antoine Floch, Erwan Raffin, Steven Derrien.

We have continued to work on the modeling problem of the run-time reconfigurable, operator-based, ROMA multimedia architecture. The ROMA processor is composed of a set of $M$ coarse-grained reconfigurable operators, $N$ data memories, a configuration memory, two interconnection networks (between operators and between operators and memories), and dedicated controllers designed for each module of the datapath. A centralized controller manages the configuration and the execution steps. The ROMA processor has three different interfaces: a data interface connected to the operator network, a control interface and a debug interface connected to the main controller. The number of operators, the number of memories and their size can be decided according to application requirements. The compilation flow of our framework rests on the use of an architecture abstract model of the targeted ROMA architecture.

During this year we have focused on the definition of the constraint model to deploy an application graph on the pipeline architecture model. The goal is here to minimize the latency of the pipeline. The main changes are at the operator and memory levels. The operators are pipelined and the dual port memories behave like circular buffers. Recall that in the case of the non pipelined model, the goal is to optimize the execution time of the application under resource constraints. We have carried out experiments to evaluate the quality of our method using different pattern libraries (patterns supported by the ROMA SWP coarse-grained reconfigurable operator, patterns extracted from the MediaBench set) [47]. In these experiments the model has no limitation in terms of number of operators and number of memories. The optimality of the solutions were proven in 93% of cases. More details can be found in [29] and in the Ph.D. thesis of Erwan Raffin [17].

In the context of the RecMotifs project, we have continued to work on a specific design flow integrating STMicroelectronics’ compiler flow. This project also allowed us to bring significant evolution to our pattern analysis software tools. The RecMotifs flow consists in a pattern analysis flow for STMicroelectronics graphml files generated by ST compiler. This flow allows pattern description (description of graphml pattern that can be used in the covering pass), type extraction, pattern generation (pattern generation on a graphml file), covering (covering of a graphml file with minimization of the parallel execution time without any resource constraints). Once the pattern analysis has been applied to the graphml files, C code regeneration can be performed using GeCos.

6.2.4. Floating-Point to Fixed-Point Conversion

Participants: Daniel Menard, Karthick Parashar, Olivier Sentieys, Romuald Rocher, Hai-Nam Nguyen.

For the fixed-point conversion process, different optimization algorithms have been tested. The aim is to minimize the implementation cost under accuracy constraint. In [54], two new algorithms for the word-length optimization procedure, based on the Greedy Randomized Adaptive Search Procedure (GRASP), are proposed. Compared to existing methods, our proposition yields better results and has a complexity between deterministic methods and stochastic methods.

6.3. Algorithm Architecture Interaction

6.3.1. Flexible hardware accelerators for biocomputing applications

Participants: Steven Derrien, Naeem Abbas, Patrice Quinton.

It is widely acknowledged that FPGA-based hardware acceleration of compute intensive bioinformatics applications can be a viable alternative to cluster (or grid) based approach as they offer very interesting MIPS/watt figure of merits. One of the issues with this technology is that it remains somewhat difficult to use and to maintain (one is rather designing a circuit rather than programming a machine).
Even though there exists C-to-hardware compilation tools (Catapult-C, Impulse-C, etc.), a common belief is that they do not generally offer good enough performance to justify the use of such reconfigurable technology. As a matter of fact, successful hardware implementations of bio-computing algorithms are manually designed at RTL level and are usually targeted to a specific system, with little if any performance portability among reconfigurable platforms.

This research work, which is part of the ANR BioWic project, aims at providing a framework for helping semi-automatic generation of high-performance hardware accelerators. In particular we expect to widen the scope of common design constraints by focusing on system-level criterions that involve both the host machine and the accelerator (workload balancing, communications and data reuse optimisations, hardware utilization rate, etc.). This research work builds upon the CAIRN research group expertise on automatic parallelization for application specific hardware accelerators and has been targeting mainstream bioinformatics applications (HMMER, ClustalW and BLAST).

Our work in 2011 extended the experiment results obtained in 2010 and led to the submission of a paper to IEEE Trans. in Parallel and Distributed Computing (the article being in revision). We also investigated another case study based on a more classical sequence comparison algorithm for which we investigated different style of architectural partitioning. This work led to a paper published in the proceedings of the ARC International Symposium [43].

### 6.3.2. Range Estimation and Computation Accuracy Optimization

**Participants:** Daniel Menard, Karthick Parashar, Olivier Sentieys, Romuald Rocher, Pascal Scalart, Aymen Chakhari, Jean-Charles Naud, Emmanuel Casseau, Andrei Banciu.

#### 6.3.2.1. Range Estimation

Efficient range estimation methods are required to optimize the integer part word-length. Our previous works based on the Karhunen-Loève Expansion (KLE) have been extended in [38]. The impulse response between the input and a variable is used to propagate the KLE parameters of the inputs. Range estimation has proven to be a difficult problem for non-linear operations especially when the input data is correlated. A stochastic approach can significantly improve the results compared to the classical methods like the interval and affine arithmetic. The aim is to obtain tight intervals by adapting the bounds to a desired probability of overflows. An approach for the analysis of range uncertainties based on the Polynomial Chaos Expansion (PCE) has been developed. The PCE representation is obtained for every input variable and an analytical description of the variability of the output is determined. Furthermore, the correlation of the inputs is captured using the Nataf transform. The range is computed using a probabilistic analysis from the probability density function (PDF).

#### 6.3.2.2. Accuracy and performance evaluation

The automation of fixed-point conversion requires generic methods to study accuracy degradation. In [51], [73] a new approach using analytical noise power propagation considering conditional structures. These structures are generated from programming language statements such as `if-then-else` or `Switch`. The proposed model takes into account two key points in fixed-point design: first, an alternative processing of noise depending on the condition; second, decision errors generated by quantization noise affecting the condition. This method is integrated in the fixed-point conversion process and uses path probabilities of execution alternatives obtained from profiling. This work extends existing analytical approaches for fixed-point conversion. Experimentations of our analytical method show that it has a fairly accurate noise power estimation compared to the real accuracy degradation. An analytical approach is studied to determine accuracy of systems including unsmooth operators. An unsmooth operator represents a function which is not derivable in all its definition interval (for example the sign operator). The classical model is no valid yet since these operators introduce errors that do not respect the Widrow assumption (their values are often higher than signal power). So an approach based on the distribution of the signal and the noise is proposed. It is applied to the sphere decoding algorithm. We also focus on recursive structure where an error influences future decision. So, the Decision Feedback Equalizer is also considered.

### 6.3.3. Reconfigurable Video Coding

**Participants:** Emmanuel Casseau, Olivier Sentieys, Arnaud Carer, Cécile Beaumin, Hervé Yviquel.
In the field of multimedia coding, standardization recommendations are always evolving. To reduce design time, Reconfigurable Video Coding (RVC) standard allows defining new codec algorithms based on a modular library of components. RVC dataflow-based specification formalism expressly targets multiprocessors platforms. However software processor cannot cope with high performance and low power requirements. Hence the mapping of RVC specifications on hardware accelerators is investigated in this work, as well as the scheduling of the functional units (FU) of the specification. Dataflow programming, such as RVC applications, express explicit parallelism within an application. Although multi-core processors are now available everywhere, few applications are able to truly exploit their multiprocessing capabilities. We describe in [69] a scheduling strategy for executing a dataflow program on multi-core architectures using distributed schedulers and lock-free communications. Actually, our goal is to design an RVC-dedicated reconfigurable architecture with various resources. Our previous results lead to the definition of a reconfigurable FIFO for optimizing cost and performance of RVC dataflow specifications by taking advantage of their dynamic behavior. We are currently working with Mickael Raulet from IETR INSA Rennes and Dr. Jani Boutellier from the university of Oulu (Finland), concerning the execution of an RVC decoder on a network of Transport Triggered Architecture (TTA) processors (proposed by the Tampere University of Technology). Thanks to its modular structure, TTA can be seen as a nice kind of CPU design to develop Application-Specific Processor. TTA processor network is connected by hardware channels so it has many similarities with RVC network. Hervé Yviquel, is expected to have a 4-month stay in 2012 in TUT to provide a functional automated flow to design TTA-based platform and compile RVC application for this platform.

6.3.4. Multi-Antenna Systems

Participants: Olivier Berder, Pascal Scalart, Quoc-Tuong Ngo.

Considering the possibility for the transmitter to get some Channel State Information (CSI) from the receiver, antenna power allocation strategies can be performed thanks to the joined optimization of linear precoder (at the transmitter) and decoder (at the receiver) according to various criteria.

A new exact solution of the maximization of the minimum Euclidean distance between received symbols has been proposed for two 16-QAM modulated symbols. This precoder shows an important enhancement of this minimum distance compared to diagonal precoders, which leads to a significant BER performance improvement. This new strategy selects the best precoding matrix among eight different expressions, depending on the value of the channel angle. Selecting only two of these expressions, this precoder was then generalized to any rectangular QAM modulation [26].

Not only the minimum Euclidean distance but also the number of neighbors providing it has an important role in reducing the error probability when a Maximum Likelihood detection is considered at the receiver. Aiming at reducing this number of neighbors, a new precoder in which the rotation parameter has no influence is proposed for two independent data streams transmitted. The expression of the new precoding strategy is less complex and the space of solution is, therefore, smaller [53], [74]. In the paper [52], we proposed the general \( \text{neighbor-dmin} \) precoder for three independent data streams and the simulation results also confirm a significant bit-error-rate improvement of the new precoder in comparison with other traditional precoding strategies.

6.3.5. Cooperative Strategies for Low-Energy Wireless Networks

Participants: Olivier Berder, Le Quang Vinh Tran, Olivier Sentieys.

During the last decade, many works were devoted to improving the performance of relaying techniques in ad hoc networks. One promising approach consists in allowing the relay nodes to cooperate, thus using spatial diversity to increase the capacity of the system. In wireless distributed networks where multiple antennas can not be installed in one wireless node, cooperative relay and cooperative Multi-Input Multi-Output (MIMO) techniques can indeed be used to exploit spatial and temporal diversity gain in order to reduce energy consumption.
Considering a system having a two-antenna source, two one-antenna relays and a one-antenna destination, MIMO simple cooperative relay model (MSCR) and MIMO full cooperative relay model (MFCR) are proposed in comparison with MIMO normal cooperative relay model (MNCR) where the relays forward signals consecutively to destination. The energy efficiency of these models is investigated by using a realistic power consumption model where the parameters are extracted from the characteristics of CC2420, a wireless sensor transceiver widely used and commercially available. For each transmission ranges, the optimal cooperative scheme in terms of energy efficiency is provided by simulation results [65], [78].

A fair analytical investigation on these cooperative protocols was also performed. A lower bound for the average symbol error probability (ASEP) of full DSTC cooperative relaying system in a Rayleigh fading environment is provided. In the case when the Signal to Noise Ratio (SNR) of the relay-relay link is much greater than that of the source-relay link, the upper bound on ASEP of this system is also derived. The effect of the distance between the relays shows that the performance does not degrade so much as the distance between relays is lower than a half of the source-destination distance. Moreover, we also show that, when the error synchronization range is lower than 0.5, the impact of the transmission synchronization error of the relay-destination link on the performance is not considerable [64].

The energy efficiency of cooperative MIMO and relay techniques is also very useful for the Infrastructure to Vehicle (I2V) and Infrastructure to Infrastructure (I2I) communications in Intelligent Transport Systems (ITS) networks where the energy consumption of wireless nodes embedded on road infrastructure is constrained. Applications of cooperation between nodes to ITS networks are proposed and the performance and the energy consumption of cooperative relay and cooperative MIMO are investigated in comparison with the traditional multi-hop technique. The comparison between these cooperative techniques helps us to choose the optimal cooperative strategy in terms of energy consumption for energy constrained road infrastructure networks in ITS applications [27].

6.3.6. Opportunistic Routing
Participants: Olivier Berder, Olivier Sentieys, Ruifeng Zhang, Jean-Marie Gorce [Insa Lyon, INRIA Swing].

However, the aforementioned approaches introduce an overhead in terms of information exchange, increasing the complexity of the receivers. A simpler way of exploiting spatial diversity is referred to as opportunistic routing. In this scheme, a cluster of nodes still serves as relay candidates but only a single node in the cluster forwards the packet. This paper proposes a thorough analysis of opportunistic routing efficiency under different realistic radio channel conditions. The study aims at finding the best trade-off between two objectives: energy and latency minimizations, under a hard reliability constraint. We derive an optimal bound, namely, the Pareto front of the related optimization problem, which offers a good insight into the benefits of opportunistic routings compared with classical multi-hop routing schemes [31]. We then provided a closed-form expression of the lower bound of the energy-delay tradeoff and of energy efficiency for different channel models (additive white Gaussian noise, Rayleigh fast fading and Rayleigh block-fading) in a linear network. These analytical results are also verified in 2-dimensional Poisson networks using simulations. The closed-form expression provides a framework to evaluate the energy-delay performance and to optimize the parameters in physical layer, MAC layer and routing layer from the viewpoint of cross-layer design during the planning phase of a network.

6.3.7. Adaptive techniques for WSN power optimization
Participants: Olivier Berder, Daniel Menard, Olivier Sentieys, Mahtab Alam, Trong-Nhan Le.

Wireless sensor networks (WSNs) have obtained a great relevancy in civil as well as military applications such as environment sensing, real-time surveillance and habitat monitoring. It is difficult to design a node that is efficient for all of these different applications. The ideal sensor node would have to dynamically adapt its behavior to various parameters such as the data traffic, the channel conditions, the amount of harvested energy, its battery level, etc. Including the capability to scavenge energy from its environment, the design of an efficient power manager able to address both hardware and software processing seems very promising.
Energy modeling is an important issue for designing and dimensioning low power wireless sensor networks (WSN). In order to help the developers to optimize the energy spent by WSN nodes, a pragmatic and precise hybrid energy model is proposed. This model considers different scenarios that occur during the communication and evaluates their energy consumption based on software profiling as well as the hardware components power profiles. The proposed model is a combination of analytical derivations and real time measurements. These experiments are particularly useful to understand the medium access control (MAC) layer mechanisms, such as wake up or data collisions for the preamble sampling category, and the energy wasted by collisions can be evaluated [18], [35].

An adaptive wake-up-interval scheme for preamble sampling MAC protocols for variable traffic in WSN is then proposed. The wake-up-interval is updated based on the traffic status register (whose content depends on the presence of messages for a particular node). The results show that the sensor node adapts and converges its wake-up-interval to the best trade-off value for fixed and variable traffic patterns. Two optimization parameters (length of traffic status register and initial wake-up-interval value) are also tuned to achieve fast convergence speed for different traffic rates and variations.

A wireless body area sensor network (WBASN) demands ultra-low power and energy-efficient protocols. MAC layer plays a pivotal role for energy management in WBASN, moreover, idle listening is the dominant energy waste in most of the MAC protocols. WBASN exhibits wide range of traffic variations based on different physiological data emanating from the monitored patient. In this context, we proposed a novel energy efficient traffic-aware dynamic (TAD) MAC protocol for WBASN [36]. A comparison with other protocols for three different widely used radio chips, i.e. cc2420, cc1000 and amis52100, is presented. The results show that TAD-MAC outperforms all the other protocols under fixed and variable traffic rates.
6. New Results

6.1. VMAD and LLVM

The goal is to provide a set of annotations (pragmas) that the user can insert in the source code to perform low level analyses (profiling) or optimizations (dynamic parallelization for example).

We are developing a virtual machine handling advanced dynamic analyses and transformations of programs. VMAD is organized as a sequence of basic operations, where external modules associated to specific strategies are dynamically loaded when required. The program binary files handled by VMAD are previously instrumented at compile time to include necessary data, instrumentation instructions and callbacks to the virtual machine. Dynamic information, such as memory locations of launched modules, are patched at startup in the binary file. The LLVM compiler has been extended to automatically instrument programs to meet the requirements both of VMAD and of the handled/chosen analysis and transformation strategies.

VMAD uses sampling and multi-versioning to limit the instrumentations time overhead. At runtime, targeted codes are launched by successive chunks that can be either original, instrumented or optimized/parallelized versions. After each chunk execution, decisions can be taken relatively to the current optimization strategy. At this time, VMAD is handling advanced memory access profiling through linear interpolation of the addresses, dynamic dependency analysis and version selection. The last developments are focusing on speculative polyhedral parallelization.

The profiling strategy interpolating the memory addresses accessed in a loop nest has been run on some of the SPEC2006 and Pointer Intensive benchmark suites, showing a very low time overhead, in most cases. More details are available in the research reports [46] and [47], and publications [15] and [16].

6.2. Dynamic version selector

Adaptive version selection between different parallel versions of code is necessary when the execution context of a program is not known. The execution contexts includes all or some of these possibly variable parameters: the target architecture, the load of the computer at execution time, and the input data.

We have developed a framework handling loops in the polyhedral model, that is able to take a runtime decision about which version to execute. It is based on:

- the generation of different code versions of a loop nest;
- an install-time profiling to take into account the architecture parameters, that builds a parametric ranking table between the versions;
- a runtime selection, predicting the load balance and the execution time of each code version, before executing the best one.

We showed that different versions of a code are required on several polyhedral loop nest benchmarks, depending on both the target architecture and the input data. And we showed speedups compared to any statically chosen version in all execution contexts. More details are available in publication [19].

6.3. Binary parallelization

Our work on parallelizing binary programs has continued in 2011, with several new results. The general principle is to analyze the binary code and extract a model of the most intense loops. The model has to include everything that is related to memory access, and also some part of the computations done in registers. Once a suitable model is extracted, it can be used to derive a new scheduling for each targeted loops, optimizing various criteria: this is where polyhedral techniques are used, providing algorithms to optimize
locality, parallelism, or both at the same time. After a new scheduling is computed, the transformed code is
generated by a polyhedron-scanning algorithm. Our approach relies on an intermediate representation whose
emphasis is on memory accesses, hiding, i.e., *outlining* all low-level details and retaining only what is needed
by the parallelization component: we use raw C constructs, and macros to denote outlined code. Starting
with the executable program, a first phase raises the code into our intermediate representation. The second
phase uses a stock parallelizing component, producing a transformed C programs. The last phase lowers
this intermediate representation into a new binary executable. The system then uses a run-time monitoring
component (generated automatically at the same time as the parallel version) that redirects execution to the
transformed loops whenever appropriate.

This year’s activity on this topic in our team has started by finalizing and presenting a paper at the IMPACT
workshop [20], held during CGO’2011, in Chamonix (France). This workshop focuses on tools and
techniques based on the polyhedral model. It has been interesting to hear the various reactions and remarks
of researchers attending our presentation: the general position is that our work opens new perspectives on the
use of the polyhedral model, and avoids to need to have a complete polyhedral tool-chain. Another major
aspect is the fact that our 3-phase strategy clearly separates the polyhedral part of the whole process, in
essence providing a basis for a polyhedral programming language that is slightly more general that what
was considered before.

The work on this topic has continued on two directions. The first was to effectively abstract the parallelizing
phase. This has been done by demonstrating the use of two distinct parallelizers: the first is PLUTO, a
polyhedral locality optimizer and parallelizer, and the second is CETUS, a “simple” parallelizer. The second
new direction was directed by the complexity of typical “real-world” executable programs. It has consisted
in developing new dependence analysis and parallelization techniques, handling more general classes of
programs. This current is currently submitted for publication.

These research results will be presented at the forthcoming HiPEAC conference, to be held in Paris in January
2012. A full-length paper has been accepted for publication in *ACM Transactions on Architecture and Code
Optimization* some time in 2012.

### 6.4. Modeling the dynamic behavior of executable programs

Modeling the dynamic behavior of a given program is useful for several reasons. First, it is a particular way
of profiling the program, targeting non trivial characteristics of the execution. As such, it helps programmers
understand the behavior of the program, and hopefully helps them optimizing it. Second, the results can be
used in a compiler, using run-time information to drive static optimizations. This path has seen considerable
development when it comes, for example, to sequence basic blocks so as to leverage branch predictors and/or
optimize the usage of instruction caches. We ambition to take it one step further, using run-time information to
help a compiler in the task of auto-parallelization. Third, modeling can be used on-line, during the execution
of the program, to drive the use of dynamic optimizations.

Our first achievement in 2011 has been the presentation of our paper at the International Symposium on
Performance Analysis of Software and Systems [17]. The paper describes an approach that e have called
*program skeletonization*. The basic idea is to perform a static analysis of the code under scrutiny, and locate
a small number of register assignments that completely determine the set of memory addresses the program
will access. By instrumenting these elementary value assignments and extracting the ensuing computations
of addresses, the amount of instrumentation can be dramatically reduced, at the cost of offloading some
computations to the profiler. On average, this provides a significant gain in the time needed to obtain a memory
trace, and is independent on the particular application, e.g., cache simulation, data race detection, and so on.

Our second main research direction on this topic has been the characterization of semi-regular memory
accesses. Semi-regular accesses are caused by the traversal of a data structure linking successive memory
cells in no particular memory order, i.e., a linked list or a tree. We have developed a modeling algorithm that is
able to detect that a set of instructions perform irregular accesses that are actually highly correlated, differing
only by an affine function of the enclosing loop indices. This has several implications in terms of potential
optimizations. First, it exhibits a kind of abstract iterator, that can later be handled, e.g., by inspector/executor techniques. Second, it reduces the number of potential dependencies that have to be tested.

The third, most recent, research project that we have started this year is the analysis of traces of parallel programs, currently MPI programs. Parallel traces offer new research challenges, because they contain events that are only partially ordered. We have extended our loop nest recognition algorithm [7] to handle parallel traces. Our preliminary results show that this algorithm is highly effective in extracting communication patterns. This has a number of potential applications that we plan to study in the coming months.

6.5. Dynamic dependence analysis

We have started a research project on dynamic dependence analysis. The principle is to observe an execution of a given program and collect dependence information. This form of profiling is similar to memory profiling, except that to goal is to directly produce data dependencies. During collection, data dependencies are abstracted into dependence graphs which, in turn, give enough information to decide whether a given code portion is parallel and, if the answer is positive, precisely constrain the set of applicable program transformations. Our implementation currently uses our own profiling infrastructure to obtain data from a running program, and models the sequence of run-time data dependencies in a unique framework. The system is of course sensitive to the fact that the input data is representative of typical inputs: the conclusion it draws cannot be applied without resorting to speculation. Currently, we restrict the system to a parallelization assistant, whose result is a set of suggestion that the programmer is free to follow or ignore.

A fundamental characteristic of a parallelization assistant is the class of parallel constructions that it is able to extract from run-time data. We have designed the profiling component and the analysis algorithm such as to be able to represent the dependencies with various abstractions, from simple boolean dependencies to full dependence polyhedra. However, even the simplest dependence model provides useful hints to parallelize a given program. In this case, our tool flags every loop in the program as either intrinsically sequential, or potentially parallel. To make this widely usable, we have decided to target common parallel programming constructions, namely OpenMP directives. In turns out that our design is flexible enough to let the system also target loops that become parallel after simple privatization transformations. The result of profiling is thus a set of complete OpenMP directives, making the loop parallel but also instructing the compiler to take special measure for local data that (falsely) renders loop iterations dependent.

Because the general design of our dynamic dependence analysis system is generic and modular, we plan to develop and distribute software implementing our approach. This project, called Parwiz, has received support from INRIA under the form of an ADT (Action de Développement Technologique). The funding should let us hire an engineer for two years. Unfortunately, we have not been able to find a suitable candidate in 2011, and plan to continue the recruitment process in 2012. In the meantime, we have supervised en Argentinian student, José Cacherosky, during 4 months (from July to October 2011), as an intern with funding from the INRIA International Internships program. The goal of the internship was to extend our framework (developed mainly to work from binary programs) to another execution environment, namely the Java Virtual Machine. Most of the time has been spent on developing an instrumentation infrastructure, but José Cacherosky has also started the implementation of some of the parallelism detection algorithms. We plan to continue this work soon, and provide a tool that could work in several, very distinct environments. Finally, Fabrice Rastello, research scientist at the École Normale Supérieure de Lyon (with the INRIA team COMPSYS), has expressed interest in collaborating with us on this research project.

6.6. Dealing with arithmetic overflows in the polyhedral model

Participants: Nicolas Magaud, Julien Narboux, Éric Violard.

Our goal in collaboration with Alexandre Pilkiewicz, PhD student, and François Pottier, senior researcher at INRIA, is to prove formally the correctness of a compiler based on the polyhedral model and to integrate it in the Compcert compiler.
But as the polyhedral transformations apply to affine loop nests in a mathematical framework where each loop variable is considered to be a mathematical integer, and not a machine integer, we must therefore warrant that no arithmetic overflow occurs when the considered loop nests are executed.

We proposed a solution to produce a compiler which does not ignore the problem of overflows. Our solution consists in generating a formula which captures the presence of overflows in the program, then asking to an external tool (i.e. the iscc calculator), for a sufficient condition about the parameters which implies the absence of overflows. Finally we check this condition dynamically. If the condition holds we can use the optimized version of the program. If it does not, in order to preserve the semantics of the program we keep the original version.

Figure 3 illustrates our solution for overcoming the problem of arithmetic overflows and for ensuring the correctness of polyhedral transformations. In addition to the polyhedral optimizer, our compiler uses an oracle and a validator.

The oracle returns a boolean expression \( b \) which denotes a sufficient condition to ensure that both the original program \((org)\) and the optimized program \((opt)\) do not produce any overflow. Our transformation then builds a program that we call the resulting program, of the shape \( \text{If } b \text{ then } opt \text{ else } org \). It dynamically evaluates the boolean expression \( b \) and executes \( org \), i.e. the original program, if the condition is not fulfilled or \( opt \), i.e. the optimized program, if the condition is true. The resulting program is then transmitted to the validator.

The validator is a function which takes the original program \((org)\), the optimized program \((opt)\) and the resulting program, and returns a boolean: if it returns true, then the resulting program is equivalent to the original one and our compiler therefore produces the resulting program.

We now have to formally prove the validator using the Coq proof assistant.

Figure 3. A solution for overcoming the problem of arithmetic overflows
6. New Results

6.1. NFS-related results

Concerning the number field sieve algorithm for the discrete logarithm problem in prime fields, Răzvan Barbulescu improved the theoretical complexity of the step called "individual logarithm", using, at a crucial point, a sequence of ECM steps with well-tuned, increasing parameters. He also proved that an approach similar to Coppersmith’s factoring factory was feasible as well for discrete logarithm, yielding an improved overall complexity if heavy precomputations are allowed [21].

In 2010, Thomas Prest and Paul Zimmermann developed a new algorithm for the polynomial selection in the Number Field Sieve (NFS). This algorithm produces two non-linear polynomials, extending Montgomery’s "two quadratics" method. For degree 3, it gives two skewed polynomials with resultant $O(N^{5/4})$, which improves on Williams $O(N^{4/3})$ recent result. The paper will appear in the Journal of Symbolic Computation [13] and its impact is assessed by the fact that two preprints extending and analyzing the algorithm have already been proposed.

6.2. Ballot stuffing in a postal voting system

In collaboration between many members of the CASSIS and CARAMEL teams, we have studied a postal voting system used by the CNRS for an election involving about 30,000 voters [16]. The structure of the material can be easily understood out of a few samples of voting material (distributed to the voters), without any prior knowledge of the system. Taking advantage of some flaws in the design of the system, we have shown how to perform major ballot stuffing, making possible to change the outcome of the election. Our attack has been tested and confirmed by the CNRS, and the system was quickly fixed for the next elections.

6.3. Symmetric cryptanalysis

Mohamed Ahmed Abdelraheem, Céline Blondeau, María Naya-Plasencia, Marion Videau, and Erik Zenner have proposed an attack against ARMADILLO2, the recommended variant of a multi-purpose cryptographic primitive dedicated to hardware which has been proposed by Badel et al. in 2010. The attack uses a meet-in-the-middle technique that allows us to invert the ARMADILLO2 core function. This makes it possible to perform a key recovery attack when used as a FIL-MAC. A variant of this attack has been applied to the stream cipher derived from the PRNG mode. A (second) preimage attack is also proposed against the hash function mode. All attacks have been validated by implementing cryptanalysis on scaled variants. The experimental results match the theoretical complexities.

The underlying idea of the attacks, the parallel matching algorithm, has also been generalized. The results are presented in the paper [14].

Thomas Fuhr, Henri Gilbert, Jean-René Reinhard, and Marion Videau have studied the security of the two most recent versions of the message authentication code 128-EIA3, which was considered for adoption (and has been adopted) as a third integrity algorithm in the emerging 3GPP standard LTE. An efficient existential forgery attack against the June 2010 version of the algorithm has been presented. This attack allows, given any message and the associated MAC value under an unknown integrity key and an initial vector, to predict the MAC value of a related message under the same key and the same initial vector with a success probability 1/2. The tweaked version of the algorithm that was introduced in January 2011 to circumvent this attack has also been analysed. While this new version offers a provable resistance against similar forgery attacks under the assumption that (key, IV) pairs are never reused by any legitimate sender or receiver, some evidence is given that some of its design features limit its resilience against IV reuse. The results are presented in the paper [18].
6.4. Implementation of cryptographic pairings

The extended version of a work on parallel architectures for the computation of the $\eta_T$ pairing over supersingular elliptic curves in characteristic 2 and 3, presented at CHES 2009 then accepted at IEEE Transaction on Computers in 2010, was finally published [3]. This paper was the result of a joint effort of Jérémie Detrey and Nicolas Estibals, in collaboration with Jean-Luc Beuchat and Eiji Okamoto (University of Tsukuba, Japan), Francisco Rodríguez-Henríquez (CINVESTAV-IPN, Mexico).

Also, the work on supersingular genus-2 pairings by Diego F. Aranha (UNICAMP, Brazil), Jean-Luc Beuchat (University of Tsukuba, Japan), Jérémie Detrey and Nicolas Estibals was accepted for publication at the Cryptographers’ Track of the RSA Conference (CT-RSA 2012) [15]. Since last year, where only the Eta pairing algorithm was described, several major revisions were undertaken to improve this paper, among which a careful and detailed analysis of the various distortion maps of the considered family of hyperelliptic curves.

This study also allowed us to exhibit a somewhat simple distortion map which would enable this curve to benefit from the shorter loop of the Ate pairing algorithm. Exploring this option is currently work in progress, and the results should eventually be submitted to a journal.

6.5. Multiple-precision arithmetic

In [25], Pascal Molin showed that the error function $\text{erf}$ can be computed very efficiently using a formula involving an integral of a form appropriate for fast evaluation using the trapezoidal scheme. A rigorous analysis of the scheme in this context allows to get precise bounds on the various errors terms, and therefore to give a proven complexity result for the multiple-precision evaluation of $\text{erf}$. The good theoretical behaviour is confirmed by an implementation in Pari.

Together with David Harvey (New York University), P. Zimmermann studied the short division of long integers, i.e., the division of a $2n$-bit integer by an $n$-bit integer where only the integer quotient is wanted, or an approximation of it. They gave detailed algorithms with rigorous errors bounds, and implemented them in GNU MPFR. Using Harvey’s integer middle product code, they obtain a speedup of up to 10% with respect to the best known implementation [20].

With Guillaume Melquiond (Proval project-team, INRIA Saclay), and Prof. W. Georg Nowak (Institute of Mathematics, Vienna), P. Zimmermann worked on the numerical approximation of the Masser-Gramain constant, following some work of Gramain and Weber in 1985. This work disproves a conjecture of Gramain, and enables one to determine the following approximation of that constant:

$$1.819776 < \delta < 1.819833.$$

This work has been completed in 2011 [12].

The article “The Great Trinomial Hunt” has been published in the Notices of the AMS [7].

6.6. Proving the complexity of computing endomorphism rings

Subsequent to the work [6] that has been finally published this year, Gaëtan Bisson has been working on rigorously proving a subexponential running time bound for computing endomorphism rings of ordinary elliptic curves over finite fields. In the end, the proof rests on only one assumption, namely the extended Riemann hypothesis (ERH) [4].

In his thesis [1], he has also made substantial advances towards the extension of these algorithms to genus 2 curves.

In studying the above-mentioned algorithms, Gaëtan Bisson, in collaboration with Andrew V. Sutherland, has designed a low-memory, Pollard-rho type algorithm for finding relations in generic groups [5].
6.7. Point counting on curves with real multiplication

Pierrick Gaudry, David Kohel and Benjamin Smith have designed a new variant of the Schoof algorithm for point counting on hyperelliptic curves, that can take advantage of the presence of the knowledge of an explicit and efficient endomorphism coming from real multiplication. In that case, the overall complexity drops from $\tilde{O}(\log^8 q)$ to $\tilde{O}(\log^5 q)$. Using our algorithm we have computed a 256-bit prime-order Jacobian, suitable for cryptographic applications, and also the order of a 1024-bit Jacobian. The corresponding paper [19] obtained the Best Paper Award at the Asiacrypt 2011 conference.

6.8. Computation of isogenies between abelian varieties

Following the work [11] of David Lubicz and Damien Robert (that has just been accepted for publication in Compositio Mathematica) about the explicit computation of isogenies using theta coordinates, Romain Cosset and Damien Robert [24] have developed further nice features. In the original paper, only $(\ell^2, \ell^2)$-isogenies between abelian surfaces were available. It is now possible to handle $(\ell, \ell)$-isogenies between genus 2 curves, thus providing a more precise tool. Two key elements were necessary: Romain Cosset gave explicit methods to transfer points between the classical representation with Mumford’s coordinates and the theta functions. This is a generalisation of the work of Van Wamelen. And Romain Cosset and Damien Robert developed an explicit algorithm to change the level in the Theta coordinates that are used to represent the geometrical objects. Many details can be found in Cosset’s thesis [2].

Using the same kind of tools, Christophe Arène and Romain Cosset [23] have constructed the first complete addition law on abelian surfaces. Although they are not yet of any practical use, completeness is a feature that is in principal interesting for cryptographic applications.

The article by Faugère, Lubicz and Robert on computing modular correspondences with Theta constants has finally appeared in Journal of Algebra [9].
6. New Results

6.1. Resource analysis by quasi-interpretation

Participants: Guillaume Bonfante, Jean-Yves Marion.

In [15], Guillaume Bonfante, Jean-Yves Marion and Jean-Yves Moyen show how quasi-interpretations can be used to deal with the resource analysis of first order functional programs. This work has been a root for several further development in implicit computational complexity.

6.2. Characterization of programs based on embedding

Participant: Guillaume Bonfante.

So far, in the implicit complexity characterizations based on the ordering MPO developed in the team, we were using the subterm relation to compare values. In [21], we have shown that the embedding relation is a generalization of these results.

6.3. Property proofs for adversary rewrite systems

Participant: Isabelle Gnaedig.

We have continued to work on rewriting property proofs in the adversary context. Our inductive proof technique, initially developed for proving termination of rewriting for systems that do not enjoy the strong termination property, was first proposed to establish termination proofs under particular strategies: the innermost, outermost, local strategies [58].

We then have tackled the proof problem of weak properties i.e., properties that do hold only on certain derivation branches. Weak property proofs are still marginal in the domain of rewriting, probably because classical proof techniques, especially for termination, work on the rules, so that the phenomena arising in the induced rewriting relation are hidden. Our technique, developing proof trees simulating rewriting trees by abstraction and narrowing, explicitly describes the behavior of the studied property on derivation branches, allowing to establish it on good branches. In addition, it is constructive, which is very useful in the programming context: the good branches are identified at compile time, when the proof is established. At run time, derivations are computed only on a good derivation branch, which avoids using the costly breadth-first strategy.

We then have proposed a procedure, based on our inductive principle, for weak termination and C-reducibility, which can be seen as a weak notion of sufficient completeness. The procedure principle is generic and can be instantiated by specific mechanisms related to both properties [20].

6.4. Computer virology: behavioral analysis

Participants: Isabelle Gnaedig, Jean-Yves Marion, Philippe Beaucamps.

Our study on behavioural malware detection has been continued. We have been developing an approach detecting suspicious schemes on an abstract representation of the behavior of a program, by abstracting program traces, rewriting given subtraces into abstract symbols representing their functionality. Considering abstract behaviors allows us to be implementation-independent and robust to variants and mutations of malware. Suspicious behaviors are then detected by comparing trace abstractions to reference malicious behaviors.
Last year, we had proposed to abstract trace automata by rewriting them with respect to a set of predefined behavior patterns defined as a regular language described by a string rewriting system [35]. We have increased the power of our approach on two aspects. We first have modified the abstraction mechanism, keeping the abstracted patterns in the rewritten traces, by just marking them. This now allows us to handle interleaved patterns. Second, we have extended the rewriting framework to express data constraints on action parameters by using term rewriting systems. An important consequence is that, unlike in [35], using the data-flow, we can now detect information leaks in order to prevent unauthorized disclosure or modifications of information [28].

The previous approach has also been extended to a probabilistic model of rewriting, in order to express uncertainty in the behavior pattern recognition. All these results on detection of malware by behavior abstraction have been given in the PhD thesis of Philippe Beaucamps, directed by Isabelle Gnaedig and Jean-Yves Marion, and defended 14 November, 2011 [11].

6.5. Randomness and ergodicity: compressibility
Participant: Mathieu Hoyrup.

In [25], we solve a problem that has been open for 15 years. It relates three notions of complexity and information: Shannon information and entropy, Kolmogorov algorithmic information and Martin-Löf randomness. We obtain that the limit rate of compressibility of a random sequence equals the entropy of the underlying ergodic measure. This result is the achievement of several years of development.

6.6. Randomness and ergodicity: decomposition
Participant: Mathieu Hoyrup.

Results about the forecasting of the long-term statistics in dynamical systems. In previous works we studied the computability of the limit-frequencies. We had proved in particular that in general they cannot be computed, we have turned to the following question: can they be computed, allowing the observation of the system as an oracle? In [24], we obtain several positive results, leaving the general problem open.

6.7. Computability and measure theory
Participant: Mathieu Hoyrup.

In [26], we study the constructive content of the Radon-Nikodym theorem, show that it is not computable in general and precisely locate its non-computability in the Weihrauch lattice.

6.8. Randomness and ergodicity: limit frequencies
Participant: Mathieu Hoyrup.

A new constructive proof of Birkhoff’s ergodic theorem, with as an application a strengthening of former results on random elements: in ergodic systems, random elements eventually reach effective closed sets of positive measure (while it was only known for a more restricted class of sets). The paper [19] is in press and will appear soon in Information and Computation.

6.9. Randomness for a class of measures
Participant: Mathieu Hoyrup.

New results about randomness for a class of measures (and not only for one particular measure) are presented in [14].

6.10. Decidability in Perturbed Dynamical Systems
Participant: Emmanuel Hainry.
We have studied the link between undecidability and robustness in dynamical systems. Indeed, undecidability occurs very easily in dynamical systems. However there exist good decision algorithms that work for most systems that are not pathological. We argue that this decidability trait may be related to their robustness to infinitesimal noise. We have proved that in smooth dynamical systems, robustness is equivalent to decidability of the reachability problem. This result relies on various hypotheses depending on the compactness of the domain and whether time is discrete or continuous [31]

6.11. Complexity in Recursive Analysis

Participant: Emmanuel Hainry.

In [30], we present a characterization of polytime computable functions in the Recursive Analysis setting. This paper in fact presents a generic framework for lifting characterizations of complexity or computability classes in the classical setting into analog characterizations in Recursive Analysis.

6.12. A soft linear logic characterization of polynomial space

Participant: Jean-Yves Marion.

Jean-Yves Marion has worked on light (soft) linear logics with Marco Gaboardi and Simona Ronchi Della Rocca in [16]. This work is based on an extension of a soft linear lambda calculus by means of a conditional construction. It provides a correspondence with the well-known result $\text{APTIME} = \text{PSPACE}$.

6.13. From control flow analysis to complexity

Participant: Jean-Yves Marion.

Jean-Yves Marion proposed a type system for an imperative programming language which certifies time bounds in [27]. It is based on secure flow information analysis as proposed for instance by Bell and La Padula. Thus, a link is done between computational complexity and security-typed languages.
CASCADE Project-Team (section vide)
6. New Results

6.1. Automated Deduction

We develop general techniques which allow us to re-use available tools in order to build a new generation of solvers offering a good trade-off between expressiveness, flexibility, and scalability. We focus on the careful integration of combination techniques and rewriting techniques to design decision procedures for a wide range of verification problems. In his habilitation, Laurent Vigneron presents his contributions to the application of automated deduction for designing decision procedures and for verifying infinite systems, with a particular focus on abstract congruence closure and on verification of security protocols [17].

6.1.1. Building and verifying decision procedures

Participants: Alain Giorgetti, Olga Kouchnarenko, Christophe Ringeissen, Elena Tushkanova.

We have developed a methodology to build decision procedures by using superposition calculi which are at the core of equational theorem provers. We are interested in developing automated deduction techniques to prove properties about these superposition-based decision procedures. To this aim, we plan to further investigate the use of meta-superposition, which has been already applied to check the termination and the combinability of superposition-based procedures [25]. We are working on the development of a framework for specifying and verifying superposition-based procedures. Since these procedures are defined as inference systems, we use the Maude system based on rewriting logic as a specification and prototyping language to implement superposition and meta-superposition.

6.1.2. Combining decision procedures

Participants: Christophe Ringeissen, Michaël Rusinowitch, Valerio Senni.

Modularity is a highly desirable property in the development of satisfiability procedures. In [59] we are interested in using a dedicated superposition calculus to develop satisfiability procedures for (unions of) theories sharing counter arithmetic. In the first place, we are concerned with the termination of this calculus for theories representing data structures and their extensions. To this purpose, we prove a modularity result for termination which allows us to use our superposition calculus as a satisfiability procedure for combinations of data structures. In addition, we present a general combinability result that permits us to use our satisfiability procedures into a non-disjoint combination method à la Nelson-Oppen without loss of completeness. This latter result is useful whenever data structures are combined with theories for which superposition is not applicable, like theories of arithmetic.

6.2. Security Protocol Verification

The design of cryptographic protocols is error-prone. Without a careful analysis, subtle flaws may be discovered several years after the publication of a protocol, yielding potential harmful attacks. In this context, formal methods have proved their interest for obtaining good security guarantees. Many analysis techniques have been proposed in the literature [79]. We have edited a book [64] where each chapter presents an important and now standard analysis technique. We develop new techniques for richer primitives, wider classes of protocols and higher security guarantees.

6.2.1. Modeling complex primitives

Participants: Mathilde Arnaud, Véronique Cortier, Michaël Rusinowitch, Mathieu Turuani.

Some attacks exploit in a clever way the interaction between protocol rules and algebraic properties of cryptographic operators. In [82], we provide a list of such properties and attacks as well as existing formal approaches for analyzing cryptographic protocols under algebraic properties.
Encryption “distributing over pairs” is employed in several cryptographic protocols. We have shown that unification is decidable for an equational theory HE specifying such an encryption [18]. We model block chaining in terms of a simple, convergent, rewrite system over a signature with two disjoint sorts: list and element. and present in [65] an algorithm for deciding the unification problem modulo this rewrite system. Potential applications of this unification procedure include flaw detection for protocols employing the CBC encryption mode. We have also proposed in [13][28] an algorithm for solving general intruder constraints in the equational theory ACI. This last result is useful for handling set datastructures and also multiple intruders.

In their seminal work Dolev and Yao used string rewriting to check protocol security against an active intruder. The main technical result and algorithm were improved by Book and Otto who formulated the security check in terms of an extended word problem for cancellation rules. We extend in [66] their main decidability result to a larger class of string rewrite systems called opt-monadic systems.

Most current techniques do not apply to protocols that perform recursive computation e.g. on a list of messages received from the network. While considering general recursive input/output actions very quickly yields undecidability, we provide NPTIME decision procedures on protocols that perform recursive tests on received messages but output messages that depend on the inputs in a standard way [26]. This is in particular the case of secured routing protocols, distributed right delegation or PKI certification paths.

We have also shown [19] that deducibility and static equivalence are decidable for the equational theories modeling trapdoor commitment and re-encryption, that are particularly relevant in the context of e-voting protocols.

6.2.2. Voting and Advanced Classes of Protocols

Participants: Mathilde Arnaud, Stefan Ciobaca, Véronique Cortier, Steve Kremer, Mathieu Turuani, Laurent Vigneron, Cyrille Wiedling.

New classes of protocols are still emerging and not all can be analysed using existing techniques. We study how to cover the emergent families of security protocols with a special focus on voting protocols.

Voting Protocols. Voting is a cornerstone of democracy and many voting systems have been proposed so far, from old paper ballot systems to purely electronic voting schemes. Although many works have been dedicated to standard protocols, very few address the challenging class of voting protocols. One major issue is the fact that privacy-related properties are stated using equivalences, which are very difficult to prove. We have studied several protocols that are currently in use:

- Helios is an open-source web-based end-to-end verifiable electronic voting system, used e.g. by UCL and the IACR association in real elections. We have discovered a vulnerability which allows an adversary to compromise the privacy of voters and we have presented a fixed version, showed to satisfy a formal definition of ballot secrecy using the applied pi calculus [39]. The vulnerability we discovered apply to some other protocols of the literature [71]. Studying further the Helios protocol, we have provided a computational proof of ballot secrecy [30].

- Norway has used e-voting in its last political election in September 2011, with more than 25,000 voters using the e-voting option. Using formal models, we have analyzed the underlying protocol w.r.t. privacy, considering several corruption scenarios [69].

- We have reviewed a postal voting system used in spring 2011 by the French research institute CNRS and designed by a French company (Tagg Informatique). We have shown how to perform major ballot stuffing, making possible to change the outcome of the election [38]. Our attack has been tested (without any prior knowledge of the system except a few samples of voting material) and confirmed by the CNRS.

Securing routing Protocols. The goal of routing protocols is to construct valid routes between distant nodes in the network. If no security is used, it is possible for an attacker to disorganize the network by maliciously interacting with the routing protocols, yielding invalid routes to be built. That is why secure versions of routing protocols are now developed. In her PhD thesis [12], Mathilde Arnaud has proposed a new model and an
associated decision procedure to check whether a routing protocol can ensure that honest nodes only accept valid routes, even if one of the nodes of the network is compromised. This result has been obtained for a bounded number of sessions, adapting constraint solving techniques.

Automated verification of indistinguishability properties. New emerging classes of protocols often require to model less classical security properties, such as anonymity properties, strong versions of confidentiality and resistance to offline guessing attacks. Many of these properties can be modelled using the notion of indistinguishability by an adversary, which can be conveniently modeled using process equivalences. In [67] we present a novel procedure to verify equivalence properties for a bounded number of sessions which is able to handle a large class of equational theories. Although, we were unable to prove termination of the resolution procedure, the procedure has been implemented in a prototype tool and has been effectively tested on examples, some of which were outside the scope of existing tools, including fully automated checking of anonymity of an electronic voting protocol by Fujioka et al.

6.2.3. Securely Composing Protocols

Participants: Stefan Ciobaca, Véronique Cortier, Steve Kremer.

Protocols are often built in a modular way. For example, authentication protocols may assume pre-distributed keys or may assume secure channel. However, when an authentication protocol has been proved secure assuming pre-distributed keys, there is absolutely no guarantee that it remains secure when executing a real protocol for distributing the keys. How the security of these protocols can be combined is an important issue that is studied in the PhD thesis of Stefan Ciobaca [15]. More precisely, we show how protocols sharing data can be safely interleaved, provided that they use disjoint primitives or that each common primitive contains some tag identifying each protocol, like e.g. the name of the protocol. As a sub-result, we provide sufficient and simple conditions for composing key distribution protocols with any protocol using secure channels or pre-distributed keys.

Moreover, we studied [35] whether password protocols can be safely composed, even when a same password is reused. The hypothesis that users do not reuse the same password for different protocols seems indeed unreasonable. More precisely, we present a transformation which maps a password protocol that is secure for a single protocol session (a decidable problem) to a protocol that is secure for an unbounded number of sessions. Our result provides an effective strategy to design secure password protocols: (i) design a protocol intended to be secure for one protocol session; (ii) apply our transformation and obtain a protocol which is secure for an unbounded number of sessions. Our technique also applies to compose different password protocols allowing us to obtain both inter-protocol and inter-session composition.

6.2.4. Soundness of the Dolev-Yao Model

Participants: Véronique Cortier, Guillaume Scerri.

All the previous results rely on symbolic models of protocol executions in which cryptographic primitives are abstracted by symbolic expressions. This approach enables significantly simple and often automated proofs. However, the guarantees that it offers have been quite unclear compared to cryptographic models that consider issues of complexity and probability. Cryptographic models capture a strong notion of security, guaranteed against all probabilistic polynomial-time attacks. A recent line of research consists in identifying cases where it is possible to obtain the best of both cryptographic and formal worlds in the case of public encryption: fully automated proofs and strong, clear security guarantees. We have proposed a survey [22] of the results obtained so far.

Existing soundness results for symmetric encryption are not satisfactory. This is due to the fact that dishonest keys may introduce many behaviors that cannot be easily captured in symbolic models. We discuss the difficulties and limitations of the available results in [37]. In particular, we provide several examples of protocols that are symbolically correct but computationally flawed if assuming IND-CCA2. Based on these findings, Guillaume Scerri has started a PhD thesis on designing more flexible symbolic models for cryptographic proofs. His first result is a computationally sound symbolic model in the presence of dishonestly generated keys, allowing a symbolic adversary to generate new equalities between terms, on-the-fly.
A soundness result is usually established for some set of cryptographic primitives and extending the result to encompass new primitives typically requires redoing most of the work. In [41], [40], we propose a notion of computational soundness, amenable to modular extensions. Specifically, we prove that a deduction sound implementation of some arbitrary primitives can be extended to include asymmetric encryption and public data-structures (e.g. pairings or list), without repeating the original proof effort. Furthermore, our notion of soundness concerns cryptographic primitives in a way that is independent of any protocol specification language.

6.3. Model-based Verification

We have investigated extensions of regular model-checking to new classes of rewrite relations on trees. We have studied specification and proof of modular imperative programs.

6.3.1. Safety Verification Techniques with Regular Fixpoint Computations

Participants: Roméo Courbis, Pierre-Cyrille Héam, Olga Kouchnarenko.

Term rewriting systems are now commonly used as a modelling language for programs or systems. On those rewriting based models, reachability analysis, i.e. proving or disproving that a given term is reachable from a set of input terms, provides an efficient verification technique. Many recent works have shown the relevance of regular approximation techniques to tackle in practice undecidable reachability problems.

We propose in [42] to exploit rewriting approximations developed in [87] for analysing properties of CCS specifications (without renaming). The approach has been implemented and used to verify properties of the Alternating Bit Protocol and of hardware components specifications expressed as CCS processes.

6.3.2. Rewriting-based Mathematical Model Transformations

Participant: Alain Giorgetti.

We have initiated a collaboration with the Department “Temps-Fréquence” of the FEMTO-ST institute (Franche-Comté Electronique Mécanique Thermique et Optique - Sciences et Technologies, CNRS UMR 6174) on the formalization of multiscale methods for MEMS arrays. Multiscale methods provide a solution for the simulation of large MEMS arrays, by approximating their mathematical model. The resulting approximated model can be rigorously derived from the exact one through a sequence of formal transformations that differs for each case. A great challenge is to generalize these formal computations and to automate them, at least in part. This exploratory research has been supported in 2011 by the University of Franche-Comté with a BQR (Research Quality Bonus) of 5000 euros, and by the CASSIS project with a 6 months post-doctoral position. A first contribution is the design of a rule-based transformation language and its implementation as a Maple package [72]. A second contribution is the specification of lazy rewriting modulo associativity and commutativity [29].

For a more scalable treatment of linearity we plan in a near future to detect the scalar nature of mathematical terms by assigning a type to each expression and then to develop a type-checker. We also plan to guide computation by goals, i.e. to adapt reachability analysis to mathematical models.

6.3.3. Algorithms for Tree Walking Automata

Participants: Pierre-Cyrille Héam, Vincent Hugot, Olga Kouchnarenko.

Tree walking automata are widely used to tackle data base algorithmic problems, particularly to analyse queries over XML documents. The emptiness problem for tree walking automata is known to be EXPTIME-complete. The general algorithm to solve this problem consists in transforming the tree walking automaton into a classical top-down tree automaton. The best known algorithm in the literature works in time $O(s2^n)$ where $n$ is the number of states of the tree walking automaton, and $s$ is the size of the alphabet. In [52] we proposed a new algorithm based on an overloop concept and working in time $O(2^{\log n})$. Then we improved our approach for deterministic tree walking automata to have in this case a $O(2^n \log n)$ time complexity. Finally, we also proposed a polynomial-time approximation based semi-algorithm for the emptiness problem, providing very promising experimentation.
6.3.4. Verification of Linear Temporal Patterns over Finite and Infinite Traces

Participants: Pierre-Cyrille Héam, Vincent Hugot, Olga Kouchnarenko.

In the regular model-checking framework, reachability analysis can be guided by temporal logic properties, for instance to achieve the counter example guided abstraction refinement (CEGAR) objectives. A way to perform this analysis is to translate a temporal logic formula expressed on maximal rewriting words into a “rewrite proposition” — a propositional formula whose atoms are language comparisons, and then to generate semi-decision procedures based on (approximations of) the rewrite proposition. In [73] we investigated suitable semantics for LTL on maximal rewriting words and their influence on the feasibility of a translation, and we proposed a general scheme providing exact results on a fragment of LTL corresponding mainly to safety formulæ, and approximations on a larger fragment.

We study in collaboration with A. Lanoix (LINA, Nantes) infinite state models of component-based systems supporting dynamic reconfigurations. To validate such complex systems, there is a need to check model consistency and also to ensure that dynamic reconfigurations satisfy integrity constraints, invariants, and also temporal constraints over reconfiguration sequences. In [55], we proposed to check the model consistency through reconfigurations by combining proof and bounded model-checking techniques. Furthermore, in [46] we proposed to specify dynamic reconfigurations by using more complex architectural constraints and linear temporal logic patterns. As component-based systems evolve at runtime, there is a need to evaluate these properties at runtime, even if only a partial information is expected. For this purpose we introduced a new four-valued logic with potential true and potential false values; they are chosen whenever an observed behaviour has not yet led to a violation or acceptance of the property under consideration. We then implemented the runtime verification of linear temporal patterns by reusing the FPath and FScript tools [83].

6.3.5. Lower Bounds for Computing the pro-Group Closure of a Regular language

Participant: Pierre-Cyrille Héam.

The profinite topology is used in rational languages classification. In particular, several important decidability problems, related to the Malcev product, reduce to the computation of the closure of a rational language in the profinite topology. It is known that given a rational language by a deterministic automaton, computing a deterministic automaton accepting its profinite closure can be done with an exponential upper bound. We prove in [23] that this upper bound is also a worst case lower bound if the alphabet contains at least three letters.

6.4. Model-based Testing

Our research in Model-Based Testing (MBT) aims to extend the coverage of tests. The coverage refers to several artefacts: model, test scenario/property, and code of the program under test [60]. The test generation uses various underlying techniques such as symbolic animation of models [61] or symbolic execution of programs by means of dedicated constraints or SMT solvers, or model-checkers.

6.4.1. Automated Test Generation from Behavioral Models

Participants: Fabrice Bouquet, Pierre-Christophe Bué, Kalou Cabrera, Jérome Cantenot, Frédéric Dadeau, Stéphane Debricon, Elizabeta Fournaret, Jonathan Lasalle.

We have introduced an original model-based testing approach that takes a behavioural view (modelled in UML) of the system under testing and automatically generates test cases and executable test scripts according to model coverage criteria. We have extended this result to SysML specifications for validating embedded systems [24], [57], [56].
In the context of software evolution, we have worked on exploiting the evolution of requirements in order to classify test sequences, and precisely target the parts of the system impacted by this evolution [49], [50]. We have proposed to define the life cycle of a test via three test classes: (i) Regression, used to validate that unimpacted parts of the system did not change, (ii) Evolution, used to validate that impacted parts of the system correctly evolved, and (iii) Stagnation, used to validate that impacted parts of the system did actually evolve. The associated algorithms are under implementation in a dedicated prototype to be used in the SecureChange European project [62]. A link with the security model proof has been started with partners of the project in [51] that allows to generate test needs associated to security properties verified on model.

6.4.2. Scenario-Based Verification and Validation

Participants: Fabrice Bouquet, Kalou Cabrera, Frédéric Dadeau, Elizabeta Fourneret.

Test scenarios represent an abstract test case specification that aims at guiding the model animation in order to produce relevant test cases. Contrary to the previous section, this technique is not fully automated since it requires the user to design the scenario, in addition to the model.

We have designed a scenario based testing language for UML/OCL that can be either connected to a model animation engine [31] or to a symbolic animation engine, based on a set-theoretical constraint solver [20]. In the context of the ANR TASCCC project, we are investigating the automation of test generation from Security Functional Requirements (SFR), as defined in the Common Criteria terminology. SFRs represent security functions that have to be assessed during the validation phase of security products (in the project, the Global Platform, an operating system for latest-generation smart cards). To achieve that, we are working on the definition of description patterns for security properties, to which a given set of SFRs can be related. These properties are used to automatically generate test scenarios that produce model based test cases. The traceability, ensured all along the testing process, makes it possible to provide evidences of the coverage of the SFR by the tests, required by the Common Criteria to reach the highest Evaluation Assurance Levels. We have proposed a dedicated formalism to express test properties [32]. A test property is first translated into a finite state automaton whose coverage by a given test suite is then measured. This makes it possible to evaluate the relevance of the test suite w.r.t. a given property.

In the context of the SecureChange project, we also investigate the evolution of test scenarios. As the system evolves, the model evolves, and the associated test scenarios may also evolve. We are currently extending the test generation and management of system evolutions to ensure the preservation of the security.

6.4.3. Mutation-based Testing of Security Protocols

Participants: Frédéric Dadeau, Pierre-Cyrille Héam.

Verification of security protocols models is an important issue. Nevertheless, the verification reasons on a model of the protocol, and does not consider its concrete implementation. While representing a safe model, the protocol may be incorrectly implemented, leading to security flaws when it is deployed. We have proposed a model-based penetration testing approach for security protocols [44]. This technique relies on the use of mutations of an original protocol, proved to be correct, for injecting realistic errors that may occur during the protocol implementation (e.g. re-use of existing keys, partial checking of received messages, incorrect formatting of sent messages, use of exponential/xor encryption, etc.). Mutations that lead to security flaws are used to build test cases, which are defined as a sequence of messages representing the behavior of the intruder. We have applied our technique on protocols designed in HLPSL, and implemented a protocol mutation tool that performs the mutations. The mutants are then analyzed by the CL-Atse [90] front-end of the AVISPA toolset [74]. Experiments show the relevance of the proposed mutation operators and the efficiency of the CL-Atse tool to conclude on the vulnerability of a protocol and produce an attack trace that can be used as a test case for implementations.

6.4.4. Code-related Test Generation and Static Analysis

Participants: Alain Giorgetti, Frédéric Dadeau, Ivan Enderlin.
In 2011 we have enriched with program slicing [33] an original combination of static analysis and structural program testing for C program debugging presented in 2010, implemented in a prototype called SANTE (Static ANalysis and TEsting). The method first calls a static value analysis which generates alarms when it cannot guarantee the absence of run-time errors. In order to simplify test generation, the method then reduces the program by program slicing and produces one or many simpler programs, while preserving a subset of the alarms. Finally the method performs an alarm-guided test generation to analyze the simplified program(s), in order to confirm or reject alarms. Experiments on real examples have shown that the verification is faster when reducing the code with program slicing. Moreover, the simplified program(s) makes the detected errors and the remaining alarms easier to analyze.

We have designed a grey-box testing and analysis tool [45] for Java programs possibly annotated by JML annotations. This tool uses a set-theoretical constraint representation of the Java code of class methods. It provides an efficient means for (i) generating structural test cases, satisfying a given code-coverage criterion (all-nodes, all-transitions, all-k-paths) and taking into account the JML annotations associated to the method, and (ii) performing static analysis on the Java code, either to detect potential runtime errors (null pointers dereferencing, division by zero, etc.) or to detect non-conformances between the Java program and its JML specifications (invariant, internal precondition or postcondition violation).

We have designed a new annotation language for PHP, named PRASPEL [48] for PHP Realistic Annotation SPEcification Language. This language relies on realistic domains which serve two purposes. First, they assign to a data a domain that is supposed to be specific w.r.t. a context in which it is employed. Second, they provide two features that are used for test generation: (i) samplability makes it possible to automatically generate a value that belongs to the realistic domain so as to generate test data, (ii) predicability makes it possible to check if the value belongs to a realistic domain. This approach is tool-supported in a dedicated framework for PHP which makes it possible to produce unit test cases using random data generators, execute the test cases on an instrumented implementation, and decide the conformance of the code w.r.t. the annotations by runtime assertion checking.

### 6.4.5. Random Testing

**Participant:** Pierre-Cyrille Héam.

The random testing paradigm represents a quite simple and tractable software assessment method for various testing approaches. When doing random testing, the main qualities required for the random sampler are that random choices must be objective and independent of tester choices or convictions: a solution is to ask for uniform random generators.

In [86] a method is proposed for drawing paths in finite graphs uniformly and it is showed how to use these techniques in a control flow graph based testing approach of C programs. Nevertheless, a finite graph often represents a strong abstraction of the system under test, and many abstract tests generated by the approach may be impossible to play on the implementation. In [53], we propose a new approach, extending previous work, to manage stack-call during the random test generation while preserving uniformity.

When doing random testing on inputs, the algorithm has to be efficient enough to allow the generation of a huge quantity of data. Moreover every programming language provides good uniform random generators (or pseudo-random to be more precise) for numbers. However, the question is more complex for non-numerical data, such as tree data structures, logical formulas, graphs, etc. In [54], we present the Seed prototype that uniformly generates recursive data structures satisfying a given grammar-like specification. The tool is easy to use, uniform and generation is uniform. Moreover, it manages some equational equivalences on data structures to shape the distribution.

### 6.5. Verification of Collaborative Systems

We investigate security problems occurring in decentralized systems. We develop general techniques to enforce read and update policies for controlling access to XML documents based on recursive DTDs (Document Type Definition). Moreover, we provide a necessary and sufficient condition for undoing safely replicated objects in order to enforce access control policies in an optimistic way.
6.5.1. Automatic Analysis of Web Services Security

Participants: Tigran Avanesov, Mohamed Anis Mekki, Michaël Rusinowitch, Mathieu Turuani, Laurent Vigneron.

Automatic composition of web services is a challenging task. Many works have considered simplified automata models that abstract away from the structure of messages exchanged by the services. For the domain of secured services (using e.g. digital signing or timestamping) we propose a novel approach to automated orchestration of services under security constraints. Given a community of services and a goal service, we reduce the problem of generating a mediator between a client and a service community to a security problem where an intruder should intercept and redirect messages from the service community and a client service till reaching a satisfying state. In his thesis Mohamed Anis Mekki [36][27] presents a tool that compiles the obtained trace describing the execution of a the mediator into its corresponding runnable code. For that the tool computes an executable specification of the mediator as prudent as possible of her role in the orchestration. This specification is expressed in ASLan language, a formal language designed for modeling Web Services tied with security policies that was developed in AVANTSSAR project. Then we can check with automatic tools that this ASLan specification verifies required security properties such as secrecy and authentication. If no flaw is found, we compile the specification into a Java servlet that can be used by the mediator to execute the orchestration.

In his thesis, Tigran Avanesov [13][28] gives a decision procedure for the satisfiability problem of general deducibility constraints. Two cases are considered: the standard Dolev-Yao theory and its extension with an associative, commutative idempotent operator. The result is applied to solve the automated distributed orchestration problem for secured Web services. As a second application a procedure is given to decide the security of a cryptographic protocol in the presence of several non-communicating intruders. It is also shown in this thesis how to detect some XML rewriting attacks on Web services.

6.5.2. Secure Querying and Updating of Recursive XML Views

Participants: Bao Thien Hoang, Houari Mahfoud, Abdessamad Imine.

Most state-of-the-art approaches for securing XML documents allow users to access data only through authorized views defined by annotating an XML grammar (e.g. DTD) with a collection of XPath expressions. To prevent improper disclosure of confidential information, user queries posed on these views need to be rewritten into equivalent queries on the underlying documents. A major concern here is that query rewriting for recursive views is still an open problem. In this work, we show that this query rewriting is possible using only the expressive power of the standard XPath [70]. We present the extension of the downward class of XPath, composed only by child and descendant axes, with some axes and operators and we propose a general approach to rewrite queries under recursive XML views. Unlike Regular XPath-based works, we provide a linear rewriting algorithm which processes the queries only over the annotated XML grammar. An experimental evaluation demonstrates that our algorithm is efficient and scales well. Finally, we plan to investigate how to combine read and update policies without revealing sensitive information to unauthorized users.

6.5.3. On the Undoability Problem in Distributed Collaborative Systems

Participants: Asma Cherif, Abdessamad Imine.

Combining Operational Transformation (OT) and undo approaches is a challenging problem. Even though various undo solutions have been proposed over the recent years, verifying their correctness still is a challenging problem due to the absence of formal guidelines to undo operations. In this work, we address the undo problem from a theoretical point of view [68]. We provide a necessary and sufficient condition for undoing replicated objects based on OT with respect to three inverse properties. To overcome the difficulty of necessity proof, we use Constraint Satisfaction Problems (CSP) theory in order to cover all possible transformation cases. As the main result, we prove that it is impossible to achieve a correct undo for objects with non-commutative operations. To relax this impossibility result, we sketch a preliminary solution that consists in adding explicitly a new form of idle operations.
5. New Results

5.1. Control-Flow Analysis by Abstract Interpretation

Control-flow analysis (CFA) of functional programs is concerned with determining how the program’s functions call each other. In the case of the lambda calculus, this amounts to computing the flow of lambda expressions in order to determine what functions are effectively called in an application \((e_1 \ e_2)\). This work shows that it is possible to use abstract interpretation techniques to derive systematically a control-flow analysis for a simple higher-order functional language. The analysis approximates the interprocedural control-flow of both function calls and returns in the presence of first-class functions and tail-call optimization. A number of advantages follow from taking this approach:

- The systematic derivation of a CFA for a higher-order functional language from a well-known operational semantics provides the resulting analysis with strong mathematical foundations. Its correctness follows directly from the general theorems of abstract interpretation.
- The approach is easily adapted to different variants of the source language. We demonstrate this by deriving a CFA for functional programs written in continuation-passing style.
- The common framework of these analyses enables their comparison. We take advantage of this to settle a question about the equivalence between the analysis of programs in direct and continuation-passing style.
- The resulting equations can be given an equivalent constraint-based presentation, providing ipso facto a rational reconstruction and a correctness proof of constraint-based CFA.

This work was presented at the Japanese Shonan workshop on Verification of higher-order functional programs in September 2011. A journal article is accepted to appear in Information and Computation.

5.2. Modular SMT Proofs for Fast Reflexive Checking inside Coq

Participants: Frédéric Besson, Pierre-Emmanuel Cornilleau, David Pichardie.

Satisfiability Modulo Theory (SMT) solvers are efficient automatic provers for combination of theories. Those solvers have proved very successful in program verification because they discharge automatically and efficiently challenging verification conditions. SMT solvers are therefore de facto part of the Trusted Computing Base of many program verification methodologies. A consequence is that a soundness bug in a SMT solver can make the whole program verification process unsound.

To tackle this problem, we propose a new methodology for exchanging unsatisfiability proofs between an untrusted SMT solver and a sceptical proof assistant with computation capabilities like Coq. We advocate modular SMT proofs that separate boolean reasoning and theory reasoning; and structure the communication between theories using Nelson-Oppen combination scheme.

We present the design and implementation of a Coq reflexive verifier that is modular and allows for fine-tuned theory-specific verifiers. The current verifier is able to verify proofs for quantifier-free formulae mixing linear arithmetic and uninterpreted functions. Our proof generation scheme benefits from the efficiency of state-of-the-art SMT solvers while being independent from a specific SMT solver proof format. Our only requirement for the SMT solver is the ability to extract unsat cores and generate boolean models. In practice, unsat cores are relatively small and their proof is obtained with a modest overhead by our proof-producing prover. We present experiments assessing the feasibility of the approach for benchmarks obtained from the SMT competition.

This work has been presented at the CPP conference [15] and the international PxTP workshop [21], [20].
5.3. Secure the Clones: Static Enforcement of Policies for Secure Object Copying

Participants: Thomas Jensen, Florent Kirchner, David Pichardie.

Exchanging mutable data objects with untrusted code is a delicate matter because of the risk of creating a data space that is accessible by an attacker. Consequently, secure programming guidelines for Java stress the importance of using defensive copying before accepting or handing out references to an internal mutable object.

However, implementation of a copy method (like clone()) is entirely left to the programmer. It may not provide a sufficiently deep copy of an object and is subject to overriding by a malicious sub-class. Currently no language-based mechanism supports secure object cloning.

We propose a type-based annotation system for defining modular copy policies for class-based object-oriented programs. A copy policy specifies the maximally allowed sharing between an object and its clone. We provide a static enforcement mechanism that will guarantee that all classes fulfill their copy policy, even in the presence of overriding of copy methods, and establish the semantic correctness of the overall approach in Coq.

The mechanism has been implemented and experimentally evaluated on clone methods from several Java libraries. The work as been presented at ESOP [18] this year and is under reviewing for a journal special issue.

5.4. Fault localization and correction in Constraint Programs

Participants: Nadjib Lazaar, Arnaud Gotlieb.

Nowadays, constraint programs are written in high-level modelling languages. Their verification is currently based on trace analysis techniques but does not integrate systematic testing techniques. In this work, we developed a Testing framework for catching the peculiarities of constraint program development, throughout the notions of conformity relations, fault localization and correction.

Within the context of the Nadjib Lazaar’s PhD (defense on 5 Dec. 2011), we explored in 2011 the testing of constraint programs written in OPL and the development of trace-based fault localization and correction techniques [19]. Lazaar’s tool called CPTEST showed impressive experimental results on four hard problems of the CP Community, leading to a publication (in progress) in the Constraints Journal.

5.5. Floating-point constraint solving

Participants: Matthieu Carlier, Arnaud Gotlieb.

Programs including floating-point computations are known to be hard-to-test. Generating test inputs for those programs requires solving constraints over floating-points computations, which led us to the development of specific constraint filtering techniques. In this work, we extended the Marre and Michel property regarding the use of internal floating-point representation to increase the filtering capabilities of addition to the case of multiplication/division. We came up with an optimized implementation of FPSE (our current FP constraint solver) that was able to deal with large C programs that include (non-linear) floating-point computations. We already got a first publication of this work [17].

5.6. Fast inference of polynomial invariants

Participants: David Cachera, Thomas Jensen, Arnaud Jobin, Florent Kirchner.
The problem of automatically inferring polynomial (non-linear) invariants of programs is still a major challenge in program verification. We have proposed an abstract interpretation based method to compute polynomial invariants for imperative programs. Our analysis is a backward propagation approach that computes preconditions for equalities like $g = 0$ to hold at the end of execution. Properties are expressed using ideals, a structure that satisfies the descending chain condition, enabling fixpoints computations to terminate without use of a widening operator. In the general case, termination would be characterized using ideal membership tests and Gröbner bases computations. In order to optimize computational complexity, we propose a specialized analysis dealing with inductive invariants which ensures fast termination of fixpoints computations. The optimized procedure has been shown by experiments to work well in practice, and to be two orders of magnitude faster than state of the art analyzers [23].
6. New Results

6.1. Foundations of information hiding

Information hiding refers to the problem of protecting private information while performing certain tasks or interactions, and trying to avoid that an adversary can infer such information.

This is one of the main areas of research in Comète, and two PhD thesis based on this topic have been defended this year in Comète [12], [11] have been defended this year. We are exploring several topics, described below. An overview of our results is contained in [24].

6.1.1. The problem of information hiding in presence of concurrency

The analysis of probabilistic concurrent systems usually relies on the notion of scheduler in order to solve the nondeterminism. Unfortunately the classical notion of scheduler, which is a mathematical function that chooses the next step depending on the history of the computation, can leak any secret information contained in the history. This creates false positives, and it is known as the problem of the allmighty scheduler. One way to solve this problem, already explored in literature, is to fix the strategy of the scheduler beforehand [31]. However this solution is considered too rigid and unrealistic. In [14] we have propose a milder restriction on the schedulers, and we have defined the notion of strong (probabilistic) information hiding under various notions of observables. Furthermore, we have proposed a method, based on the notion of automorphism, to verify that a system satisfies the property of strong information hiding, namely strong anonymity or no-interference, depending on the context.

6.1.2. Modeling the knowledge of the adversary

In [15] we have developed a game semantics for process algebra with two interacting agents. The purpose of our semantics is to make manifest the role of knowledge and information flow in the interactions between agents and to control the information available to interacting agents. We have defined games and strategies on process algebras, so that two agents interacting according to their strategies determine the execution of the process, replacing the traditional scheduler. We have shown that different restrictions on strategies represent different amounts of information being available to a scheduler. We have also shown that a certain class of strategies corresponds to the syntactic schedulers of Chatzikokolakis and Palamidessi [32], which were developed to overcome problems with traditional schedulers modeling interaction. The restrictions on these strategies have an explicit epistemic flavor.

6.1.3. Opacity

Opacity is a security property formalizing the absence of secret information leakage and we have addressed in [30] the problem of synthesizing opaque systems. A secret predicate $S$ over the runs of a system $G$ is opaque to an external user having partial observability over $G$, if s/he can never infer from the observation of a run of $G$ that the run belongs to $S$. We have chosen to control the observability of events by adding a device, called a mask, between the system $G$ and the users. We have first investigated the case of static partial observability where the set of events the user can observe is fixed a priori by a static mask. In this context, we have shown that checking whether a system is opaque is PSPACE-complete, which implies that computing an optimal static mask ensuring opacity is also a PSPACE-complete problem. Then, we have introduced dynamic partial observability where the set of events the user can observe changes over time and is chosen by a dynamic mask. We have shown how to check that a system is opaque with respect to a dynamic mask and we have also addressed the corresponding synthesis problem: given a system $G$ and secret states $S$, compute the set of dynamic masks under which $S$ is opaque. Our main result is that the set of such masks can be finitely represented and can be computed in EXPTIME and this is a lower bound. Finally we have also addressed the problem of computing an optimal mask.
6.1.4. Interactive systems

In [13], we have considered systems where secrets and observables can alternate during the computation. We have shown that the information-theoretic approach which interprets such systems as (simple) noisy channels is not valid anymore. However, the principle can be recovered if we consider more complicated types of channels, that in Information Theory are known as channels with memory and feedback. We have shown that there is a complete correspondence between interactive systems and such kind of channels. Furthermore, we have shown that the capacity of the channels associated to such systems is a continuous function of the Kantorovich metric.

6.1.5. Differential privacy

Differential privacy is a notion that has emerged in the community of statistical databases, as a response to the problem of protecting the privacy of the database’s participants when performing statistical queries. The idea is that a randomized query satisfies differential privacy if the likelihood of obtaining a certain answer for a database \( x \) is not too different from the likelihood of obtaining the same answer on adjacent databases, i.e. databases which differ from \( x \) for only one individual.

In [17], [16], we have analyzed critically the notion of differential privacy in light of the conceptual framework provided by the Rényi min information theory. We have shown that there is a close relation between differential privacy and leakage, due to the graph symmetries induced by the adjacency relation. Furthermore, we have considered the utility of the randomized answer, which measures its expected degree of accuracy. We have focused on certain kinds of utility functions called “binary”, which have a close correspondence with the Rényi min mutual information. Again, it turns out that there can be a tight correspondence between differential privacy and utility, depending on the symmetries induced by the adjacency relation and by the query. Depending on these symmetries we can also build an optimal-utility randomization mechanism while preserving the required level of differential privacy. Our main contribution is a study of the kind of structures that can be induced by the adjacency relation and the query, and how to use them to derive bounds on the leakage and achieve the optimal utility.

6.2. Concurrent constraint programming

6.2.1. Bisimilarity

Bisimilarity is one of the main representative equivalences for concurrent behaviour. It captures our intuitive notion of process equivalence; two processes are equivalent if they can match each other’s moves. Furthermore, it provides an elegant co-inductive proof technique based on the notion of bisimulation. Nevertheless, there have been few attempts to define a notion of bisimilarity for concurrent constraint programming (ccp). The ones we were aware of are those in [40] and [36] but they are not completely satisfactory: The first one may tell apart processes with identical observable behaviour, while the second quantifies over all possible inputs from the environment, and hence it is not clear whether it can lead to a feasible proof technique.

Bisimilarity relies on labelled transitions: each evolution step of a system is tagged by some information aimed at capturing the possible interactions of a process with the environment. In [18] we have provided a labelled transition system for ccp and we have proposed a notion of ccp bisimilarity. Intuitively, in this transition system the labels represent the minimal information that processes require from the environment to execute. Furthermore we have shown that, unlike previous approaches, our notion of bisimilarity coincides with the standard notion of equivalence for (deterministic) ccp. This way we have provided ccp with an alternative co-inductive proof technique, coherent with previous equivalences, for process behaviour.

When the state space of a system is finite, the ordinary notion of bisimilarity can be computed via the well-known partition refinement algorithm, but unfortunately, this algorithm does not work for ccp bisimilarity. In [19] we have proposed a variation of the partition refinement algorithm for verifying ccp bisimilarity. To the best of our knowledge this is the first work providing for the automatic verification of program equivalence for ccp.
6.2.2. Modeling cellular signaling systems

The molecular mechanisms of cell communication with the environment involve many concurrent processes governing dynamically the cell function. This concurrent behavior makes traditional methods, such as differential equations, unsatisfactory as a modeling strategy since they do not scale well when a more detailed view of the system is required.

In [19] we have described a modeling strategy for cellular signaling systems based on a temporal and probabilistic extension of CCP. Starting from an abstract model, we have built refinements adding further details coming from experimentation or abstract assumptions. The advantages of our approach are: due to the notion of partial information as constraints in CCP, the model can be straightforwardly extended when more information is available; qualitative and quantitative information can be represented by means of probabilistic constructs of the language; finally, the model is a runnable specification and can be executed, thus allowing for the simulation of the system. We have outlined the use of this methodology to model the interaction of G-protein-coupled receptors with their respective G-proteins that activates signaling pathways inside the cell. Finally, we have presented simulation results obtained from an implementation of the framework.

6.3. Session types

In [22] we have presented a type checking algorithm for establishing a session-based discipline in the pi calculus of Milner, Parrow and Walker [37]. Our session types are qualified as linear or unrestricted. Linearly typed communication channels are guaranteed to occur in exactly one thread, possibly multiple times; afterwards they evolve as unrestricted channels. Session protocols are described by a type constructor that denotes the two ends of one and the same communication channel. We have proved the soundness of the algorithm by showing that processes consuming all linear resources are accepted by a typing system preserving typings during the computation and that type checking is consistent w.r.t. structural congruence.
6. New Results

6.1. Introduction

This section presents the results obtained by Compsys in 2011. For clarity, some earlier results are also recalled, when they were continued or extended during the year 2011.

6.2. Studying Optimal Spilling in the Light of SSA

Participants: Florian Brandner, Quentin Colombet, Alain Darte.

Recent developments in register allocation, mostly linked to static single assignment (SSA) form, have shown that it is possible to decouple the problem in two successive phases: a first spilling phase places load and store instructions so that the register pressure at all program points is small enough, a second assignment and coalescing phase maps the remaining variables to physical registers and reduces the number of move instructions among registers. We focused on the first phase, for which many open questions remained: in particular, we studied the notion of optimal spilling (what can be expressed?) and the impact of SSA form (does it help?).

To identify the important features for optimal spilling on load-store architectures, we developed a new integer linear programming formulation, more accurate and expressive than previous approaches. Among other features, we can express SSA \(\phi\)-functions, memory-to-memory copies, and the fact that a value can be stored simultaneously in a register and in memory. We implemented this formulation in LAO and analyzed in details the static and dynamic results obtained for the SPEC INT 2000 and EEMBC 1.1 benchmarks. We can draw, among others, the following conclusions: a) rematerialization is extremely important, b) SSA complicates the formulation of optimal spilling, especially because of memory coalescing when the code is not in CSSA, c) micro-architectural features are significant and thus have to be accounted for, d) significant savings can be obtained in terms of static spill costs, cache miss rates, and dynamic instruction counts.

This work has been presented at the CASES’11 conference [8].

6.3. Copy Elimination on Data Dependence Graphs

Participants: Florian Brandner, Quentin Colombet.

Register allocation recently regained much interest due to new decoupled strategies that split the problem into separate phases: spilling, register assignment, and copy elimination.

A common assumption of existing copy elimination approaches is that the original ordering of the instructions in the program is not changed. We worked on an extension of a local recoloring technique that we developed earlier, called Parallel Copy Motion [30]. We perform code motion on data dependence graphs in order to eliminate useless copies and reorder instructions, while at the same time a valid register assignment is preserved. Our results show that even after traditional register allocation with coalescing our technique is able to eliminate an additional 3% (up to 9%) of the remaining copies and reduce the weighted costs of register copies by up to 25% for the SPECINT 2000 benchmarks. In comparison to Parallel Copy Motion, our technique removes 11% (up to 20%) more copies and up to 39% more of the copy costs.

This work will be presented at the conference SAC’12 [6].

6.4. Graph-Coloring and Treescan Register Allocation Using Repairing

Participants: Quentin Colombet, Benoît Boissinot, Philip Brisk [University of California, Riverside], Sebastian Hack [Saarland University], Fabrice Rastello.
Graph coloring and linear scan are two appealing techniques for register allocation as the underlying formalism are extremely clean and simple. Our previous work advocated the use of a decoupled approach that first lowers the register pressure by spilling variables, then performs live-range splitting, coalescing, and coloring in a separate phase. This enables the design of simpler, cleaner, and more efficient register allocators.

In this context, we introduced a new and more general approach to deal with register constraints. This approach, called repairing, does not require any preliminary live-range splitting and does not introduce additional spill code. It ignores register constraints during coloring/coalescing and repairs the violated constraints afterwards. We applied this method to develop both a graph-based and a scan-based decoupled approach: one based on the iterated register coalescer (IRC) and the other on a scan algorithm (the treescan) that uses static single assignment (SSA) properties.

Our experimental evaluation shows that, for the graph-based approach, we reduced the number of vertices (edges) in the interference graph by 26% (33%) without compromising the quality of the generated code. The treescan algorithm improved the compile time of the allocation process by $6.97 \times$ over IRC while providing comparable results for the quality of the generated code.

This work was part of a collaboration with the Saarland University and the University of California, Riverside. It has been presented at the conference CASES’11 [7].

### 6.5. Decoupled Graph-Coloring Register Allocation with Hierarchical Aliasing

**Participants:** Andre Tavares [UFMG, Brazil], Quentin Colombet, Mariza Bigonha [UFMG, Brazil], Christophe Guillot [STMicroelectronics], Fernando Pereira [UFMG, Brazil], Fabrice Rastello.

Decoupling spilling from register assignment, as mentioned in previous sections, has the main advantage of simplifying the implementation of register allocators. However, the decoupled model faces many problems when dealing with register aliasing, a phenomenon typical in architectures usually seen in embedded systems, such as ARM.

We introduced the semi-elementary form, a program representation that brings decoupled register allocation to architectures with register aliasing. The semi-elementary form is much smaller than program representations used by previous decoupled solutions, which leads to register allocators that perform better in terms of time and space. Furthermore, this representation reduces the number of copies that traditional allocators insert into assembly programs. We have empirically validated our results by showing that how our representation improves two well-known graph-coloring-based allocators, namely the iterated register coalescer (IRC) and Bouchez et al.’s brute force (BF) method, both augmented with Smith et al. extensions to handle aliasing. Running our techniques on SPEC CPU 2000, we have reduced the number of nodes in the interference graphs by a factor of 4 to 5, hence speeding-up the allocation time by a factor of 3 to 5. In addition, the semi-elementary form reduces by 8% the number of copies that IRC leaves uncoalesced.

This work is part of a collaboration with the Federal University of Minas Gerais. It has been presented at SCOPES’11 [13].

### 6.6. A Non-Iterative Data-Flow Algorithm for Computing Liveness Sets in Strict SSA Programs

**Participants:** Benoît Boissinot, Florian Brandner, Alain Darte, Benoît Dupont de Dinechin [Kalray], Fabrice Rastello.

We revisited the problem of computing liveness sets, i.e., the sets of variables live-in and live-out of basic blocks, for programs in strict static single assignment (SSA). In strict SSA, aka SSA with dominance property, the definition of a variable always dominates all its uses. We exploited this property and the concept of loop-nesting forest to design a fast two-phase data-flow algorithm: a first pass traverses the control-flow graph (CFG), propagating liveness information backwards, a second pass traverses the loop-nesting forest, updating liveness sets within loops. The algorithm is proved correct even for irreducible CFGs.
We analyzed its algorithmic complexity and evaluated its efficiency on SPEC INT 2000. Compared to traditional iterative data-flow approaches, which perform updates until a fixed point is reached, our algorithm is 2 times faster on average. Other approaches are possible that propagate from uses to definitions, one variable at a time, instead of unioning sets as in data-flow analysis. Our algorithm is 1.43 times faster than the fastest alternative on average, when sets are represented as bitsets and for optimized programs, i.e., when there are more variables and larger live-sets and live-ranges.

This work has been presented at the conference APLAS’11 [5].

6.7. SSI Revisited: A Program Representation for Sparse Dataflow Analyses

**Participants:** Andre Tavares [UFMG, Brazil], Mariza Bigonha [UFMG, Brazil], Roberto Bigonha [UFMG, Brazil], Benoit Boissinot, Fernando Pereira [UFMG, Brazil], Fabrice Rastello.

Dataflow analyses usually associate information about variables to program regions. Informally, if these regions are too small, e.g., a point between two consecutive statements, we call the analysis dense. On the other hand, if these regions include many such points, then we call it sparse. We developed a systematic method to build program representations that support forward and/or backward sparse analyses. To pave the way that leads to this framework, we first clarified the literature on intermediate program representations. We revisited the static single information (SSI) form introduced in the 90s and showed how to simplify the construction of program representations for unidirectional dataflow analyses. We showed how to cope with live-ranges that have multiple uses/definitions without losing the equivalence property with the initial dataflow analysis problem. This allows us to simplify, for unidirectional problems, the SSI construction algorithm.

We also showed that our approach, up to a parameter choice, subsumes other program representations such as the SSA, SSI, and e-SSA forms. We can produce intermediate representations isomorphic to the sparse evaluation graphs (SEGs) of Choi et al. This data structure enables sparse solutions to the class of dataflow problems called partitioned dataflow analysis (PDA). However, contrary to SEGs, we can handle - sparsely - problems that are not PDA. We have implemented this framework in the LLVM compiler and have empirically compared different program representations in terms of size and construction time.

This work is part of a collaboration with the Federal University of Minas Gerais and is under reviewing process for the journal Science of Computer Programming.

6.8. Incremental Spilling

**Participants:** Albert Cohen [Inria, Parkas], Boubacar Diouf [Université Paris Sud, Parkas], Fabrice Rastello.

This work addresses the minimization of the spill code overhead in the contexts of both coupled and decoupled register allocation. We devised a heuristic approach called *stacking*; it incrementally allocates clusters of variables, as opposed to the conventional incremental spilling approach. We describe two polynomial methods, a stacking-optimal allocator and a greedy stacking-independent-set allocator. The first method is very close to the optimal allocation; the second method outperforms state-of-the-art heuristics for just-in-time compilation.

This work has been submitted for publication.

6.9. Program Analysis and Communication Optimizations for HLS

**Participants:** Christophe Alias, Alain Darte, Alexandru Plesco.

High-level synthesis (HLS) tools are now getting more mature for generating hardware accelerators with an optimized internal structure, thanks to efficient scheduling techniques, resource sharing, and finite-state machines generation. However, interfacing them with the outside world, i.e., integrating the automatically-generated hardware accelerators within the complete design, with optimized communications, so that they achieve the best throughput, remains a very hard task, reserved to expert designers. The goal of our research on HLS is to study and to develop source-to-source strategies to improve the design of these interfaces, trying to consider the HLS tool as a back-end for more advanced front-end transformations.
Using the C2H HLS tool from Altera, which can synthesize hardware accelerators communicating to an external DDR-SDRAM memory, we showed that it is possible to automatically restructure the application code, to generate adequate communication processes in C, and to compile them all with C2H, so that the resulting application is highly-optimized, with full usage of the memory bandwidth.

These transformations and optimizations, which combine techniques such as double buffering, array contraction, loop tiling, software pipelining, among others, were incorporated in an automatic source-to-source transformation tool, called CHUBA (see Section 5.7), based on the polyhedral model representation. Our study shows that HLS tools can indeed be used as back-end optimizers for front-end optimizations, as it is the case for standard compilation with high-level transformations developed on top of assembly-code optimizers. We believe this is the way to go for making HLS tools viable. The complete automation of the process will be presented at PPoPP'12 [3] and Impact’12 [15].

We also showed how to extend this method to programs with irregular control and array accesses. The main difficulty arises when some data may be redefined in the accelerator but this is not sure. We showed that techniques based on parametric polyhedral optimizations can be used to generate the sets of data to be loaded (resp. stored) just before (resp. after) each tile. An interesting feature is that the previous method appears nicely as a particular case when no approximation is needed. This work is fully described in a research report [23], but is not yet published in a conference or journal.

6.10. Compilation of Hardware Accelerators with Pipelined Arithmetic

Participants: Christophe Alias, Bogdan Pasca [PhD student, Arénaire Inria Team], Alexandru Plesco.

By nature, source-level optimizations cannot perform fine optimizations of the datapath and the control, sometimes mandatory to obtain performances. In high-level synthesis, the circuit generated must be efficient and must produce quality results. This last point is the specialty of the Arénaire Inria-team, which develops the tool FLoPoCo, an open-source FPGA-specific generator of pipelined floating-point arithmetic operators, from a functional description as \((x, y, z) \mapsto e^x + y \times \cos(z)\). The user can specify the precision (mantissa, exponent) and the maximum frequency, then FLoPoCo generates the corresponding operator.

We have developed an algorithm to automatically generate, from a C program, an hardware accelerator that efficiently uses these pipelined operators. The main issue is to reschedule the initial program execution in order to keep the operator’s pipeline as busy as possible, while minimizing memory access. Then, the new execution schedule is used to generate the VHDL code of finite state machines (FSM) controlling the data-flow through the arithmetic operator. This work has been published at the conference ARC’11 [4].

We also showed how our method can be used as a tool to generate control FSMs of multiple parallel computing cores accelerating the same application [25]. This work has been submitted to the journal Microprocessors and Microsystems.

This is still a work in progress and many issues need to be addressed, for example how to extend the program model to general nested loops with more general dependences. These extensions will require to handle properly the communications between the operators and temporary buffers. We believe that the array contraction technique developed in Compsys can be helpful in this context too.

6.11. FPGA Optimized Table Maker’s Dilemma Architecture

Participants: Alexandru Plesco, Florent De Dinechin [Assistant Prof., ARENAIRE Inria Team], Jean-Michel Muller [Research Director, ARENAIRE Inria Team], Bogdan Pasca [PhD student, ARENAIRE Inria Team].

In this work, with some members of the Arénaire team, we developed an algorithm that enables to perform the table maker’s dilemma on a very regular architecture such as an FPGA. The algorithm is crucially different from the algorithm implemented on a standard PC and exploits efficiently the FPGA optimized high-performance arithmetic operators.
The core component (TaMaDi core) of our design is the polynomial evaluator based on the tabulated differences method that uses Remez algorithm. We instantiated multiple TaMaDi cores that work on $2^n$ disjoint intervals obtained from the splitting of the input interval. The TaMaDi cores are connected using a pipelined communication architecture based on the credit communication methodology. We focused on the communication architectures internal to a FPGA that can scale with limited impact on the frequency of the generated data processing architecture. We defined a new pipelined credit based communication interface that leads to a full-rate pipeline data transmission with the cost of small transfer initialization latency. This design has been proven to be effective on parallel design of TaMaDi FPGA specific algorithms. This strategy can be reused in any parallel FPGA design.

This work has been published at ASAP’11 [14] where it received a best paper award.

6.12. Termination of Big Programs

Participants: Christophe Alias, Laure Gonnord, Guillaume Andrieu [Undergraduate Internship, Polytech’Lille].

In a previous work with Alain Darte and Paul Feautrier, we showed how to prove the termination of a certain class of irregular programs with while loops [29]. Our technique amounts to compute a sequential schedule -- called ranking function -- for the program, reinvesting most of the techniques from [37] to schedule static loops. Our termination method is based on the resolution of a linear programming instance, and does not scale. In most of big programs, all the information is not relevant for proving termination. Only a few slices need a termination proof. Moreover, inside a loop nest, termination can be proved incrementally. Inner loops are proved to terminate, and then are replaced by a summary. Our contribution is thus a reduction to prove a single loop termination, the price being the approximation made while computing the summary of a given (nest of) loop(s). This method can also be used to schedule irregular programs, which is clearly a need for program optimization.

A prototype has been designed and is currently under testing. A paper is in preparation. This method has been developed during the undergraduate internship of Guillaume Andrieu, from the Engineering School Polytech’Lille.

6.13. Simplification of Boolean Affine Formulas

Participant: Paul Feautrier.

Boolean Affine Formulas, in which affine inequalities are combined by boolean connectives, are ubiquitous in computer science: static analysis, code and hardware generation, symbolic model checking and many other techniques use them as a compact representation of large or infinite sets. Common algorithms tend to generate large and highly redundant formulas, hence the necessity of a simplifier for keeping the overall complexity under control. Simplification is a difficult problem, at least as hard as SMT solving, with a worst case complexity exponential in the number of affine inequalities. Paul Feautrier has proposed a new method, based on path cutting in Ordered Binary Decision Diagrams, which is able to take advantage of any regularity in the subject formula to speed up simplification. The method has been implemented and tested on benchmarks from several application domains.

The method has been presented at the 4th “Rencontres de la communauté française de compilation”. A detailed description has been submitted for publication; see also [28].

6.14. Retiming for Faust

Participants: Alain Darte, Alexandre Isoard [Master 1 student, ENS-Lyon], Yann Orlarey [Grame].

Faust (Functional Audio Stream) is a formal specification stream-like language designed for real-time signal processing and synthesis. Faust programs are compiled into equivalent C++ programs and optimized for parallel execution. One of the core optimizations in Faust is a preliminary phase called normalization, which amounts to change the delays between operators into an equivalent normalized form used for, among others, redundancy elimination.
We showed that this normalization is actually a special form of retiming, a well-known technique in circuit design and loop transformations. However, an important subtlety needs to be considered: the problem of initialization of signals, which is usually not even mentioned in the retiming literature. We proved that the problem comes from “time-dependent” operators in Faust and that all these operators can be decomposed into combinations of regular operators (in the sense that standard retiming applies) and a single elementary time-dependent operator, called init, which transfers any signal with no modification for time $t \geq 0$ and outputs a constant for time $t < 0$.

This collaboration between Compsys and Grame did not go beyond the Master internship of A. Isoard yet and did not lead to an actual implementation within Faust.
6. New Results

6.1. Constraint Handling Rules and Linear Logic

**Participants:** François Fages, Thierry Martinez, Sylvain Soliman.

Implementations of Constraint Handling Rules (CHR) follow a committed-choice forward chaining execution model: the non-determinism of the abstract semantics is partly refined with extra-logical syntactic convention on the program order and possibly notations for weighted semantics (with priorities or probabilities), and partly left unspecified in the underlying compiler. In [13], we propose an alternative execution model which explores all the possible choices, by opposition to the committed-choice strategy. This execution model is angelic in the sense that if there exists a successful execution strategy (with respect to a given observable), then this strategy will be found. Formally, the set of computed goals is complete with respect to the set of the logical consequences of the interpretation of the initial goal in linear logic. In practice, this paper introduces a new data representation for sets of goals, the derivation nets. Sharing strategies between computation paths can be defined for derivation nets to make execution algorithmically tractable in some cases where a naive exploration would be exponential. Control for refined execution is recovered with the introduction of user constraints to encode sequencing, fully captured in the linear-logic interpretation. As a consequence of angelic execution, CHR rules become decomposable while preserving accessibility properties. This decomposability makes natural the definition in angelic CHR of meta-interpreters to change the execution strategy. More generally, arbitrary computation can be interleaved during head matching, for custom user constraint indexation and deep guard definition.

6.2. Rule-based Modeling Language for Constraint Programming

**Participants:** François Fages, Raphaël Martin, Thierry Martinez, Sylvain Soliman.

Rules2CP is a rule-based modeling language which allows easy modeling of constraint satisfaction problems, together with specifications for search strategies and heuristic choice criterias by pattern matching. In [23], we study a new compilation scheme for Rules2CP which allows us to deal with dynamic ordering criteria and to generate procedural constraint programming code instead of flattened constraints. The comparison with the static expansion of Rules2CP models shows that the overhead at runtime is limited, with a gain in the size of the generated program which could be exponentially larger by static expansion.

The language Rules2CP is currently extended to deal with hybrid discrete and continuous domains and packing problems with complex shapes, in the framework of the ANR Net-WMS-2 project with KLS-Optim and EMN Inria EPI TASC. The compiler of Rules2CP is currently rewritten in Java in the framework of a collaboration with KLS-Optim supported by Inria DTI.

6.3. Trace Development Methodology

**Participants:** Pierre Deransart, Armando Gonçalves da Silva Junior.

We are working on a general theory of traces design taking traces as primary objects of study. It is based on the observation of the way trace files are accumulated as knowledge bases and elaborated in different fields of activity like software engineering, rule based systems and resolution, learning in context, or personal experience storing systems.

We worked on two main points: the development of an experimental tracer of CHR\(^e\) [11] (see TODAS project) and an application the notion of generic trace to standardization of constraints. In [17] we analyze, and occasionally correct, shortcomings of the former approach based on the generic trace format GENTra4CP, and show the interest that a generic tracer may bring to develop portable applications or to standardization efforts, in particular in the field of constraints.
6.4. Railway Time Tabling Optimization

Participants: François Fages, David Fournier, Sylvain Soliman.

Sustainable development is a key issue for our society. Optimization of resources, energy and costs (as admirably done in living organisms) has thus grown significantly over years to become a major field in industry. In collaboration with General Electric Transportation France which is a key-actor in the field of transportation all around the world, we investigate energy reduction for train and metro service providers through time tabling optimization. In [31], we describe and compare different optimization methods to reduce energy in mass rapid transits (MRT). Most of the literature deals with a special problematic arising in train services: the maximum traction energy. They show that for reducing costs and energy consumption, one method is the reduction of the peak energy over a time period. This objective function has been chosen for this study and the thesis depicts how to implement it on different paradigms, such as mixed-integer linear programming, constraint programming or local search. We conclude on promising approaches in terms of optimization methods for time tabling computation and real-time scheduling.

6.5. Petri Net Analysis of Biochemical Networks

Participants: François Fages, Thierry Martinez, Faten Nabli, Sylvain Soliman.

Bridging the gap between quantitative and qualitative models, Petri nets (also known as place/transition graphs) have recently emerged as a promising tool for modeling and analysis of biochemical networks. In [14], we present a method to compute the minimal siphons and traps of a Petri net as a Constraint Satisfaction Problem (CSP). In our case, siphons and traps are purely structural properties that brings us information about the persistence of some molecular species. We present a constraint program that finds minimal siphons and traps containing specific set of places in a Petri net. This method is compared on models of the biomodels.net repository with other methods based on Mixed Integer Linear Programming (MILP) and Boolean Satisfiability (SAT). The flexibility brought by constraint programming, for instance in the declarative choice of variable enumeration heuristics, seems promising in further improving those results.

6.6. Theory of Subgraph Epimorphisms

Participants: François Fages, Steven Gay, Thierry Martinez, Sylvain Soliman.

The operations of deleting and merging vertices are natural operations for reducing a graph. While graph reductions through a sequence of vertex deletions (resp. mergings) characterize subgraph isomorphisms (resp. graph epimorphisms), sequences of both vertex deletion and merging operations characterize subgraph epimorphisms. Our proposal is thus to use subgraph epimorphism for comparing graphs in applications where a more flexible notion than the classical notion of subgraph isomorphism is required.

In collaboration with Christine Solnon (INSA Lyon), we have developed the theory of subgraph epimorphisms in [5]. We have shown that SEPIs preserve graph completeness and arc symmetry and that, just like SISO and EPI, SEPI is not a well quasi order. We have defined the SEPI, EPI and SISO distances between two graphs as the size of the largest SEPI (resp. EPI, SISO) lower bound graphs. These distances are equal to the minimum number of respectively vertex deletion and/or merging operations that are necessary to obtain isomorphic graphs. They are also metrics on graphs and we have $d_d \geq d_{md}$ and $d_m \geq d_{md}$.

From a computational point of view, we have shown that the existence of a SEPI between two graphs is an NP-complete problem and have presented a constraint satisfaction algorithm for solving it. This algorithm is implemented in BIOCHAM.

It is worth noticing that, given two graphs $G$ and $G'$, the greatest lower SEPI bounds and the least upper SEPI bounds are also interesting to compute since they represent “intersection” and “union” graphs for the SEPI relation. For instance, in our motivating application in systems biology, these objects correspond to the intersection (resp. union) of models at different levels of details for a given biochemical process. These graphs are not unique but we are confident that the constraint satisfaction algorithm presented in [10] can be interestingly generalized to compute them.
6.7. Parameter Search under Temporal Logic Constraints

**Participants:** Grégory Batt, Elisabetta De Maria, François Fages, Domitille Heitzler, Aurélien Rizk, Sylvain Soliman, Jannis Uhlendorf.

Our method for solving temporal logic constraints in the quantifier-free fragment of first-order linear time logic QFLTL(R), opens up the field of model-checking to optimization through the definition of a continuous degree of satisfaction for temporal logic formulae \[8\], \[2\]. This satisfaction degree can be used in a number of ways, e.g. as a fitness function with continuous optimization methods\(^2\) to find unknown parameter values in a model with temporal logic constraints formalizing biological properties \[4\], \[6\], or to control a system from a temporal specification of its behavior \[15\], or to compute the robustness of a system w.r.t. a temporal property and a perturbation of the parameters.

This approach is implemented in BIOCHAM and is one unique feature of this modeling environment.

6.8. Model-based Optimization of Cancer Chronotherapies

**Participants:** Elisabetta De Maria, François Fages, Aurélien Rizk, Sylvain Soliman, Denis Thieffry.

Recent advances in cancer chronotherapy techniques support the evidence that there exist some links between the cell cycle and the circadian clock genes. One purpose for modeling the entrainment in period of the cell cycle by the circadian clock is to better understand how to efficiently target malignant cells depending on the phase of the day and patient characteristics. This is at the heart of our participation in collaboration with the EPI BANG in the EraNet SysBio project C5Sys, follow up of the former EU STREP project TEMPO. In \[4\] we show how temporal logic constraints, and the new features of BIOCHAM for parameter search (running on a cluster of 10000 processors at the GENCI) can be used to couple dynamical models in high dimension and more precisely to build a coupled model composed of:

- a four phases model of the mammalian cell cycle by Novak and Tyson,
- a circadian clock model by Leloup and Goldbeter,
- a DNA damage repair model by Ciliberto et al.,
- a model of irinotecan metabolism by Dimitrio and Ballesta,
- a simple model of drug administration control.

This coupled model allows us to minimize the toxicity of irinotecan on healthy cells, using BIOCHAM’s parameter search method applied on the drug administration control law.

Our technology is ready to calibrate models on real patient data, evaluate model predictions and optimize patient-tailored chronotherapeutics. The collaboration currently focuses on the obtaining of consistent data in the C5Sys project and on the improvement of the cell cycle model.

6.9. Analysis of FSH and Angiotensine Signaling

**Participants:** François Fages, Domitille Heitzler, Aurélien Rizk, Sylvain Soliman.

In \[6\] in collaboration with Eric Reiter (UMR CNRS-INRA 6175) and Frédérique Clément (SISYPHE) in the framework of the Initiative Action REGATE, we have combined experimental approaches with computational modeling to decipher the molecular mechanisms as well as the hidden dynamics governing ERK activation by the angiotensin II type 1A receptor (AT1AR) in HEK293 cells. We have built in BIOCHAM a dynamical model that captures available knowledge and experimental data. The unknown kinetic parameters have been inferred using a temporal logic specification of experimental data in both control and perturbed conditions, using a cluster of 10000 processors at the GENCI.

\(^2\) we use the Covariance Matrix Adaptation Evolutionary Strategy CMAES of Nikolaus Hansen from the EPI TAO. Moreover, this year we have implemented a second method by Particle Swarm Optimization, PSO.
The mathematical model predicts and experiments confirm that, for the AT1AR expressed in HEK293 cells: i) GRK2/3 and 5/6 regulate switching between the G protein and β-arrestin pathways as well as their distinct dynamics by phosphorylating the C-terminal region of the activated receptor; ii) GRK2/3 not only mediates desensitization of G protein activation but also exerts a strong restraining influence on β-arrestin signaling; iii) GRK5/6 exert little effect on G protein-stimulated ERK but are required for β-arrestin-mediated ERK activation; iv) the β-arrestin-dependent ERK pathway undergoes both activation and deactivation through amplified enzymatic processes.

These results convincingly illustrate the value of using computational modeling to decipher the complex signaling mechanisms elicited by 7TMRs [1]. This approach is applied more generally to G protein-coupled receptor signaling which is of great importance in pharmacology.

6.10. Multi-affine Hybrid Automaton Model of Cardiac Cells

Participant: Grégory Batt.

A fundamental question in the treatment of cardiac disorders, such as tachycardia and fibrillation, is under what circumstances does such a disorder arise? To answer to this question, in collaboration with E. Bartocci and R. Grosu at SUNY, Stony Brook, USA, we develop a multi-affine hybrid automaton (MHA) cardiac-cell model, and restate the original question as one of identification of the parameter ranges under which the MHA model accurately reproduces the disorder [12]. The MHA model is obtained from the minimal cardiac model of one of the authors, Fenton from Cornell University, by first bringing it into the form of a canonical genetic regulatory network, and then linearizing its sigmoidal switches in an optimal way. By leveraging the Rovergene tool for genetic regulatory networks, we are then able to successfully identify the parameter ranges of interest.

6.11. Real-time Control of Gene Expression in Yeast

Participants: Grégory Batt, François Fages, Jannis Uhlendorf.

To decipher the dynamical functioning of cellular processes, the method of choice is to observe the time response of cells subjected to well controlled perturbations in time and amplitude. Efficient methods, based on molecular biology, are available to monitor quantitatively and dynamically many cellular processes. In contrast, it is still a challenge to perturb cellular processes - such as gene expression - in a precise and controlled manner. In collaboration with Pascal Hersen at MSC lab (Paris Diderot University), in the framework of the Iceberg ANR project, we propose a first step towards in vivo control of gene expression: in real-time, we dynamically control the activity of a yeast signaling cascade thanks to an experimental platform combining a micro-fluidic device, an epi-fluorescence microscope and software implementing control approaches [15]. We experimentally demonstrate the feasibility of this approach, and we investigate computationally some possible improvements of our control strategy using a model of the yeast osmo-adaptation response fitted to our data.

6.12. Artificial Tissue Homeostasis in Mammalian Cells

Participants: Grégory Batt, François Bertaux, Xavier Duportet, François Fages, Szymon Stoma.

Cell-based gene therapy aims at creating and transplanting genetically-modified cells into a patient in order to treat a disease. Ideally, actively-growing cells are used to form a self-maintaining tissue in the patient, thus permanently curing the disease. However, before any real therapeutic use, robust mechanisms enforcing tissue homeostasis, that is, that the size of the newly-introduced tissue remains within admissible bounds, need to be developed. We proposed various designs and tested their robustness using in silico approaches. Preliminary results demonstrated that cell-to-cell variability plays a crucial role for tissue long-term maintenance. More extensive in silico characterizations require the development of efficient multiscale simulation methods. In parallel to the in silico work, done in collaboration with the Bang research group (Dirk Drasdo), we started the construction and in vitro experimental characterization of the most promising designs in collaboration with the Weiss lab (MIT) [24].
5. New Results

5.1. Co-Modeling for HP-SoC with MARTE

5.1.1. Diagonal mesh modeling with MARTE

As a continuation of this work on modeling at system level, a methodology for modeling concepts of NoC-based architectures is proposed especially the modeling of all kinds of topologies (regular, irregular or hierarchical) and routing algorithms. This contribution includes a VHDL code generation. On the other side we proposed a VLSI implementation of a new NoC topology called diagonal mesh that it designed to offer a good tradeoff between hardware cost and theoretical quality of service (QoS). This NoC is based on a new router architecture called FeRoNoC (Flexible, extensible Router NoC).

5.1.2. MARTE extension for reconfigurable hardware models

Reconfigurable System-on-Chip (RSoC), mainly FPGAs, offer several advantages such as flexibility, adaptivity and especially their capability to allow switching several implementations at run-time, i.e., PDR.

PDR feature requires multiple run-time changes in RSoC such as:
- QoS factors: changes in executing functionalities due to designer requirements, or changes due to resource constraints of targeted hardware/platforms.
- The changes can also take place due to other environmental criteria such as communication quality, time and area consumed for reconfiguration and energy consumption.

In previous work [86], we provided an initial contribution to the modeling of these systems by extending UML MARTE profile to incorporate significant design criteria such as power consumption. Furthermore, high flexibility of RSoC implies high design complexity of the control of such system. This makes designing a robust control for managing reconfiguration a studious task. In [25], we present a high level design approach using UML MARTE for modeling dynamic reconfiguration controllers. Our proposed controller is based on distributed monitoring of runtime changes and distributed decision making. Our approach allows to increase flexibility and design reusability compared to centralized solution.

Indeed, in its current version, UML MARTE profile lacks dynamic reconfiguration concepts and requirements for the reconfiguration control mechanism. Even these later are necessary to model and implement rapid prototypes for complex systems. We can only model a state machine at high abstraction levels which is responsible for switching between the available configurations.

So we define a new design methodology using the proposed version of RecoMARTE (extended MARTE) to model PDR concepts at different abstraction levels, mainly architecture (structural and physical models) and allocation (software to Hardware allocation (Sw/Hw Allocate) and Hardware to Hardware allocation (Hw/Hw Allocate)). We also define necessary requirements for the reconfiguration control mechanism in order to manage reconfiguration at every design level. In addition, our solution allows to describe global contracts and constraints for combining automata. As future works, we plan to carry out model transformations to enable automatic code generation of configuration files. The code can then be used as input for commercial tools for final FPGA synthesis.

5.1.3. Comparaison of SAC and ArrayOL for parallelism expression

In this join work with the University of Hertfordshire, we compare and analyse two such schemes. One of them is a domain-specific language, ArrayOL, to OpenCL. The other one is a transformation mechanism for mapping a image/signal processing transformation route for mapping a high-level general purpose array processing language, Single Assignment C (SaC) to CUDA. Using a real-world image processing application as a running example, we demonstrate that albeit the fact of being general purpose, the array processing language be used to specify complex array access patterns generically. Performance of the generated CUDA code is comparable to the OpenCL code created from domain-specific language.
5.1.4. Gaspard Modeling Improvements

Gaspard2 is the IDE proposed by the DaRT team. Its usage can be painful for beginners as well as for experts. We try to improve the usage of Gaspard in different ways:

- By allowing modifications at any model level, and let propagate the modifications to the higher and lower models (Amen Souissi).
- By providing missing diagrams in Papyrus (Amine El Kouen) By customizing the Gaspard User Interface (UI). Modeling in Gaspard is done with the Papyrus Modeler. We participate to the Papyrus development, which allow us to propose some customization tools. These later, are used to provide a modeling UI more adapted to embedded system co-modeling. This work is done by Rahma Yangui (INRIA engineer).
- By allowing to adapt dynamically the UML modeler environment according to the steps of the modeling process (Amine El Kouen’s thesis). This allows to guide the user in its development process, and to propose a simplified UI, oriented to the current development step.

Also, we have migrated from Papyrus I to Papyrus Eclipse.

5.2. Formal Methods for general-purpose and domain-specific languages

5.2.1. Formal Semantics for Domain Specific Modeling Languages

Domain-Specific Modelling Languages (DSMLs) are languages dedicated to modelling in specific application areas. Recently, the design of DSMLs has become widely accessible to engineers trained in the basics of Model-Driven Engineering (MDE): one designs a metamodel for the language’s abstract syntax; then, the language’s operational semantics is expressed using model transformations over the metamodel.

The democratisation of DSM design catalysed by MDE is likely to give birth to numerous languages. One can also reasonably expect that there shall be numerous errors in those languages. Indeed, getting a language right (especially its operational semantics) is hard, regardless of whether the language is defined in the modern MDE framework or in more traditional ones.

Formal approaches can benefit language designers by helping them avoid or detect errors. But, in order to be accepted by nonexpert users, formal approaches have to operate in the background of a familiar language design process, such as the MDE-based one mentioned above.

In 2011 we have migrated from [21], which uses the general Maude semantic framework, towards using the more language-definition specific the K-semantic framework to formalise the basic MDE ingredients used in DSML definition: models, metamodels, and model transformations. We have implemented a prototype tool that takes as input any DSML described in using MDE, and generates formal K definitions for the language’s syntax, static semantics, and operational semantics. Since the definitions are executable, we get execution and formal verification engines for free [44]. A subproject of this work has been a formal definition for a substantial fragment of the OCL language [45].

5.2.2. A new abstraction for signal programs, and improvement of the compilation process

In this work we propose a sound abstraction for an efficient static analysis of synchronous programs describing multi-clock embedded systems in Signal. This abstraction combines the Boolean theory and numeric interval approximation to adequately address clock relations defined as combinations of logical and numerical expressions. Through a few examples, we show how the proposed solution is used to determine absence of reaction captured by empty clocks; mutual exclusion captured by two or more clocks whose associated signals never occur at the same time; or hierarchical control of component activations via clock inclusion. We also show this analysis improves the quality of the code generated automatically by the Signal compiler, e.g., a code with smaller footprint, or a code executed more efficiently thanks to optimizations enabled by the new abstraction [38].
5.2.3. Using bounded model checking to focus fixpoint iterations

Two classical sources of imprecision in static analysis by abstract interpretation are widening and merge operations. Merge operations can be done away by distinguishing paths, as in trace partitioning, at the expense of enumerating an exponential number of paths. In this article, we describe how to avoid such systematic exploration by focusing on a single path at a time, designated by SMT-solving. Our method combines well with acceleration techniques, thus doing away with widenings as well in some cases. We illustrate it over the well-known domain of convex polyhedra [40].

5.2.4. A formal definition of a compiler for the Kermeta metamodeling language in K

Kermeta [109] is a DSL designed as a kernel for metamodel engineering. It unifies metamodeling, constraints, semantics and transformation features into a statically typed language. It is object-oriented and allows for metamodeling features such as attributes, associations, and multiplicities. It also includes design-by-contract, aspect-oriented features, and genericity. This makes Kermeta a large and complex language: indeed, combining all these features into one language may easily lead to inconsistencies.

Christophe’s postdoctoral work, starting in September 2010, has been to formally specify Kermeta. He did so via a specification of compiler for Kermeta in K [117]. K formal specifications are executable, hence, Christophe’s compiler can be used to actually compile Kermeta programs. The compiler it completely self-contained and generates bytecode for an abstract machine also formally specified in K.

This work led to the discovery of several errors and inconsistencies in Kermeta’s manual and existing interpreter. The errors are reported to the Kermeta designers (Triskell project-team at Inria Rennes-Bretagne Atlantique), who, as it turns out, are also writing a compiler of Kermeta in the traditional, informal way. We are planning to make them benefit from the experience we gained in formal compilation.

5.2.5. A generic approach and tool for tracing executions back to a DSML’s operational semantics

Model-driven engineering allows users to define abstract syntaxes for their own DSMLs in terms of metamodels. Several approaches for defining operational semantics for DSMLs have also been proposed. These approaches allow, in principle, for model execution and for formal analyses of the DSMLs. However, most of the time, the executions/analyses are performed via transformations to other languages: code generation, resp. translation to the input language of a model checker. The consequence is that the results (e.g., a program crash log, or a counterexample returned by a model checker) may not be straightforward to interpret by the users of a DSML. We propose a formal and operational framework for tracing such results back to the original DSML’s syntax and operational semantics. We implement the approach in a generic tool written in Kermeta, and illustrated in on the xSPEM language, a timed language for expressing the execution of activities constrained by time, resources, and precedences [31].

5.3. Optimization and compilation techniques

5.3.1. Generated Code Optimization

Performing a model-to-source transformation, whereby a high-level language is mapped to CUDA or OpenCL, is an attractive option. In particular, it enables to harness the power of GPUs without any expertise on the GPGPU programming. In this work, we add a new compilation option for the Gaspard2 transformation chain: UMi2OpenCL to detect shareable data zone. The tilers from ArrayOL, which allow express the data parallelism from repetitive tasks, are analyzed in time compilation to create areas of shared data. The identification of these areas is crucial to allow us loading data on shared areas of memory that have high throughput. Consequently, programs automatically generated shall have performances comparable to manually well written programs.
5.3.2. Methodology to generate OpenCL code from MARTE models

In order to reduce design complexity, we propose an approach to generate code for OpenCL API, an open standard for parallel programming of heterogeneous systems. This approach is based on Model Driven Engineering (MDE) and Modeling and Analysis of Real-Time and Embedded Systems (MARTE) standard proposed by Object Management Group (OMG). The aim is to provide resources to non-specialist in parallel programming to implement their applications. Moreover, concepts like reuse and platform independence are present. Since we have designed an application and execution platform architecture, we can reuse the same project to add more functionalities and/or change the target architecture. Consequently, this approach helps industries to achieve their time-to-market constraints. The resulting code, for the host and compute devices, are compilable source files that satisfy the specifications defined on design time.

5.3.3. Profiling into Models

Regarding the models fine tuning, we propose integrating software-profiling results to higher-level specification models [56]. The aim is to optimize the models and, consequently, the generated code. The model optimization approach relies on the Gaspard2 branch dedicated to code generation for OpenCL and GPUs [58]. We offer software execution feedback, based on models transformation traceability [75], to model designers. These feedbacks enable the designers to tune their models in order to improve the software performances even if they do not have in-depth knowledge on the running platform (GPU). First, the code is generated from a first designed model using Gaspard2. The resulting code is then executed within an existing profiling environment. Afterwards, profiling results are delivered directly to designer as annotations in the model. Basically, we move up two types of information, using traceability. The first type directly results from the profiler, e.g. processor occupancy, onto specific regions in the model, enlightening the regions that requires tuning. The second type correspond to results of an expert system analysis that we provide. Information of this second type is delivered to designers as advice in the model annotations. The expert system generates these advice from platform features and running results. For example, it can suggest changing the shape of a task in order to optimize the processor occupancy. The more we feed the knowledge base and engine of the expert system, the more it is able to give better advice.

The model optimization relies on the hypothesis that the high level models are error free. Since these models are complex, it is difficult for the designers to conceive them correctly the first time. We propose a new approach, enabling the model designer to debug its models. For this purpose, we offer a quick and automatic code instrumentation to the model designer. As for the model optimisation, we take advantage of the model transformation traceability to keep the link between models and software execution and to provide execution information feedback. Hence, the information produced in the running environment during the software execution is moved up directly onto the models, allowing the model designers verifying the behavior of their software, directly on the high level models.

5.3.4. Static Analysis of Polychronous Specifications with SMT Theory

As opposed to single clocked synchronous programming paradigms, polychronous formalism allows specification of concurrent data flow computation on signals such that various data flows can evolve asynchronous with respect to each other. We formulated the clock analysis in Signal compilation [38] and the detection of false loops in MRICDF as a decision problem in Satisfiability Modulo Theory (SMT) [30][59]. Due to recent interests in SMT solvers, a number of efficient solvers are available which offer a greater expressiveness in dealing with non Boolean constraints and allow us to discard false loops from realizable causalities in reasonable computation time. We demonstrated that several polychronous specifications rejected by current compilers due to their inability to identify only true causal loops, can be synthesized as correct sequential embedded software.

5.3.5. Programming functional and real-time aspects simultaneously

An embedded system is usually required to respect real-time constraints related to physical constraints, either those of its environment or those of the physical devices it controls. First, it is often multi-periodic since its devices have different physical characteristics and must therefore be controlled at different rates. Second,
the system must respect deadline constraints, which may correspond for instance to a maximum end-to-end latency requirement between observations (inputs) and the corresponding reactions (outputs). A correct implementation must respect all the real-time constraints and must also be functionally deterministic, meaning that the outputs of the system are always the same for a given sequence of inputs. Current practice often deals with these two aspects separately, while our objective is to deal with them simultaneously.

To this intent, we must first introduce real-time primitives at the programming language level. We carried on previous work on the PRELUDE language [19], which provides such primitives in a synchronous data-flow language. We produced a complete end-to-end framework for the design and the implementation of embedded systems on a symmetric multicore: the PRELUDE-SCHEDMCORE toolset [32]. We recently started a Master research project to study how real-time aspects could be introduced in more traditional programming paradigms with the SCALA a language.

The PRELUDE compiler translates a program into a set of dependent periodic tasks. We proposed a new dynamic priority-based scheduling policy capable of dealing with the extended precedence constraints (constraints between tasks of different periods) of such systems in [36], [48].

Finally, as PRELUDE'X semantics defines formally both the functional and the temporal behaviour of a system, we studied temporal formal verification in [46].

5.3.6. Chaining Localized Model Transformation

Usually, two transformations can only be chained if the output metamodel of the first one is included into the input metamodel of the second one. This compliance issue forces to design either tailored fine-grain model transformations for a dedicated chain or large and complex transformations. In both cases, transformations are not reusable and hardly maintainable.

In order to solve this problem, we have introduced localized transformations which apply to a (typically very small) subset of an input metamodel of a transformation. Each localized transformation is designed and implemented to accomplish a specific transformation task, and involves and is applicable to a few concepts. The input and output metamodels of these transformations are not disjoint contrary to traditional transformations; new chaining constraints have to be defined. We have thus defined new chaining constraints based on a type analysis to specify when two transformations can be chained in one, both or any order [96]. In some cases, this analysis concludes that the transformations can be chained in both order but with some input models, the two output models resulting of the two chaining, are not the same. We have introduce an intermediary abstraction level independent of any transformation language that focuses on read, modified, created and deleted metaelements. We are pursuing our investigations with this new abstraction level.

5.4. Green computing on SoC

5.4.1. Correct and Energy-Efficient Design of a Multimedia Application on SoC

We studied the design and analysis of multimedia applications such as the JPEG encoder on multiprocessor architectures [55] [24] [13]. A model-based approach was adopted by using the UML Marte specifications [54]. An abstract clock analysis has been proposed to deal with the correctness of system behaviors and to find the most suitable execution platform configurations regarding performance and energy consumption. Our approach offers a rapid and reliable design space analysis, which is crucial when implementing complex systems [37].

5.4.2. Design Space Exploration for Efficient Data Intensive Computing on SoCs

Finding efficient implementations of data intensive applications, such as radar/sonar signal and image processing, on a system-on-chip is a very challenging problem due to increasing complexity and performance requirements of such applications. One major issue is the optimization of data transfer and storage micro-architecture, which is crucial in this context. We proposed a comprehensive method to explore the mapping of high-level representations of applications into a customizable hardware accelerator [52]. The high-level representation is given in a language named Array-OL. The customizable architecture uses FIFO queues and
a double buffering mechanism to mask the latency of data transfers and external memory access. The mapping of a high-level representation onto a given architecture is achieved by applying loop transformations in Array-OL. A method based on integer partition is used to reduce the space of explored solutions. Our proposition aims at facilitating the inference of adequate hardware realizations for data intensive applications. It is illustrated on a case study consisting in implementing a hydrophone monitoring application.

5.4.3. Power Estimation

Within the context of the OPEN-PEOPLE project, we aim at addressing the power estimation challenges of embedded system design with a new approach, combining Functional Level Power Analysis with advanced SystemC – Transaction Level Modeling (TLM) simulation techniques, in order to formally prove qualitative and quantitative properties of the final system power estimation. This approach requires the construction of a power models from FLPA for different embedded boards (FPGA and ASIC) and building up system level simulation environment for the analysis of power model and proofs of properties of the simulated system [42].

As a main contribution, we propose a new hybrid system-level power consumption estimation methodology for complex embedded systems [41]. A key word in our contribution is hybridization between abstraction levels. Almost all the previous studies focus on power estimation for a given abstraction level without overcoming the wall of speed/accuracy trade-off. The idea here is to build up a hybrid power estimation tool that combines Functional Level Power Analysis (FLPA) for hardware power modeling and Transactional Level Modeling (TLM) simulation technique for rapid system prototyping and fast power estimation. Basically, the FLPA is used for processor power modeling. In the frame of this work, it will be extended to cover the other hardware components used in the MultiProcessor System-on-Chip (MPSoC) such as the memory and the reconfigurable logic. After that, we go further in terms of scalability to target heterogeneous multiprocessor architectures. The functional power estimation part is coupled with a fast SystemC simulator in order to obtain the needed micro-architectural activities for power models, which allows us to reach a superior bargain between accuracy and speed[43].

5.5. Dynamic reconfiguration for HP-SoC

5.5.1. Context switching for volatile IP

Dynamic reconfigurations require configuration decisions from smart controllers. Such a decision implies context saving of an existing IP or switching from an IP to another (loading a new bitstream). The store/restore operations can be managed by the operating system or by using a dedicated hardware component. In this work, a new model for hardware IP context storage and management is proposed. The approach is based on a flexible hardware wrapper which can make IP reconfigurable. In fact, these wrappers contain a naming system supporting efficient runtime context switching.

5.5.2. A generic broadcast network for HP-SoC architecture

The hNoC model proposes a specific network on chip dedicated to the massively parallel architecture SCAC. This model is composed of huge number of complex routers, called node elements (the NEs), communicating and working in perfect synchronizations. Each NE is potentially connected to its neighbors via a regular connection. Furthermore, each NE is connected to a heterogeneous set of computing groups (clusters) allow asynchronous processing. Each group includes a combination of processors programmable, the PEs (software processing units) and specialized hardware accelerators (hardware processing units) to perform critical tasks demanding the more performance. All the system is controlled by a Network Controller Unit, the NCU. The NCU and The PEs are implemented with the Forth processor.

The aim of our works is to design a new kind of communication network model for SCAC architecture to overcome firstly the overlapping communications with computations and secondly to increase significantly the external performances in terms of throughput. The difficulty of designing hNoC is a compromise between an optimal quality of broadcasting, high bandwidth and important flexibility of use, while reducing power consumption and silicon area.
Our first contributions defined a broadcast with mask model integrated in the communication network hNoC of SCAC architecture. This model is based on subnetting the network of processing nodes which separate the control of communication and processing. Our model was implemented in synthesizable VHDL code that is simulated and targeted Xilinx Virtex6 (XC6VLX240T) board.

5.5.3. Distributed control for dynamic reconfiguration

The aim of our current work is to propose a distributed approach for reconfiguration control on FPGAs. The main reason for choosing a distributed control approach is that, with the ever growing complexity and size of the modern reconfigurable systems, the traditional centralized approach is no more efficient. Instead, a distributed control has many advantages in terms of performance and design efficiency. Indeed, the distributed control allows to avoid communication bottlenecks and to increase the parallelism compared to the centralized one, allowing a better performance which is a critical issue especially for high-performance applications. At the design level, the distributed control has many advantages. It allows to decrease the design complexity of the control by dividing the intelligence between the controllers, which allows a shorter design time and an easier verification. It also facilitates the reuse of the controllers instead of redesigning a centralized controller for different systems, which allows also a higher scalability in order to adapt to the growing size of the modern SoC. Our approach for reconfiguration control is an event-driven control, where events come from a variety of sources in order to ensure a high adaptability of the reconfigurable systems. Reconfiguration can be triggered by a user input, a change in the environment condition (e.g. changes in lightening condition) or a change in performance or power consumption requirements, etc. Therefore, we propose a modular structure of each controller allowing three major tasks: monitoring, decision making and reconfiguration realization. In order to respect the global constraints of the system, the controllers communicate their decisions to each other in order to handle cooperation and conflicts. In [25], we proposed a high-level design of our approach using Model Driven Engineering aiming to combine the advantages of the distributed control with the high-level design in order to decrease design complexity and automate code generation increasing thus design productivity. At the physical level, the distributed control has been implemented for simple applications in order to test the different modules of the controllers (monitoring, decision making, communication between controllers). As a future work, we plan to implement the distributed control for more complex applications in order to highlight the advantages of our approach and study its limits.

5.5.4. Avionic test bench on heterogenous reconfigurable platform

The aim goal of this thesis is to design the next Eurocopter avionic test bench generation. For the past 20 years, Test Systems have always be considered as a must do in the avionic development cycle. In early 2008, the Eurocopter research department has undertaken a profound reflection on the vocation Pro-Active Test Systems [15]. Hitherto, the test systems were based on real time specific CPU boards that run proprietary real time operating systems and plugged with Input/Output (I/O) boards to communicate with the equipments under test. In current industrial practice, the well-spread VME CPU boards are widely used. Due to the present test system performance requirement, an increase in the computation rates is needed, but it cannot be delivered by the VME CPU boards any-more. Furthermore, this solution is considered as an expensive maintainable technology. To overcome these drawbacks, the usage of multicore hosts (PC or workstation) allows an immediate increase in the capacity of computation. An important outcome of this transition is the refusal of the obsolete CPU boards. However, this solution cannot guarantee the real-time criteria while the execution of concurrent tasks due to the lack of an appropriate Operating System (OS) environment. In addition, this solution brings new communication latencies between the CPUs and I/O boards plugged in the VME backplane. In this work, our proposal is to make profit from the new available hardware computing resource (FPGA) and to make up hybrid avionic test systems [27]. Indeed, FPGA technology could offer a higher computation rates comparing to CPUs up to 10x. It could implement heavy models in a hardware fashion with the management of the parallelism degree to answer the real-time constraints of the application. The main challenge of hybrid (CPU/FPGA) architectures concerns the programming model and the design methodology. We need to deal with the heterogeneity of both hardware and software parts in order to obtain a fast system prototyping. In current industrial practice, manual coding is still widely adopted in the development of hybrid architectures, which is clearly not suited to manage the complexity intrinsic in these systems. For designers,
this approach is very tedious, error-prone and expensive. In the first part of our work we emphasized the usage of Model Driven Engineering (MDE) for heterogeneous systems in order to reduce the design complexity of CPU-FPGA architectures [72]. In ReCoSoC paper, we focused on the prototyping environment and the related development tools in order to map existing software into CPU-FPGA architectures by detecting all data dependencies and get the parallelism degree. Moreover, we presented communication solutions comparing fast links such as Ethernet and PCIe. Secondly Multi-Core optimizations in different environments such as Linux with Open Source real-time patches (Xenomai) and processor affinities capabilities. Then, we presented in [28] a new generation of adaptive and generic avionic test benches using FPGA reconfigurability capabilities. Indeed, nowadays, each Eurocopter test bench is related to a specific embedded part and a specific aircraft. Proposing such generic architecture will reduce the helicopter design cycle significantly by Testing different embedded systems at the same time.

5.6. Application case-studies

5.6.1. Experimentations for electromagnetism simulations

The electrical and electronic engineering has used parallel programming to solve its large scale complex problems for performance reasons. However, as parallel programming requires a non-trivial distribution of tasks and data, developers find it hard to implement their applications effectively. Thus, we used our approach, based on Model Driven Engineering (MDE) and the MARTE profile, to generate code for a sparse solver and achieve a good speed-up. Moreover, thanks to model reuse capacity, we can add/change functionalities or the target architecture and still keep a good scalability.

5.6.2. H.264 modeling on NoC, implementation and synthesis

In addition, the H.264 coder is modeled using MARTE profile and a hardware description for all components is proposed. Particularly the Motion Estimation (ME), adopts many new features to increase the coding performance such as block matching algorithm, motion vector prediction, variable block size motion estimation, etc. However, VBSME is utilized in the MPEG4-AVC/H.264 standard which leads to a higher computational complexity and a higher data dependence that makes the hardware implementation very complex. The aim of our work is to propose a VLSI architecture for full-search VBSME (FSVBSME). This contribution allows the reusing smaller sub-blocks for the computation results, sharing sub-blocks comparator and offering low power consumption.

5.7. Axellience

Based on the good results of the localized transformations coupled with MDFactory in Gaspard in term of reusability, modifiability and understandability, Alexis Muller (Expert Engineer) decided to study the opportunity to create a start-up company from these works. Due to his past experiment in the domain of information system and the maturity of the model usage by the enterprises, the idea was to target the automatic generation of information system from UML model and no more to address embedded systems. Joined by Thomas Legrand (Software Engineer), they developed new localized transformations and new chains. The results and the first feed-backs from enterprises in the domain of information systems are very encouraging. Furthermore, the Axellience project has won the national competitive examination of helping to the creation of innovating enterprises (Oseo). A technological transfer concerning MDFactory is foreseen between the DaRT team and the Axellience project in order to create the start-up company in the early beginning of the 2012 year. Straight collaborations between the company and the DaRT team should continue via the works around localized transformations.
5. New Results

5.1. Polychrony as open-source toolset

Participants: Loïc Besnard, Thierry Gautier, Paul Le Guernic.

A major event for us is that the open-source distribution of the Polychrony toolset has been effective since Summer 2011. The Polychrony toolset is described in Section 4.1. Following the considered part of the software, the distribution is made with the GPL V2 or EPL license. One of the objectives of this opening is to make possible a distribution of the software “by apartment” corresponding to a given functionality or to a group of functionalities, for users or developers that would be interested by only a given part of the whole software. To make this possible, a deep restructuration of the whole software has been undertaken. This takes several forms:

- One is related to the polychronous semantics and the transformations that are applied by a compilation process. A typical example is that of the representation type of the Data Control Graph (DCG), for which different levels of representation are distinguished. Some of them are based on the level of representation of the clock hierarchy.
  - The $DCGBasic$ level is the general type. The DCG represents a program with all dependencies set. Clocks are represented as signals of event type.
  - The $DCGPoly$ level is the subtype of $DCGBasic$ such that the clock hierarchy in the DCG is the result of the clock calculus. Specific clocks such as $tick$ are created, but the clock hierarchy, in the general case, has several roots.
  - The $DCGEndo$ level is the subtype of $DCGPoly$ such that the clock hierarchy in the DCG is a tree (it is provided with a single root which is $tick$). The program is endochronous.
  - The $DCGBool$ level is the subtype of $DCGEndo$ such that all clock expressions are boolean extractions. Clocks are represented as Boolean signals (no event type is used). Boolean signals representing clocks have themselves clocks represented as Boolean signals (the clock hierarchy still exists).
  - The $DCGSeq$ level is the subtype of $DCGBool$ such that all nodes of the graph are statically sorted.
  - The $DCGFlat$ level is the subtype of $DCGBool$ such that the clock hierarchy in the DCG is flat: every boolean clock signal is a direct child of the $tick$. Moreover, each state variable (corresponding to delayed signals) is defined at $tick$.

- Another aspect of the reorganization is the automatic reconstruction of the toolset from basic components. For that purpose, a new tool, called $pKmake$, has been developed, that allows the architect of the software to describe its structure and construction independently of external tools (such as $emacs$ that was used previously). It is especially useful for portability reasons, considering the different systems on which the toolset is provided.

- A third aspect that has required special attention is the automatic generation of the documentation of the source, which is realized using $cmake$, with an automatic management of cross-references.

In the context of the ITEA2 OPEES project, the Polychrony toolset is being provided as base component of the open-source toolchain of the Polarsys platform and Industry Working Group of the Eclipse consortium. A qualification plan will be defined in this context.
5.2. New features of Polychrony

**Participants:** Loïc Besnard, François Fabre, Thierry Gautier, Paul Le Guernic.

Some new features have been implemented in the Signal toolbox of the Polychrony toolset:

- It is now possible to declare *virtual* objects (types, constants and process models), which are distinguished from external objects, though objects declared as external may also be redefined in the context of declaration. The actual value of an object declared as virtual is provided in the syntactic context of declaration or in a module. A module provides a context of definition for some of the objects described as virtual in the model or the module containing the module importation command. These virtual objects are *overridden* in this way if they are imported (as corresponding objects with the same name) from an imported module, or transitively, from a module imported in an imported module.

- Process models as (static) parameters have been implemented: the formal parameters of the interface of a process model can contain process model parameters, that appear as a formal name of process model typed with a process model type. The call of a process model sets up an expansion context in which an effective process model, designated by its name, is associated with each formal model.

- The connection to the SynDEx tool ([http://www.syndex.org/](http://www.syndex.org/)) has been completed as follows. So far, only the functional part of a given application described in Signal was translated as a corresponding “algorithm” in SynDEx. The multicomponent architecture (typically, processors interconnected through communication medias) and the mapping of the algorithm onto the architecture had to be provided directly within SynDEx. As the polychronous model may be used as intermediate common formalism for applications described in languages where these aspects may be specified (this may be the case in AADL, for instance), they have to be taken into account in the translation. Thus, required elements of the architecture and distribution constraints are described using specific “pragmas” in Signal. These features are then translated into the SynDEx formalism. Using all these information, SynDEx can explore the possible implementations of the algorithm onto the multicomponent.

Moreover, we have redefined the meta-model of Signal in Ecplise, now called SSME (for Signal Syntax meta-model under Eclipse). The SME meta-model, that was used previously, suffered from several drawbacks. It was not fully complete in some parts of the language and, due to design choice, required a strict separation between clock and data flow relations. Thus specific program transformations had to be applied, which did not facilitate traceability. SSME is a full syntax oriented meta-model of Signal, very close to the abstract syntax that is used in the Signal toolbox. Compared to SME, this facilitates model transformations and traceability requirements, which are the primary objectives for its use. The transformation of AADL models into Signal, for instance, now uses the SSME meta-model.

5.3. Extensions of the language and the model

**Participants:** Loïc Besnard, Thierry Gautier, Paul Le Guernic, Jean-Pierre Talpin.

The different works on using the polychronous model as semantic median model (which has also a syntactic instance) for different effective models (AADL, Simulink via GeneAuto, UML via CCSL...) lead us to study various possible extensions of the semantic model as well as the syntactic one. Some of them have already been defined while for others, the study is still ongoing. In particular, we plan to add to the Signal language a new syntax for automata, partly inspired from AADL mode automata and hierarchic automata existing in other formalisms. An automaton is considered as an instance of a new process model and the “and” composition is the Signal composition.
A fundamental issue that we wish to address in a new way is that of globally asynchronous, locally synchronous (GALS), or globally asynchronous, locally polychronous systems. The idea we have is to extend Signal with a syntactic structure that encapsulates a polychronous (or synchronous) process \( P \) in a system, \( S \), that creates a continuous temporal domain providing a real-time clock presented in different time units (..., fs, ..., ms, ..., sec, mn, ...). Such a real-time clock can be used as a usual “synchronous” signal in the process \( P \) encapsulated in \( S \). Systems \( S_1, ..., S_n \) may be composed (with the standard composition of Signal) in a same system \( S \), but the \( ms \) of a given system \( S_i \) is a priori not synchronous with the \( ms \) of another system \( S_j \). Then it is possible to specify standard Signal constraints in the system \( S \) on these different signals, to express for instance some variation limits of different clocks.

We have also started a new work on causality aspects in order to express and operate more elaborate dependencies than instantaneous dependencies currently computed on the graph of a program. This theoretical work allows one to express dependencies that cross several instants, in a formal framework of word automata and graph algebra.

5.4. Source to source traceability in Polychrony

Participants: Loïc Besnard, François Fabre, Thierry Gautier.

To fulfill a mandatory requirement for adoption and qualification of Polychrony environment on the open-source industrial platform of the Polarsys IWG, we have integrated source to source traceability features into the Polychrony toolset. The implementation of traceability is based on the definition of structures of data and algorithms allowing to follow the transformation of objects since the Eclipse modeler of the SME Platform until the generated code. These elements have a direct application with our industrial partners, as, for example, Geensoft with whom, within the framework of the ANR project Spacify, we implemented a simulator of embedded software for satellite applications. We have also integrated such a simulator mode in the Polychrony toolset. Moreover, the error messages from the Signal compiler (Signal Toolbox) are now directly visible on the SME Graphical User Interface and on the Synoptic model (Synoptic is a satellite domain-specific modeling language).

5.5. A simulation infrastructure for CCSL, the timing model of UML MARTE

Participants: Huafeng Yu, Loïc Besnard, Thierry Gautier, Jean-Pierre Talpin, Paul Le Guernic.

Clock Constraint Specification Language (CCSL) [32] is defined in an annex of the UML MARTE profile [48]. We are interested in the analysis, synthesis and code generation of multi-clocked/polychronous systems specified in CCSL. Timed systems subject to clock expressions or relations can be modeled, specified, analyzed, and simulated within the software environments, such as SCADE [41], TimeSquare [44] and Polychrony. However, code generation from a multi-clocked system is far from obvious. For instance, SCADE always uses a reference or master clock (the fastest); all clocks and all conditions are defined as a functional sampling of this master clock, from the highest specification down to the lowest generated code. In TimeSquare, clock constraints are solved using a heuristic algorithm, which is generally non-deterministic. On the contrary, in Polychrony, a formally defined refinement process yields to the generation of (sequential or concurrent) code by the addition of control variables to get a deterministic behavior satisfying the constraints and allowing the desired amount of concurrency.

The motivation of our work, to address the simulation and code generation of polychronous systems, is to take advantage of the formal framework of Polychrony in the context of a high-level specification formalism, MARTE CCSL [22]. Yet, our work considers a novel approach with regards to previous approaches: to generate executable specifications by considering discrete controller synthesis (DCS) [50], [45], [46]. Clock constraint resolution is addressed by DCS, which does not necessarily require a master clock to address polychronous clocks. In our approach, polychronous (CCSL) specifications are first partitioned: clock relations are considered as control objectives, other constraints are considered as the system to be controlled. The all the constraints are translated into, via SIGNAL, polynomial dynamical systems (PDSs). A PDS represents the transition system of a specification as well as the constraints (invariants) it must satisfy. The Sigali tool
is then used to generate the controller. Finally, the generated controller, together with the original system, is composed to complete the code generation for simulation. In our approach, the temporal semantics of CCSL is mapped onto a polychronous model of computation, on which effective synthesis is carried out to meet constraint requirements. This approach provides both a useful mapping in theory and a flow, which is practical in the generation of reactive controllers.

5.6. The CESAR demonstrator and reference technology platform

**Participants:** Huafeng Yu, Yue Ma, Loïc Besnard, Thierry Gautier, Jean-Pierre Talpin, Paul Le Guernic.

The design of embedded systems from multiple views and heterogeneous models is ubiquitous in avionics as, in particular, different high-level modeling standards are adopted for specifying the structure, hardware and software components of a system. The system-level simulation of such composite models is necessary but difficult task, allowing to validate global design choices as early as possible in the system design flow. Inspired by the Ptolemy [40], MoBIES [31], SML-Sys [47], etc., we propose an approach to the issue of composing, integrating and simulating heterogeneous models in a system co-design flow [21]. First, the functional behavior of an application is modeled with synchronous data-flow and Statechart diagrams using Simulink/Gene-Auto [54], [55]. These high-level, synchronous and asynchronous, models are then translated into a common model, based on a polychronous model of computation, allowing for a Globally Asynchronous Locally Synchronous (GALS) interpretation of the composed models. This translation is implemented as an automatic model transformation within Polychrony. Simulation, including profiling, value change dump demonstration [24], Syndex adequation [43], etc., is carried out based on the common model within Polychrony.

Polychrony has been integrated to the Reference Technology Platform (RTP) V2 and V3 of CESAR to serve as a framework for co-modeling and architecture exploration. ModelBus [49] is used for the integration of Polychrony into the RTP. ModelBus [25], an integration platform based on Service-Oriented Architecture (SOA), connects different services offered by tools connected to ModelBus. In the demonstration, we participated in the pilot application of Sub-Project 3 (SP3), whose aim is to use the RTP to define a complete software design flow for the doors management system (DMS) of an Airbus A350 in the framework of ModelBus. In the pilot application of the DMS, functional components are modeled in the synchronous model of computation of Simulink, whereas the architecture is modeled in the asynchronous model of computation of AADL [14], [18]. These high-level models are transformed into Signal programs via SME models. Additional models, which are used in the simulation of a closed system, are coded manually in Signal and synchronously composed with the Signal programs transformed from Simulink and AADL models. Finally, C or Java code is generated from Signal programs. Simulation can then be carried out for the purpose of performance evaluation and VCD (Value Change Dump) based demonstration in RTP V2. In RTP V3, Syndex adequation is also integrated to demonstrate real-time scheduling and distribution. Our whole model transformation and simulation chain has been implemented with Galileo Eclipse and attached to ModelBus as a provider of registered remote service. This demonstration also shows the integration of Polychrony with other tools, such as OSATE (AADL), Simulink, Gene-Auto, TimeSquare, ATL, Kermeta, etc.

5.7. Modeling AADL in a polychronous model of computation

**Participants:** Loïc Besnard, Thierry Gautier, Paul Le Guernic, Jean-Pierre Talpin, Huafeng Yu.

Architecture Analysis and Design Language (AADL) is an SAE standard aimed at high level design and evaluation of architecture of embedded systems. We are interested in the analysis, simulation and verification of timed systems specified in AADL. Polychrony is well suited for the GALS architecture, and it enables deterministic specifications and formal analysis for the design of safety-critical systems. In order to benefit from the advantages provided by Polychrony, a proposition of a methodology for system-level modeling and validation of embedded systems specified in AADL via the polychronous model of computation is proposed.
By studying the different timing semantics of AADL and Polychrony, we have proposed an approach that automatically translates AADL models to a polychronous model of computation (SSME model). In the Polychrony framework, the Signal program can be generated, and an executable model can be obtained. The systems can be analyzed by tools and technologies associated with Polychrony allowing early simulation, testing and verification.

We implemented a plug-in for Eclipse framework to perform model transformation from AADL to SSME (new meta model of Signal). This transformation is implemented in Java. The following new features have been developed this year:

- Temporal interpretation of AADL model. Due to the different timing semantics between AADL and Signal, we keep the ideal view of instantaneous computations of polychronous model, moving computing latencies and communication delays to specific memory process, that introduce delays and well suited synchronizations. Each component modeled in Polychrony is composed of a behavior process (which models the functional behaviors) and a property process (which models the temporal properties).
- Architecture restructures. The architecture of the transformation is optimized. Functional architecture and meta architecture are described to give a global view of the transformation. The translation is recursive. Each AADL component is separated into a java class. The hierarchy of classes are reserved.
- Library developments. We define a Signal library containing the Signal process models representing some basic AADL concepts.
- Documentation. A new technical documentation of the transformation from AADL to SSME has been developed to accompany its implementation. This document aims to provide a global view of our implementation, from a high-level structural view to low-level implementation technical details of components.
- Programming language updates. This version of model transformation uses Java as the programming language. It avoids the disadvantages of dependent on other model transformation languages, and it provides more conveniences and flexibility. The new version is integrated as a plug-in in the Eclipse platform.
- Papers published. Three papers [14], [18], [21] are published this year.

5.8. Composing Simulink and AADL

Participants: An Phung-Khac, Jean-Pierre Talpin, Benoit Combemale, Jean-Marc Jezequel.

The goal of this work is to improve an import function of the Polychrony environment proposed by the team. Particularly, Polychrony comprises a co-modeling tool supporting the import a high-level Simulink (functional) and AADL (architectural) specifications [21]. This import function is currently implemented by two different transformations, namely Simulink-to-Signal, and AADL-to-Signal. To integrate the Signal programs resulting from these transformations, some Signal interfaces are manually implemented. The composition of Simulink and AADL models thus depends on system designers who implement the interfaces, making difficult its maintenance and validation. To deal with this issue, the model composition approach proposed by the Triskell team, namely ModMap [37], could be used to build a new Simulink and AADL model composition framework.

In ModMap, model composition is considered as a pair of a mapping and an interpretation. A mapping aligns concepts of two meta-models, while the interpretation describes the composition goal. As a model mapping framework, ModMap provides an extensible modeling language supporting the definition of generic mappings and the definition of interpretations. Together with this language, the ModMap kernel is also implemented as an extensible set of mapping processing functions. Model composition frameworks are then built by extending the language and the kernel according to specific composition purposes.
As mentioned above, we intend to apply the ModMap approach to the development of the Simulink and AADL model composition framework. To this end, we need to extend the ModMap mapping language to obtain an other one that allows system designers to align elements between Simulink and AADL models regarding the purpose of co-simulation in Signal. Then, a transformation, namely ModMap-to-Signal, needs to be implemented by extending the ModMap kernel. This transformation uses mappings provided by system designers as inputs to generate Signal interfaces. The three transformations (i.e., Simulink-to-Signal, AADL-to-Signal, and ModMap-to-Signal) form the new model composition framework. Compared to the previous one, this framework will more automated. On the other hand, existing transformations will also be reused.

5.9. From affine-related dataflow models to Safety-critical Java

Participants: Adnan Bouakaz, Jean-Pierre Talpin, Jan Vitek.

The objective of this work is to investigate a dataflow concurrency model in order to help specifying, analyzing, and synthesizing functionally deterministic and schedulable SCJ applications. Indeed, the SCJ shared-memory concurrency model makes proving functional determinism and schedulability of applications quite hard if not impossible.

The new model is called the firing related dataflow (FRDF) model in which actors are connected to each other by means of bounded channels. The operational semantics of this model is based on the notion of firing relations. Each actor is associated with a firing clock (an infinite set of activation ticks). The proportionality of the rates of two clocks is expressed by a firing relation. A special and enough expressive case of firing relations is the class of affine relations. Some results about the canonical form of affine relations are already developed by the ESPRESSO team.

Our first study was about synthesizing affine relations between firing clocks in such a way that overflow and underflow exceptions cannot occur during execution. This synthesis is conducted by minimizing the overall of buffer sizes. It is proven that the operational semantics of the dataflow graph based on the computed affine relations is equivalent to the Kahn semantics. This implies that functional determinism is guaranteed.

The previous analysis step (called affine relations synthesis) aims to produce an abstract schedule of the dataflow graph. The computed schedule is abstract in the sense that it is independent from the implementation code of actors and from the target machine. Executing the graph on a mono-processor system using EDF scheduling algorithm is investigated in our study. We synthesize the timing characteristics of each actor (i.e., its period and phase) in such a way the set of tasks is schedulable. In this timing synthesis, we use the worst-case execution times computed from the Java implementation code of actors.

Our objective is to automatically generate a SCJ application from a dataflow specification. Currently, we work on increasing the expressivity of the underlying dataflow model together with providing the necessary analysis tool for generating deterministic and schedulable SCJ code.

5.10. Translation validation of Polychronous Equations with an iLTS Model-checker

Participants: Van-Chan Ngo, Jean-Pierre Talpin, Loïc Besnard.

Synchronous languages such as SIGNAL have been introduced and used successfully for the design and implementation of embedded and critical real-time systems. They rely on the fact that programs are modeled as data-flow equations or finite state machines that allow formal reasoning on designs. In consequence of that, a full toolset of synchronous languages provides formal transformation, automatic code generation, formal verification...
In general, the synchronous language’s compiler takes several translations from the source program before generating the target code (e.g. C/C++ or Java code), thus we present an approach to verify these translations of synchronous language compiler. Our approach adopts the translation validation notion \[49\]. The idea of translation validation is the following: rather than proving in advance that the compiler always produces correct translations, each individual translation (e.g. every run of the compiler) is followed by a validation phase which verifies that the final output of this run correctly implements the input source program. This method avoids the drawback of freezing the potential improvements and/or developments of the compiler of the traditional compiler verification. For every small change in the compiler, the verification must be redoing the proof, that is an extremely complex task.

The validation phase is made automated which consists of:

(i) Represent both the input source and output target SIGNAL programs as Polynomial Dynamical Systems - PDSs.

(ii) Propose a refinement relation for the PDS models of the source and target programs.

(iii) Use a syntactic simulation-based proof method which automatically verifies the refinement. This automated proof is done by extending the functionality of the model checker SIGALI in the Polychrony toolset.

5.11. PDSs for translation validation: from SIGNAL to C

Participants: Van-Chan Ngo, Jean-Pierre Talpin, Loïc Besnard.

Synchronous programming languages provide a formal and abstract model of concurrency to facilitate the implementation of concurrent embedded software by automating the most complex tasks of verification, validation and code generation. They also guarantee the reliability of the design/implementation of concurrent embedded software by providing either the proof of compiler’s correction or the validation of each run of the compiler. Adopting the translation validation approach \[49\], we provide an automatic process to formally verify the code C generation task of the SIGNAL’s compiler.

The verification framework will take the SIGNAL program and the generated C code program as the input and proves whether the generated C code correctly implements the SIGNAL program. It also allows to automatically generate the refinement and counterexamples of the generated C code. Polynomial dynamical system - PDS is used as a common semantic framework to model the behavior of both the SIGNAL program and its generated C code. First, the generated C code is translated into the target SIGNAL program \[34\] thanks to the intermediate SSA forms. An appropriate relation called refinement for PDSs is proposed to represent the correct implementation relation between the SIGNAL program and its generated C code. The generated code C correctly implements the SIGNAL program if and only if there is a refinement for their PDSs and we say that the generated C code’s PDS refines the SIGNAL program’s PDS. A proof method which allows to generate the refinement or counterexamples, and then proposes a refining process for the generated C code.

5.12. Synchronous symbolic translation systems for translation validation

Participants: Van-Chan Ngo, Jean-Pierre Talpin.

We propose a framework for verification of the correct implementation of the SIGNAL compiler’s generation code task. In order to present the formal semantics of SIGNAL and generated code programs we introduce synchronous symbolic transition system (SSTS) which is the computational model of our formal verification approach. We denote \(D_V = \prod_{i \in [1,n]} D_{v_i}\) as the domain of a set of variables \(V = (v_1, ..., v_n)\). A set of states \(P \subseteq D_V\) is defined as a predicate over the set of variables \(V\) such that the predicate is held in \(P\). An assignment \(A\) is a function \(A : D_V \rightarrow D_V\) that the values of the variable set \(V\). A SSTS is a tuple \(L = (V, \Theta, \Gamma, \mathcal{E})\) where:

- \(V = (v_1, ..., v_n)\) is a set of variables,
- \(\Theta \subseteq D_V\) is a predicate on \(V\) defining the initial condition on the variable set,
- \(\Gamma\) is a finite set of symbolic transitions \(\gamma = (P_\gamma, A_\gamma)\) where:
– \( P_\gamma \subseteq D_V \) is a predicate on \( V \), which guards \( \gamma \)
– \( A_\gamma : D_V \to D_V \) is the assignment function of \( \gamma \)

– \( \mathcal{E} \subseteq V \) is a set of externally observable variables.

The generated code correctly implements the SIGNAL program if and only if there is a refinement for their SSTSs and we say that the generated code’s SSTS refines the SIGNAL program’s SSTS. This framework also works with SIGNAL programs which is considered as infinite state systems. To obtain the verification results, we apply abstraction interpretation techniques \([39]\) which provide over-approximations of the refinement relation between the input SIGNAL program’s model and the output generated code’s model.

### 5.13. An integrated environment for Esterel/Quartz and Polychrony/Signal

**Participants:** Jens Brandt, Ke Sun, Jean-Pierre Talpin.

The design of modern embedded software architectures relies on models and programs built and reused from engineering teams with specific skills and know-how. Each of these skills and backgrounds correspond to specific tools and processes that help implement the viewpoint under consideration with mathematically grounded foundations.

It is not uncommon, for instance, that the design of the only functional views of a system may require the use of tools as heterogeneous and exotic as Catia, Scade, Matlab or Rhapsody. The same holds for design objectives that may range from that of mapping the functional design on specific hardware architectures to that of virtual prototyping for simulation or performance or energy usage evaluation.

Co-modeling itself encompasses the variety of engineering activities that cross the border between the functional and physical views of system design. It is typically the system architects, who will put together functional components and explore different metrics for an effective and efficient mapping on target systems.

We wish to further and scale the framework and experiments developed within the CESAR and OPEES projects in that respect, by thinking a new, domain-specific language, built from synchronous modules designed with Quartz, an imperative synchronous programming language, and connected by data-flow networks described in Signal, the polychronous data-flow language at the core of Polychrony. The combination of viewpoints or paradigms offered by these two design environments provides powerful abstractions and easy to use concepts in order to address two design challenges of utmost importance:

– To provide a natural and dependable specification of elementary synchronous functionalities, most of them algorithmic and control-intensive, in the imperative framework offered by Quartz.

– To synthesize the scheduling of computations and communications among these functionalities starting from the multi-clocked synchronous abstractions offered by the Signal data-flow language.

The remaining long-term goal will then be homogenize this programming framework by further extending it with the capability to control polychromous networks, seen as modes of execution, with a Quartz module, which would control mode changes.

### 5.14. Quality assessment and qualification of Polychrony on the open-source Polarsys IWG platform

**Participants:** Loïc Besnard, Thierry Gautier, Paul Le Guernic, Jean-Pierre Talpin.
Since the open-source release of Polychrony and in the context of the ITEA2 OPEES project, we are collaborating with CS to the integration of Polychrony on the Polarsys platform. This integration proceeds according to guidelines and requirements under definition within the OPEES project and aims first, at putting them to the test. The qualification process of Polychrony in Polarsys consists of checking the maturity level of its implementation and documentation (a standard software engineering assessment) but is also concerned with its capabilities to be composed, inter-operated and mapped with other components on the platform to form a application-specific design toolchain, by using model-driven engineering technologies for model transformation and orchestration. The last phase of this assessment is with regards to its qualifiability, as a simulation tool, as a verification tool, as a code generation tools, which needs to adhere standards such as these defined the in DO178 documents. In parallel, the quality assessment of Polychrony is complemented with a case study, of the APOTA network protocol, whose aim is to document, as a tutorial, the use and added-value of the toolset for its future users on the Polarsys platform.
6. New Results

6.1. Simulation

6.1.1. Simulation of vector architecture
Participants: Vania Joloboff, Yang Yu.

Many architectures including PowerPC and ARM now have vectorized instructions, that is, instructions that can execute on several data items in parallel (e.g., 8 simultaneous additions) on specific vector data.

We have implemented the ALTIVEC extension of the PowerPC to support the vector instructions.

6.1.2. Native translation using LLVM
Participants: Vania Joloboff, Xinlei Zhou, Zuyu Zhang.

We have started to implement a new technique of dynamic translation. This new method consists in decompiling the binary object code into an abstract representation and recompiling it to native host code.

The decompilation of the program amounts to reconstructing the simulated program Control Flow Graph using an intermediate representation. We have chosen LLVM (Low Level Virtual Machine), defined by University of Illinois, and now widely adopted in many projects, as our representation format. Using LLVM allows us to directly use the LLVM Intel code generator.

The SimSoC binary decoder has been modified to identify basic blocks (blocks of sequential instructions ending with a branch instruction). After instructions have been grouped into basic blocks, they are translated into an LLVM representation and finally the LLVM compiler is called to generate native code.

A first version of this technique has been implemented for both the ARM and Power Architecture. We have reach a considerable speed improvement in the generated code, with the execution speed multiplied by factor of 2 to 8. However the translation time from binary to LLVM and from LLVM to native code is significant (translation speed is roughly 1000 instructions per second). Consequently the overall speed is improved by only a factor of 20 to 50 percent when the simulation are relatively short test programs [20].

In order to reach still higher simulation speed we need to use a more sophisticated analysis of the control flow graph. The idea is to do an edge profiling analysis of the basic blocks in order to identify larger blocks. This work is under development.

Another idea is to use multi-processor hosts machine to parallelize translation from LLVM to native code. This is also under investigation.

6.1.3. Trace Analysis
Participants: Guillaume Merle, Vania Joloboff.

Simulation sessions produce huge trace files, sometimes now in hundreds of gigabytes, that are hard to analyze with a quick response time. This comes down to two sub-problems:

- The trace file size. Trace files are huge because they include lots of information. But when looking for a specific problem, one does not need all of this information. To search one given defect, one may ignore a large amount of the data in the trace file. One would like the trace file to contain only relevant information to the concerned problem.

- The expressive power of the language to analyze the trace, and its usability. If the language is limited to expression search, it is easy to use but hard to construct sophisticated formulas. If the language used is Linear Temporal Logic (LTL), there is a lot of expressive power but many engineers are unable to write a LTL formula and to maintain it over time.
We would like to build a trace analysis tool that includes a language which allows expression of time-related formulas but is simple to formulate expressions. When this language is compiled, ideally the compiler is smart enough to identify independent formulae, the search of which can be parallelized, and it is also smart enough to generate "filter scripts".

When compiling one trace language input file, it would generate, from one input file, N filter scripts and N analyzers. Then during the simulation, the huge raw trace file is actually split into N smaller trace files, each relevant to one problem only, filtering out all unnecessary data. Hence trace files sizes would be considerably reduced.

We have started to design a trace language and a compiler, and extended the SimSoC simulator to support generation of trace files with a filter.

A first version of the trace language compiler has been coded in OCAML.

In the current version under development, the filters are not generated but coded manually, and filters are not parallelized.

6.1.4. Generation of simulators from vendor specification

Participants: Frédéric Blanqui, Vania Joloboff, Jean-François Monin, Xiaomu Shi, Frédéric Tuong.

Starting last year, we undertook the task of generating automatically an instruction set simulator (ISS) from the vendor specification in a PDF file. In order to generate the C code of the simulator, it is assumed such vendor specification contains at least some formal definitions of the instruction set that can be analyzed. It is the case to a wide extent for the ARM, the PowerPC and the SH architectures.

The process of generating the simulator consists of 4 major steps, first eliminating from the PDF file irrelevant information, next construct from the relevant data an abstract syntax representation of the instruction set, then to generate the C code of the simulator, using some additional data provided manually to complete the vendor specification.

This work was completed last year for the ARM architecture with the documentation form ARM corporation [35]. This year, we did similar work for the SH architecture from specification from RENESAS corporation. We have indeed generated a simulator for the SH4 architecture [31], which has not been fully tested yet.

However, this works has proved that the abstract syntax we have defined is powerful enough to describe two different architectures with significant differences in the way they are described by the vendor.

6.1.5. First steps towards the certification of an ARM simulator

Participants: Frédéric Blanqui, Jean-François Monin, Xiaomu Shi, Frédéric Tuong.

The simulation of Systems-on-Chip (SoC) is nowadays a hot topic because, beyond providing many debugging facilities, it allows the development of dedicated software before the hardware is available. Low-consumption CPUs such as ARM play a central role in SoC. However, the effectiveness of simulation depends on the faithfulness of the simulator. To this effect, in [24], we propose here to prove significant parts of such a simulator, SimSoC. Basically, on one hand, we develop a Coq formal model of the ARM architecture while on the other hand, we consider a version of the simulator including components written in Compcert-C [61]. Then we prove that the simulation of ARM operations, according to Compcert-C formal semantics, conforms to the expected formal model of ARM. Size issues are partly dealt with using automatic generation of significant parts of the Coq model and of SimSoC from the official textual definition of ARM. However, this is still a long-term project. We report here the current stage of our efforts and discuss in particular the use of Compcert-C in this framework.

6.2. Type and rewriting theory

6.2.1. A type theory for Coq

Participants: Jean-Pierre Jouannaud, Qian Wang.
In this joint work with Bruno Barras and Pierre-Yves Strub [17], we describe an abstract model of CoqMT [73] called CoqMTU, which puts together the Calculus of Inductive Constructions, decidable first-order theories, and an infinite hierarchy of universes which are all predicative but the first impredicative universe of propositions. We have shown its consistency, strong normalization and decidability of type checking in presence of weak elimination (and absence of strong elimination). An important feature of this work is that the first-order theory is abstract, characterized by the three natural axioms that (i) it is non-degenerated (its models have at least two elements), (ii) constructors are free, and (iii) defined symbols are completely defined. On the theoretical side, this allows us to give an abstract elimination principle for such non-canonical theories. On the practical side, this justifies the implementation of CoqMT in which decidable theories can be dynamically downloaded. It should be noticed that these proofs are done in Coq, except for the strong normalization part. Qian Wang is now continuing this work at Ecole Polytechnique with Bruno Barras and Pierre-Yves Strub, the target being strong normalization.

6.2.2. **Confluence by decreasing diagrams**

**Participants:** Jean-Pierre Jouannaud, Huiying Luo, Jiaxiang Liu.

Invented by Vincent Van Oostrom, decreasing diagrams capture both kinds of diagrams arising from Newman’s Lemma and Hindley’s Lemma: they indeed allow to reduce all known confluence methods to critical pairs computations, and a search of decreasing diagrams for them all, where decreasingness is measured by a well-founded order on proof steps.

In [55], we give a new simple proof of Van Oostrom’s main theorem, and extend the method of decreasing diagrams to rewrite relations on a term algebra. We prove that the union of a terminating left-linear systems, and a non-terminating linear system is confluent provided the various critical pairs existing in in their combination have decreasing diagrams (with respect to some order built from the respective orders of both systems).

During this year, we have further simplified and generalized these results in order to get rid of the left-linearity assumption for the first system, and of the right-linearity assumption for the second. This yields a true generalization of the well-known Knuth-Bendix-Huet confluence result for terminating systems, and at the same time of various critical-pair based results found in the literature for non-terminating systems.

6.2.3. **Confluence of normal rewriting**

**Participants:** Jean-Pierre Jouannaud, Jianqi Li.

Confluence results for first-order and higher-order rewriting differ in many ways: by the rewriting relation used, and by the strong normalization assumption made. We believe that these differences hide the strong similarities of these (and other) kinds of rewriting.

In this work, we introduce a new notion of rewriting, **normal rewriting**, which aims at capturing all known results reducing confluence to critical (and extension) pair computations in presence of some termination assumption.

We achieve this goal in the following way. First, we consider theories made of a set $R$ of rules, a set $S$ of simplifiers, and a set $E$ of equations. Rewriting operates on terms in $S$ modulo $E$ normal forms, and uses $S$ union $E$-pattern matching for firing the rules in $R$, before to normalize the result with respect to $S$ modulo $E$. Termination is assumed for the union of $S$ modulo $E$ and $R$ modulo $S$ union $E$. Second, we introduce relations on an abstract set of terms, and an abstract, well-founded set of positions, and reduce the Church-Rosser property of abstract normal rewriting to abstract notions of critical pairs and extensions. We can then apply this result to first-order rewriting, as well as to various forms of higher-order rewriting. These results capture plain rewriting ($S$ union $E$ is empty), Stickel’s rewriting modulo ($S$ is empty), Nipkow’s higher-order rewriting ($S$ is made of beta-reduction and eta-expansion, and $E$ is alpha-conversion), and allow to describe new forms of first and higher-order rewriting relations.

6.2.4. **Argument filterings and usable rules in higher-order rewrite systems**

**Participant:** Frédéric Blanqui.
Joint work with Keiichirou Kusakari and Sho Suzuki from Nagoya University, Japan.

The static dependency pair method is a method for proving the termination of higher-order rewrite systems à la Nipkow [62]. It combines the dependency pair method introduced for first-order rewrite systems with the notion of strong computability introduced for typed lambda-calculi [52]. Argument filterings and usable rules are two important methods of the dependency pair framework used by current state-of-the-art first-order automated termination provers [51], [53]. In [12], we extend the class of higher-order systems on which the static dependency pair method can be applied. Then, we extend argument filterings and usable rules to higher-order rewriting, hence providing the basis for a powerful automated termination prover for higher-order rewrite systems.

6.3. Decision procedures

6.3.1. A certificate framework for DPLL(T)

Participants: Min Zhou, Fei He, Bow-Yaw Wang, Wenrui Meng.

Satisfiability Modulo Theories (SMT) techniques are widely used nowadays. SMT solvers are used to decide the satisfiability of first-order formulas. When an SMT solver is invoked, it is important to ensure correctness of the result. For this purpose, we proposed a certificate framework based on DPLL(T), including generation of certificates and verification of certificates. Some properties are discussed and proved theoretically. The certificate is easy to generate because it only needs minor modification to the existing SMT solvers. Experiment results show that the overhead for certificates generation is only 10%. Moreover, verifying the certificate requires few memory and time, which outperforms other approaches.

6.3.2. Automated verification of termination certificates

Participants: Frédéric Blanqui, Kim-Quyen Ly, Sidi Ould Biha.

The research community on rewriting developed a grammar for termination certificates called CPF [29] (given by a XML Schema file). Our goal is to develop a safe, modular and efficient termination certificate verifier based on the formal library of mathematical results on termination called CoLoR that has been developed for the proof assistant Coq [11].

Because the CPF format is regularly modified and extended with new features, it is useful to have a tool that can automatically generate data structures, parsers and pretty-printers for that format. Hence, we developed a first version of such a tool in OCaml.

Once we got a representation of termination certificates in Coq, we could start defining a boolean function checking the correctness of a certificate, and formally prove its correctness. For the moment, we only considered the case of polynomial interpretations on integers. The proof is almost finished. To do so, we had to modify some of the CoLoR files to be able to use its results (transformation of modules into records that are first-class objects). The use of dependent types in CoLoR makes also definitions and proofs much more difficult.

6.3.3. Proving computational geometry algorithms in TLA+2

Participants: Hui Kong, Hehua Zhang, Ming Gu.

Geometric algorithms are widely used in many scientific fields like computer vision, computer graphics. To guarantee the correctness of these algorithms, it is important to apply formal method to them. In this work, we propose an approach to proving the correctness of geometric algorithms [22]. The main contribution is that a set of proof decomposition rules is proposed which can help improve the automation of the proof of geometric algorithms. We choose TLA+2, a structural specification and proof language, as our experiment environment. The case study on a classical convex hull algorithm shows the usability of the method.

6.4. Compositional verification

6.4.1. BDD-based assume-guarantee reasoning through implicit learning

Participants: Fei He, Bow-Yaw Wang, Lei Zhu.
We present a purely BDD-based assume-guarantee reasoning technique to improve the scalability of symbolic model checking. The new technique adopts a BDD learning algorithm to generate BDD’s as contextual assumptions. A new witness analysis algorithm is proposed to exploit the multitude of traces returned by symbolic model checkers. Using the classification tree-based BDD learning algorithm to generate contextual assumptions, we compare assume-guarantee reasoning with monolithic symbolic model checking. The new technique always infers smaller contextual assumptions than contexts in our experiments.

6.4.2. Predicate generation for learning-based loop invariant inference

**Participant:** Bow-Yaw Wang.

We address the predicate generation problem in the context of loop invariant inference. Motivated by the interpolation-based abstraction refinement technique, we apply the interpolation theorem to synthesize predicates implicitly implied by program texts. Our technique is able to improve the effectiveness and efficiency of the learning-based loop invariant inference algorithm in [21]. Experiments excerpted from Linux, SPEC2000, and Tar source codes are reported.

This is a joint work with Yungbum Jung, Wonchan Lee, and Kwangkuen Yi of Seoul National University, South Korea.

6.4.3. Thread-modular model checking with iterative refinement

**Participants:** Wenrui Meng, Fei He, Bow-Yaw Wang.

Thread-modular analysis is an incomplete compositional technique for verifying concurrent systems. The heuristic works rather well when there is limited interaction among system components. In this project, we develop a refinement algorithm that makes thread-modular model checking complete. Our algorithm refines abstract reachable states by exposing local information through auxiliary variables. The experiments show that our complete thread-modular model checking can outperform other complete compositional reasoning techniques.

6.5. Specification and verification of TLA+ and PLC systems

6.5.1. Formal semantics of PLC programming languages

**Participants:** Sidi Ould Biha, Litian Xiao, Ming Gu.

We formalized a semantics of the Instruction List (IL) language, one of the five programming languages defined in the IEC 61131-3 standard for PLC programming [23]. This semantics support a significant subset of the IL language that includes on-delay timers. This semantics was used in a joint work to with Jan Olaf Blech from Fortiss (Germany) to prove some safety properties for a real industrial example of PLC program [18].

A second widely used language for programming PLC is the graphical language Ladder Diagrams (LD). We defined a formal semantics of LD in the proof assistant Coq. Based on this semantics and the IL one, we developed a translation function from LD to IL. We also proved a semantic preservation property for this translation function. We have now a certified compilation function from the graphical language LD to IL. This work opens the way for the development of a certified compilation chain for PLC. A journal paper about this work and others aspects of PLC certification is under reviewing.

In [16], [15], we study the definition of denotational semantics on PLC program language, which is convenient to PLC programs modeling and model checking. The purpose of the work is the correctness verification on PLC programs by formal methods. Based on the extended \( \lambda \)-calculus definition, this work has defined the configuration of PLC program architecture, denotational semantics of PLC programs and functions of denotational semantics. It is the basis of model checking and theorem proving.

6.5.2. Formalization and verification of PLCs

**Participants:** Hai Wan, Litian Xiao, Ming Gu.
PLCs are widely used in embedded systems. Timers play a pivotal role in PLC real-time applications. The formalization of timers is of great importance. In [13], we present a formalization of PLC timers in the theorem proving system Coq, in which the behaviors of timers are characterized by a set of axioms at an abstract level. The authors discuss how to model timers at a proper and sound abstract level. PLC programs with timers are modeled. As a case study, a quiz machine problem with a timer is investigated. This work demonstrates the complexity of formal timer modeling.

In [25], we modeled kernel data type and basic statements and the denotational semantics of PLC program in Coq. It has given the correctness proof of PLC program based on theorem proving, i.e., based on semantics function the relationship of configuration between the before codes execution and the after is proved. The main purpose is to prove whether a PLC program satisfies certain nature within a scan period.

6.5.3. Synthesis of PLC programs
Participants: Rui Wang, Ming Gu.

PLCs are complex cyber-physical systems which are widely used in industry. In [14], we present a robust approach to design and implement PLC-based embedded systems. Timed automata are used to model the controller and its environment. We validate the design model with resort to model checking techniques. We propose an algorithm to generate PLC code from timed automata and implement this algorithm with a prototype tool. This method can condense the developing process and guarantee the correctness of PLC programs. A case study demonstrates the effectiveness of the method.

6.5.4. Domain-driven probabilistic analysis of PLCs
Participants: Hehua Zhang, Yu Jiang, Ming Gu.

Programmable Logic Controllers are widely used in industry. Reliable PLCs are vital to many critical applications. We present a novel symbolic approach for analysis of PLC systems [27]. The main components of the approach consists of: (1) calculating the uncertainty characterization of the PLC systems, (2) abstracting the PLC system as a Hidden Markov Model, (3) solving the Hidden Markov Model using domain knowledge, (4) integrating the solved Hidden Markov Model and the uncertainty characterization to form an integrated (regular) Markov Model, and (5) harnessing probabilistic model checking to analyze properties on the resultant Markov Model. The framework provides expected performance measures of the PLC systems by automated analytical means without expensive simulations. Case studies on an industrial automated system are performed to demonstrate the effectiveness of our approach.

6.5.5. Edola: a domain modeling and verification language for PLCs
Participants: Hehua Zhang, Ming Gu.

Formal modeling and verification of PLC systems become paramount in engineering applications. The work presents a novel PLC domain-specific modeling language Edola [26]. Important characteristics of PLC embedded systems, such as reactivity, scan cycling, real-time and property patterns, are embodied in the language design. Formal verification methods, such as model checking and automatic theorem proving, are supported in Edola modeling. The TLA+ specification language constitutes an intermediate language layer between Edola and the verification tools, enhancing a large degree of reusability. A prototype IDE for Edola and its seamless integration of a model checker TLC and an automatic theorem prover Spass are implemented. A case study illustrates and validates the applicability of the language.

6.6. Distributed algorithms

6.6.1. Formal model and proofs for Netlog protocols
Participants: Meixian Chen, Jean-François Monin.

Joint work with Yuxin Deng (Jiaotong University, Shanghai) and Stéphane Grumbach (LIAMA/Netquest).
Netlog is a language designed and implemented in the Netquest project for describing protocols. Netlog has a precise semantics, provides a high level of abstraction thanks to its Datalog flavor and benefits from an efficient implementation. This makes it a very interesting target language for proofs of protocols. Netlog comes with two possible semantics: a synchronous semantics, better suited to tightly coupled parallel systems and an asynchronous semantics, better suited to distributed systems.

We designed a formal model of Netlog in Coq, where the two possible semantics are derived from common basic blocks. In a fully certified framework, a formal proof of the Netlog engine (running on each node) would be required. We don’t attack this part at the moment: we assume that the implementation respects the general properties stated in our model and focus on the issues raised by the distributed model of computation provided by Netlog.

As a proof of concept, we applied in 2010 this framework to an algorithm constructing a Breadth-First Search Spanning Tree (BFS) in a distributed system [46]. This work has been slightly improved this year and published in [19].

Moreover, we generalized the model in order to take the removal of datalog facts into account, and started to use this feature for more complicated protocols. In main one under study is Prim’s algorithm (publication under submission), and we target next GHS, which still resists to palatable proof techniques.
6. New Results

6.1. Algebraic Algorithms for Geometric Computing

6.1.1. An Algebraic Approach to Continuous Collision Detection for Ellipsoids

Participant: Bernard Mourrain.

In [28], we present algebraic conditions for characterizing three configurations of two ellipsoids in R^3 that are the most relevant to collision detection: separation, external touching and overlapping. These conditions are given in terms of explicit formulae expressed by the subresultant sequence of the characteristic polynomial of the two ellipsoids and its derivative. For any two ellipsoids, the signs of these formulae can easily be evaluated to classify their configuration. Furthermore, based on these algebraic conditions, an efficient method is developed for continuous collision detection for two moving ellipsoids under arbitrary motion.

This is a joint work with Xiaohong Jia, Yi-King Choi and Wenping Wang from the university of Hong Kong.

6.1.2. On Continued Fraction Expansion of Real Roots of Polynomial Systems, Complexity and Condition Numbers

Participants: Angelos Mantzaflaris, Bernard Mourrain.

In [29], we elaborate on a correspondence between the coefficients of a multivariate polynomial represented in the Bernstein basis and in a tensor-monomial basis, which leads to homography representations of polynomial functions, that use only integer arithmetic (in contrast to Bernstein basis) and are feasible over unbounded regions. Then, we study an algorithm to split this representation and we obtain a subdivision scheme for the domain of multivariate polynomial functions. This implies a new algorithm for real root isolation, MCF, that generalizes the Continued Fraction (CF) algorithm of univariate polynomials. A partial extension of Vincent’s Theorem for multivariate polynomials is presented, which allows us to prove the termination of the algorithm. Bounding functions, projection and preconditioning are employed to speed up the scheme. The resulting isolation boxes have optimized rational coordinates, corresponding to the first terms of the continued fraction expansion of the real roots. Finally, we present new complexity bounds for a simplified version of the algorithm in the bit complexity model, and also bounds in the real RAM model for a family of subdivision algorithms in terms of the real condition number of the system. Examples computed with our C++ implementation illustrate the practical aspects of our method.

This is a joint work with E. Tsigaridas, from the Department of Computer Science, University of Aarhus.

6.1.3. Matrix-based representations of rational hypersurfaces

Participants: Laurent Busé, Nicolas Botbol.

This ongoing work is related to matrix-based representations of rational hypersurfaces whose theoretical funda-ments has been recently developed by our team and several other authors in the context of the implicitization problem. Being given a parameterized curve or hypersurface, this method consists in building a matrix whose entries are typically linear forms in the variables of the ambient space and such that the ideal generated by the maximal minors of this matrix provide a good approximation of the original curve or hypersurface.

We aim to study and determine the geometric informations that are contained in a representation matrix. In particular, we are currently adressing the two following questions:

1) understand the extraneous components that are added by taking the initial Fitting ideal of a representation matrix, with respect to the original curve or surface. Indeed, these extraneous components appear because of the good behavior of Fitting ideals under change of bases. Therefore, one can expect that these extraneous components yields some geometric properties of a curve or surface as a member of a certain family. 2) examine the extraction of singularities from a representation matrix, similarly to the recent results on what is called "singular factors" for the case of rational curves.
6.1.4. The surface/surface intersection problem by means of matrix based representations  
Participants: Laurent Busé, Thang Luu Ba.

Evaluating the intersection of two rational parameterized algebraic surfaces is an important problem in solid modeling. In this work, we made use of some generalized matrix based representations of parameterized surfaces in order to represent the intersection curve of two such surfaces as the zero set of a matrix determinant. As a consequence, we extended to a dramatically larger class of rational parameterized surfaces the applicability of a general approach to the surface/surface intersection problem due to J. Canny and D. Manocha. In this way, we obtained compact and efficient representations of intersection curves allowing to reduce some geometric operations on such curves to matrix operations using results from linear algebra.

See the preprint version at http://hal.inria.fr/inria-00620947/en/

6.2. Symbolic-Numeric Analysis

6.2.1. A Subdivision Method for Computing Nearest Gcd with Certification  
Participants: André Galligo, Bernard Mourrain.

A new subdivision method for computing the nearest univariate gcd is described and analyzed in [24]. It is based on an exclusion test and an inclusion test. The exclusion test in a cell exploits Taylor expansion of the polynomial at the center of the cell. The inclusion test uses Smale’s alpha-theorems to certify the existence and unicity of a solution in a cell. Under the condition of simple roots for the distance minimization problem, we analyze the complexity of the algorithm in terms of a condition number, which is the inverse of the distance to the set of degenerate systems. We report on some experimentation on representative examples to illustrate the behavior of the algorithm.

This is a joint work with Guillaume Chèze and Jean-Claude Yakoubsohn (University Paul-Sabatier, Toulouse).

6.2.2. An Adapted Version of the Bentley-Ottmann Algorithm for Invariants of Plane Curves Singularities

Participant: Bernard Mourrain.

In [34], we report on an adapted version of the Bentley-Ottmann algorithm for computing all the intersection points among the edges of the projection of a three-dimensional graph. This graph is given as a set of vertices together with their space Euclidean coordinates, and a set of edges connecting them. More precisely, the three-dimensional graph represents the approximation of a closed and smooth implicitly defined space algebraic curve, that allows us a simplified treatment of the events encountered in the Bentley-Ottmann algorithm. As applications, we use the adapted algorithm to compute invariants for each singularity of a plane complex algebraic curve, i.e. the Alexander polynomial, the Milnor number, the delta-invariant, etc.

This is a joint work with Madalina Hodorog and Joseph Schicho, from RICAM, Linz, Austria.

6.2.3. Virtual Roots of a Real Polynomial and Fractional Derivatives

Participant: André Galligo.

After the works of Gonzales-Vega, Lombardi, Mahé, and Coste, Lajous, Lombardi, Roy, we consider the virtual roots of a univariate polynomial $f$ with real coefficients. Using fractional derivatives, we associate to $f$ a bivariate polynomial $P_f(x,t)$ depending on the choice of an origin $a$, then two type of plan curves we call the FDcurve and stem of $f$. We show, in the generic case, how to locate the virtual roots of $f$ on the Budan table and on each of these curves. The paper [32] is illustrated with examples and pictures computed with the computer algebra system Maple. It is a joint work with Daniel Bembe.

6.2.4. Computing monodromy via continuation methods on random Riemann surfaces

Participant: André Galligo.
In [25], we consider a Riemann surface $X$ defined by a polynomial $f(x, y)$ of degree $d$, whose coefficients are chosen randomly. Hence, we can suppose that $X$ is smooth, that the discriminant $\delta(x)$ of $f$ has $d(d-1)$ simple roots $\Delta$ and that $\delta(0) \neq 0$, i.e. the corresponding fiber has $d$ distinct points $\{y_1, \ldots, y_d\}$. When we lift a loop $0 \in \gamma \subset \mathbb{C} - \Delta$ by a continuation method, we get $d$ paths in $X$ connecting $\{y_1, \ldots, y_d\}$, hence defining a permutation of that set. This is called monodromy.

Here we present experimentations in Maple to get statistics on the distribution of transpositions corresponding to loops around each point of $\Delta$. Multiplying families of "neighbor" transpositions, we construct permutations and the subgroups of the symmetric group they generate. This allows us to establish and study experimentally two conjectures on the distribution of these transpositions and on transitivity of the generated subgroups.

Assuming that these two conjectures are true, we develop tools allowing fast probabilistic algorithms for absolute multivariate polynomial factorization, under the hypothesis that the factors behave like random polynomials whose coefficients follow uniform distributions. It is a joint work with Adrien Poteaux (University of Lille).

6.3. Algebraic representations for geometric modeling

6.3.1. Multihomogeneous Polynomial Decomposition using Moment Matrices

Participants: Alessandra Bernardi, Jérôme Brachat, Bernard Mourrain.

In [33], we address the important problem of tensor decomposition which can be seen as a generalisation of Singular Value Decomposition for matrices. We consider general multilinear and multihomogeneous tensors. We show how to reduce the problem to a truncated moment matrix problem and we give a new criterion for flat extension of Quasi-Hankel matrices. We connect this criterion to the commutation characterization of border bases. A new algorithm is described: it applies for general multihomogeneous tensors, extending the approach of J.J. Sylvester on binary forms. An example illustrates the algebraic operations involved in this approach and how the decomposition can be recovered from eigenvector computation.

This is a joint work with Pierre Comon (I3S, CNRS).

6.3.2. On the variety parametrizing completely decomposable polynomials.

Participant: Alessandra Bernardi.

The purpose of the paper [15] is to relate the variety parameterizing completely decomposable homogeneous polynomials of degree $d$ in $n+1$ variables on an algebraically closed field, called $\text{Split}_d(\mathbb{P}^n)$, with the Grassmannian of $n-1$ dimensional projective subspaces of $\mathbb{P}^{n+d-1}$. We compute the dimension of some secant varieties to $\text{Split}_d(\mathbb{P}^n)$ and find a counterexample to a conjecture that wanted its dimension related to the one of the secant variety to $\mathcal{G}(n-1, n+d-1)$. Moreover by using an invariant embedding of the Veronese variety into the Plücker space, then we are able to compute the intersection of $\mathcal{G}(n-1, n+d-1)$ with $\text{Split}_d(\mathbb{P}^n)$, some of its secant variety, the tangential variety and the second osculating space to the Veronese variety.

This is a joint work with Enrique Arrondo (Universidad Complutense de Madrid, Spain)

6.3.3. Computing symmetric rank for symmetric tensors.

Participant: Alessandra Bernardi.

In [21] we consider the problem of determining the symmetric tensor rank for symmetric tensors with an algebraic geometry approach. We give algorithms for computing the symmetric rank for $2 \times \ldots \times 2$ tensors and for tensors of small border rank. From a geometric point of view, we describe the symmetric rank strata for some secant varieties of Veronese varieties.

This is a joint work with Alessandro Gimigliano and Monica Idà (Università di Bologna, Italy).

6.3.4. Higher secant varieties of $\mathbb{P}^n \times \mathbb{P}^m$ embedded in bi-degree $(1, d)$.

Participant: Alessandra Bernardi.
Let $X^{(n,m)}_{(1,d)}$ denote the Segre-Veronese embedding of $\mathbb{P}^n \times \mathbb{P}^m$ via the sections of the sheaf $\mathcal{O}(1,d)$. In [20] we study the dimensions of higher secant varieties of $X^{(n,m)}_{(1,d)}$ and we prove that there is no defective $s^{th}$ secant variety, except possibly for $n$ values of $s$. Moreover when $\binom{m+n+1}{d}$ is multiple of $(m+n+1)$, the $s^{th}$ secant variety of $X^{(n,m)}_{(1,d)}$ has the expected dimension for every $s$.

This is a joint work with Enrico Carlini (Politecnico di Torino, Italy, Maria Virginia Catalisano (Università di Genova, Italy).

**6.3.5. On the $X$-rank with respect to linear projections of projective varieties.**

**Participant:** Alessandra Bernardi.

In [17] we improve the known bound for the $X$-rank $R_X(P)$ of an element $P \in \mathbb{P}^N$ in the case in which $X \subset \mathbb{P}^n$ is a projective variety obtained as a linear projection from a general $\nu$-dimensional subspace $V \subset \mathbb{P}^{n+\nu}$. Then, if $X \subset \mathbb{P}^n$ is a curve obtained from a projection of a rational normal curve $C \subset \mathbb{P}^{n+1}$ from a point $O \subset \mathbb{P}^{n+1}$, we are able to describe the precise value of the $X$-rank for those points $P \in \mathbb{P}^n$ such that $R_X(P) \leq R_C(O) - 1$ and to improve the general result. Moreover we give a stratification, via the $X$-rank, of the osculating spaces to projective cuspidal projective curves $X$. Finally we give a description and a new bound of the $X$-rank of subspaces both in the general case and with respect to integral non-degenerate projective curves.

This is a joint work with Edoardo Ballico (Università di Trento, Italy).

**6.3.6. Decomposition of homogeneous polynomials with low rank.**

**Participant:** Alessandra Bernardi.

Let $F$ be a homogeneous polynomial of degree $d$ in $m+1$ variables defined over an algebraically closed field of characteristic zero and suppose that $F$ belongs to the $s$-th secant varieties of the standard Veronese variety $X_{m,d} \subset \mathbb{P}^{\binom{m+d}{d}-1}$ but that its minimal decomposition as a sum of $d$-th powers of linear forms $M_1, \ldots, M_r$ is $F = M_1^r + \cdots + M_r^r$ with $r > s$. In [16] we show that if $s + r \leq 2d + 1$ then such a decomposition of $F$ can be split in two parts: one of them is made by linear forms that can be written using only two variables, the other part is uniquely determined once one has fixed the first part. We also obtain a uniqueness theorem for the minimal decomposition of $F$ if the rank is at most $d$ and a mild condition is satisfied.

This is a joint work with Edoardo Ballico (Università di Trento, Italy).

**6.3.7. On the $X$-rank with respect to linearly normal curves.**

**Participant:** Alessandra Bernardi.

In [18] we study the $X$-rank of points with respect to smooth linearly normal curves $X \subset \mathbb{P}^n$ of genus $g$ and degree $n + g$.

We prove that, for such a curve $X$, under certain circumstances, the $X$-rank of a general point of $X$-border rank equal to $s$ is less or equal than $n + 1 - s$.

In the particular case of $g = 2$ we give a complete description of the $X$-rank if $n = 3, 4$; while if $n \geq 5$ we study the $X$-rank of points belonging to the tangential variety of $X$.

This is a joint work with Edoardo Ballico (Università di Trento, Italy).

**6.3.8. Symmetric tensor rank with a tangent vector: a generic uniqueness theorem.**

**Participant:** Alessandra Bernardi.

Let $X_{m,d} \subset \mathbb{P}^N$, $N := \binom{m+d}{m} - 1$, be the order $d$ Veronese embedding of $\mathbb{P}^m$. Let $\tau(X_{m,d}) \subset \mathbb{P}^N$, be the tangent developable of $X_{m,d}$. For each integer $t \geq 2$ let $\tau(X_{m,d},t) \subset \mathbb{P}^N$, be the join of $\tau(X_{m,d})$ and $t - 2$ copies of $X_{m,d}$. In [19] we prove that if $m \geq 2$, $d \geq 7$ and $t \leq 1 + \lfloor \frac{m+2}{m+1} \rfloor + (m+1)$, then for a general $P \in \tau(X_{m,d},t)$ there are uniquely determined $P_1, \ldots, P_{t-2} \in X_{m,d}$ and a unique tangent vector $\nu$ of $X_{m,d}$ such that $P$ is in the linear span of $\nu \cup \{P_1, \ldots, P_{t-2}\}$, i.e. a degree $d$ linear form $f$ (a symmetric tensor $T$ of order $d$) associated to $P$ may be written as
\[ f = L_{t-1}^{d-1}L_t + \sum_{i=1}^{t-2} L_i^d, \quad T = v_{t-1}^{(d-1)}T_1 + \sum_{i=1}^{t-2} v_i^{(d)} \]

with \( L_i \) linear forms on \( \mathbb{P}^m \) (\( v_i \) vectors over a vector field of dimension \( m + 1 \) respectively), \( 1 \leq i \leq t \), that are uniquely determined (up to a constant).

This is a joint work with Edoardo Ballico (Università di Trento, Italy).

### 6.3.9. Parametrization of computational domain in isogeometric analysis: methods and comparison

**Participants:** André Galligo, Bernard Mourrain.

Parameterization of computational domain plays an important role in isogeometric analysis as mesh generation in finite element analysis. In this paper, we investigate this problem in the 2D case, i.e., how to parametrize the computational domains by planar B-spline surface from the given CAD objects (four boundary planar B-spline curves). Firstly, two kinds of sufficient conditions for injective B-spline parameterization are derived with respect to the control points. Then we show how to find good parameterization of computational domain by solving a constraint optimization problem, in which the constraint condition is the injectivity sufficient conditions of planar B-spline parametrization, and the optimization term is the minimization of quadratic energy functions related to the first and second derivatives of planar B-spline parameterization. By using this method, the resulted parameterization has no self-intersections, and the isoparametric net has good uniformity and orthogonality. After introducing a posteriori error estimation for isogeometric analysis, we propose \( r \)-refinement method to optimize the parameterization by repositioning the inner control points such that the estimated error is minimized. Several examples are tested on isogeometric heat conduction problem to show the effectiveness of the proposed methods and the impact of the parameterization on the quality of the approximation solution. Comparison examples with known exact solutions are also presented. This joint work with Régis Duvigneau (EPI OPALE) and Gang Xu (Hangzhou Dianzi University, China) is published in [31].

### 6.3.10. Variational Harmonic Method for Parameterization of Computational Domain in 2D Isogeometric Analysis

**Participants:** André Galligo, Bernard Mourrain.

In isogeometric analysis, parameterization of computational domain has great effects as mesh generation in finite element analysis. In this paper, based on the concept of harmonic map from the computational domain to parametric domain, a variational approach is proposed to construct the parameterization of computational domain for 2D isogeometric analysis. Different from the previous elliptic mesh generation method in finite element analysis, the proposed method focus on isogeometric version, and converts the elliptic PDE into a nonlinear optimization problem. A regular term is integrated into the optimization formulation to achieve more uniform grid near convex (concave) parts of the boundary. Several examples are presented to show the efficiency of the proposed method.

This joint work with Régis Duvigneau (EPI OPALE) and Gang Xu (Hangzhou Dianzi University, China) is published in [36].

### 6.3.11. Warp-based Helical Implicit Primitives

**Participant:** Evelyne Hubert.

Implicit modeling with skeleton-based primitives has been limited up to now to planar skeletons elements, since no closed-form solution was found for convolution along more complex curves. We show that warping techniques can be adapted to efficiently generate convolution-like implicit primitives of varying radius along helices, a useful 3D skeleton found in a number of natural shapes. Depending on a single parameter of the helix, we warp it onto an arc of circle or onto a line segment. For those latter skeletons closed form convolutions are known for entire families of kernels. The new warps introduced preserve the circular shape of the normal cross section to the primitive.
This is joint work with Cédric Zanni and Marie-Paule Cani from the project-team EVASION (INRIA Grenoble Rhône-Alpes / LJK Laboratoire Jean Kuntzmann) which is published in [37].

6.4. National Initiatives

6.4.1. PlantScan3D

PlantScan3D is an ARC between coordinated by the EPI Virtual Plants (UMR DAP, INRIA-CIRAD, Montpellier), with the EPI Galaad (INRIA, Méditerranée) and Evasion (INRIA Rhône-Alpes, Grenoble).

A close collaboration between specialists in plant structures modelling, algebraic geometry, and 3D computer graphic is required to address plant structure reconstruction from laser scanned point clouds. Indeed it is required to take into account efficiently knowledge from topology and geometry to allow mapping and reconstruction of data despite noise, occlusions, and thinness of structure. The objective of the project is to provide as output a compact geometrical model that model smoothly branching point of tubular structure and organs (like leaves). At the end, this model should make it possible an interactive visualisation and automatize different measurement operators needed by biological partners.


6.4.2. GEOLMI

GEOLMI - Geometry and Algebra of Linear Matrix Inequalities with Systems Control Applications - is an ANR project working on topics related to the Geometry of determinantal varieties, positive polynomials, computational algebraic geometry, semidefinite programming and systems control applications.

The partners are LAAS-CNRS, Univ. de Toulouse (coordinator), LJK-CNRS, Univ. Joseph Fourier de Grenoble; INRIA Sophia Antipolis Méditerranée; LIP6-CNRS Univ. Pierre et Marie Curie; Univ. de Pau et des Pays de l’Adour; IRMAR-CNRS, Univ. de Rennes.


6.4.3. ANEMOS

ANEMOS - Advanced Numeric for ELMs : Modeling and Optimized Schemes - is an ANR project devoted to the numerical modelling study of such ELM control methods as Resonant Magnetic Perturbations (RMPs) and pellet ELM pacing both foreseen in ITER. The goals of the project are to improve understanding of the related physics and propose possible new strategies to improve effectiveness of ELM control techniques. The study of spline spaces for isogemetric finite element methods is proposed in this context.

The partners are IRFM, CEA, Cadarache; JAD, University of Nice - Sophia Antipolis; INRIA, Bacchus; Maison de la Simulation CEA-CNRS-INRIA-University of Orsay- University of Versailles St Quentin.

6.5. European Initiatives

6.5.1. FP7 Projects

6.5.1.1. TERRIFIC

Title: Towards Enhanced Integration of Design and Production in the Factory of the Future through Isogeometric Technologies
Type: COOPERATION (ICT)
Defi: PPP FoF: Digital factories: Manufacturing design and product lifecycle manage
Instrument: Specific Targeted Research Project (STREP)
Duration: September 2011 - August 2014
Coordinator: SINTEF, Oslo (Norway)
Abstract: The project aims at significant improvement of the interoperability of computational tools for the design, analysis and optimization of functional products. An isogeometric approach is applied for selected manufacturing application areas (cars, trains, aircraft) and for computer-aided machining. Computer Aided Design and numerical simulation algorithms are vital technologies in modern product development, yet they are today far from being seamlessly integrated. Their interoperability is severely disturbed by inconsistencies in the mathematical approaches used. Efficient feedback from analysis to CAD and iterative refinement of the analysis model is a feature of isogeometric analysis, and would be an essential improvement for computer-based design optimization and virtual product development. Our vision is to provide and disseminate tangible evidence of the performance of the isogeometric approach in comparison to traditional ones in four important application areas as well as addressing interoperability and other issues that necessarily arise in a large-scale industrial introduction of isogeometry.

6.5.1.2. EXCITING

Title: Exact geometry simulation for optimized design of vehicles and vessels
Type: FP7-CP-SST-2007-RTD-1-218536, COOPERATION (TRANSPORTS)
Instrument: Specific Targeted Research Project (STREP)
Duration: October 2008 - September 2011
Coordinator: Jozef Kepler universitet, Linz (Austria)

Abstract: This project focuses on computational tools for the optimized design of functional free-form surfaces. Specific applications are ship hulls and propellers in naval engineering and car components, frames, and turbochargers in the automotive and railway transportation industries. The objective is to base the corresponding computational tools on the same exact representation of the geometry. This should lead to huge benefits for the entire chain of design, simulation, optimization, and life cycle management, including a new class of computational tools for fluid dynamics and solid mechanics, simulations for vehicles and vessels based. This seamless integration of CAD and FEM will have direct applications in product design, simulation and optimization of core components of vehicles and vessels.

6.5.1.3. SAGA

Title: ShApe, Geometry and Algebra, 2008-2012
Type: FP7-PEOPLE-2007-1-1-ITN.
Instrument: Initial Training Network (ITN)
Duration: November 2008 - October 2012
Coordinator: SINTEF (Norway)
Others partners: University of Oslo (Norway); Johannes Kepler Universitaet Linz (Austria); Universidad de Cantabria, Santander (Spain); Vilniaus Universitetas (Lithuany); National and Kapodistrian University of Athens (Greece); INRIA Méditerranée (France); Graphtech (Italy); Kongsberg SIM GmbH (Austria); Missler Software (France);
See also: http://saga-network.eu/.

Abstract: The project aims at promoting the interaction between Geometric Modeling and Real Algebraic Geometry and, in general, at strengthening interdisciplinary and inter-sectorial research and development concerning CAD/CAM. Its objective is also to train a new generation of researchers familiar with both academic and industry viewpoints, while supporting the cooperation among the partners and with other interested collaborators in Europe.

6.5.1.4. DECONSTRUCT

Title: Decomposition of Structured Tensors, Algorithms and Characterization.
Type: PEOPLE (FP7-PEOPLE-2009-IEF)
Instrument: Marie Curie Intra-European Fellowships for Career Development (IEF)
Duration: November 2010 - November 2012
Coordinator: INRIA (France)
Others partners: No.
See also: http://www-sop.inria.fr/teams/galaad/joomla/index.php/international-collaborations-147/172-deconstruct.html

Abstract: Tensors play a wide role in numerous application areas as Signal Processing for Telecommunications, Arithmetic Complexity or Data Analysis. In some applications tensors may be completely symmetric, or symmetric only in some modes, or may not be symmetric. In most of these applications, the decomposition of a tensor into a sum of rank-1 terms is relevant, since tensors of interest have a reduced rank. Most of them are structured i.e. they are either symmetric or enjoy some index-invariance. Lastly, they are often real, which raises open problems concerning the existence and calculation of the decompositions. These issues build the basic bricks of the research program we propose. The classes of tensors described above have a geometric translation in terms of classical algebraic varieties: Segre, Veronese, Segre-Veronese varieties and Grassmannians and their secant varieties. A complete description of equations for those secant varieties and their dimensions is still not known (only dimensions of secant varieties to Veronesean are classified), although they have been studied by algebraic and differential geometers and algebraists for a long period up to now. The aim of this research project is:

• to attack both the description of the ideal of those secant varieties and their dimensions, starting from low dimensions and low degrees,
• to propose algorithms able to compute the rank of structured tensors.

6.6. International Initiatives

6.6.1. Visits of International Scientists

Annie Cuyt and Wen-Shin Lee (University of Antwerpen, Belgium) visited from January 24th to February 2nd to initiate a collaboration on the topic of shape from moments.
George Labahn (University of Waterloo, Canada) visited February 14th-18th and October 3rd-8th to collaborate with Evelyne Hubert on scaling invariants and their application to symmetry reduction of dynamical system (with parameters).
Mark Hickman (University of Canterbury, New Zealand) visited from March to June, as part of his sabbatical year, to collaborate with Evelyne Hubert on the topic of integral and moment invariants and their applications in computer vision.
Nelly Villamizar (University of Oslo) visited us from March 15 to May 15, to collaborate with B. Mourrain on splines spaces, for her secondement in the context of the ITN Marie-Curie SAGA.

Nguyen Tuan Thien (JKU, LINZ) visited us from March to May, to collaborate with B. Mourrain and A. Galligo on parameterization problems in isogeometric analysis, for his secondement in the context of the ITN Marie-Curie SAGA.

6.6.2. Participation In International Programs

6.6.2.1. PAI STAR South Corea collaboration

Participants: Laurent Busé, André Galligo, Evelyne Hubert, Angelos Mantzaflaris, Bernard Mourrain.

The objective of this collaboration is to conduct research in algebraic techniques for solving geometric modeling problems. More specially, we are interested in developing efficient and robust methods to solve non-linear constraints which appear in geometric computation. These methods will be used in applications such as shape design and reconstruction for solving interpolation or approximation problems. A typical area in which we will apply our methods is ship design. Experimentation and validation will lead to open source software implementation.

Collaborators from Seoul National University: Tae-Wan Kim, Sharma Rajiv, Hur Seok, Yeong-hwa Seo.

Tae-Wan Kimm visited INRIA-GALAAD from April 17 to April 23.
6. New Results

6.1. Formal verification of compilers and static analyses

6.1.1. The Compcert verified compiler for the C language

Participants: Xavier Leroy, Sandrine Blazy [project-team Celtique], Alexandre Pilkiewicz.

In the context of our work on compiler verification (see section 3.3.1), since 2005 we have been developing and formally verifying a moderately-optimizing compiler for a large subset of the C programming language, generating assembly code for the PowerPC, ARM, and x86 architectures [5]. This compiler comprises a back-end part, translating the Cminor intermediate language to assembly and reusable for source languages other than C [4], and a front-end translating the CompCert C subset of C to Cminor. The compiler is mostly written within the specification language of the Coq proof assistant, from which Coq’s extraction facility generates executable Caml code. The compiler comes with a 50000-line, machine-checked Coq proof of semantic preservation establishing that the generated assembly code executes exactly as prescribed by the semantics of the source C program.

This year, we improved the Compcert C compiler in several ways:

- The formal semantics for the CompCert C source language was made executable and turned into a reference interpreter. This interpreter is proved sound and complete with respect to the formal semantics. It makes it possible to animate the semantics on test programs, identifying undefined behaviors and enumerating all possible execution orders. Another application is to provide an experimental validation of the semantics itself.

- The top-level statements of compiler correctness were strengthened. In particular, semantic preservation is shown to hold even in the presence of a non-deterministic execution context. Also, we showed that if the source program goes wrong after performing some input/output actions, the compiled code performs at least these actions before continuing with an arbitrary behavior.

- A new optimization pass, redundant reload optimization, was added, improving performance by up to 10% on the x86 architecture.

- A general annotation mechanism was added to observe the values of local program variables at user-specified program points, such observations being guaranteed to produce the same results in the source code and in the compiled code. These annotations can be used to improve the precision of worst-case execution time (WCET) analysis over the compiled code. They can also provide stronger evidence of traceability for code qualification purposes.

Three versions of the Compcert development were publically released, integrating these improvements: versions 1.8.1 in March, 1.8.2 in April, and 1.9 in August.

In parallel, we continued our collaboration with Jean Souyris, Ricardo Bedin França and Denis Favre-Felix at Airbus. They are conducting an experimental evaluation of CompCert’s usability for avionics software, and studying the regulatory issues (DO-178 certification) surrounding the potential use of CompCert in this context. Preliminary results were reported at the Predictability and Performance in Embedded Systems workshop [20]. More detailed results will be presented at the 2012 Embedded Real-Time Software and Systems conference (ERTS’12) [19].

6.1.2. Formal specification and verified compilation of C++

Participants: Tahina Ramananandro, Gabriel Dos Reis [Texas A&M University], Xavier Leroy.
This year, under Xavier Leroy’s supervision and with precious C++ advice from Gabriel Dos Reis, Tahina Ramananandro tackled the issue of formally specifying object construction and destruction in multiple-inheritance languages, especially the C++ flavour featuring non-virtual and virtual inheritance (allowing repeated and shared base class subobjects), and also structure array fields. This formalization consists in specifying, in Coq, a small-step operational semantics for a subset of C++ featuring multiple inheritance, static and dynamic casts, field accesses, and object construction and destruction, and mechanically proving properties about resource management, thus obtaining a formal account of the RAII (Resource Acquisition is Initialization) principle. Moreover, this formalization also studies the impact of object construction and destruction on the behaviour of dynamic operations such as virtual function dispatch, introducing the notion of generalized dynamic type. These results were accepted for publication at the POPL 2012 symposium [29].

Finally, this formalization includes a verified realistic compiler for this subset of C++ to a CFG-style 3-address intermediate language featuring low-level memory accesses in the style of the CompCert RTL language. Following usual compilation schemes and techniques inspired from the Common Vendor ABI for Itanium (which has since been reused and adapted by GNU GCC), the target language additionally features virtual tables to model object-oriented features, and virtual table tables to model the generalized dynamic type changes during object construction and destruction. This verified compiler reuses and extends the results of a previous work on verified C++ object layout by Tahina Ramananandro, Gabriel Dos Reis and Xavier Leroy published this year at the POPL 2011 symposium [28].

6.1.3. Validation of polyhedral optimizations

Participants: Alexandre Pilkiewicz, François Pottier.

The polyhedral representation of loop nests with affine bounds is a unified way to compute and represent a large set of optimizations, including loop fusion, skewing, splitting, peeling, tilling etc. Polyhedral optimizers usually rely on heavily optimized C tools and libraries to manipulate polyhedrons. Those C libraries are, like any other programs, bug prone, which can easily lead to erroneous optimizations. Those two facts—powerful yet error prone—make the formal proof of such optimizations appealing. Proving a full optimizer however would probably be unrealistic: the proof would be terribly challenging, but even writing in Coq an optimizer efficient enough to handle non trivial loop nest might be impossible.

Another option is to write and prove in Coq a validator: after each run of the unproved optimizer—considered as a black box—the validator is used to compare the program before and after optimization to make sure that its semantics—the meaning of the program—has not been change. If the validator does not report an error, we have formal certitude that no bug has been introduced by the optimization.

Alexandre Pilkiewicz, under François Pottier’s supervision, has implemented and proved in Coq such a validator.

6.1.4. A formally-verified parser for CompCert


During a 6-month Master’s internship (M2), Jacques-Henri Jourdan built a formally-verified parser for the C99 language. This parser was obtained through a general method for checking that an LR(1) parser produced by the parser generator Menhir is correct and complete, that is, it conforms exactly to the specification represented by the context-free grammar. This check is carried out by a validator that is implemented in Coq and proved correct, so that, in the end, there is no need to trust Menhir. A paper describing this work was accepted for presentation at the ESOP 2012 conference [24].

6.1.5. Formal verification of an alias analysis

Participants: Valentin Robert, Xavier Leroy.
As part of his 5-month Master’s internship, Valentin Robert developed and proved correct a static analysis for pointers and non-aliasing. This alias analysis is intraprocedural and flow-sensitive, and follows the “points-to” approach of Andersen [40]. An originality of this alias analysis is that it is conducted over the RTL intermediate language of the CompCert compiler: since RTL is essentially untyped, the traditional approaches to field sensitivity do not apply, and are replaced by a simple but effective tracking of the numerical offsets of pointers with respect to their base memory blocks. Using the Coq proof assistant and techniques inspired from abstract interpretation, Valentin Robert proved the soundness of his alias analysis against the operational semantics of RTL.

6.2. Type systems

6.2.1. A type-and-capability calculus with hidden state

Participants: François Pottier, Jan Schwinghammer [Saarland University, Saarbrücken], Lars Birkedal [IT University of Copenhagen], Bernhard Reus [University of Sussex, Brighton], Kristian Støvring [University of Copenhagen], Hongseok Yang [University of Oxford].

During the year 2010, François Pottier developed a machine-checked proof of an expressive type-and-capability system. Such a system can be used to type-check and prove properties of imperative ML programs. The proof, which follows a “syntactic” method, is carried out in Coq and takes up roughly 20,000 lines of code. It confirms that earlier publications by Charguéraud and Pottier [1], [7] were indeed correct, offers insights into the design of the type-and-capability system, and provides a firm foundation for further research. In the first half of 2011, François Pottier wrote a paper that describes the system and its proof in detail. This paper has been submitted for publication [37].

Together with Jan Schwinghammer and other co-authors, François Pottier also worked on a (pencil-and-paper) proof of this type-and-capability system. This proof is based on a “semantic” method and is quite different from the proof mentioned in the previous paragraph. It offers somewhat different insights, and proves (for the first time) that the ideas presented in an unpublished note by Pottier (“Generalizing the higher-order frame and anti-frame rules”, 2009) were correct. A paper that describes this proof has been submitted for publication [39].

6.2.2. Fine-grained static control of side effects in HaMLet

Participants: Jonathan Protzenko, François Pottier.

In the past ten years, the type systems community and the separation logic community, among others, have developed highly expressive formalisms for describing ownership policies and controlling side effects in imperative programming languages. In spite of this extensive knowledge, it remains very difficult to come up with a programming language design that is simple, effective (it actually controls side effects!) and expressive (it does not force programmers to alter the design of their data structures and algorithms). Jonathan Protzenko and François Pottier have recently made significant progress on this topic. They are designing a programming language, tentatively called HaMLet, in the tradition of ML and Caml-Light. The language offers immutable and mutable algebraic data structures and first-class functions. It allows very fine-grained control of ownership and side effects. The project is still at a preliminary stage and no publications have appeared yet.

6.2.3. Partial type inference with first-class polymorphism

Participants: Didier Rémy, Boris Yakobowski [CEA, LIST laboratory], Gabriel Scherer.

The language MLF uses optional type annotations of function parameters and instance bounded polymorphism—quantification over all types that are instances of a given type—to smoothly combine the simple type inference mechanism of ML with the expressive types of System F. In MLF, programs need only type annotations on parameters of functions that are used polymorphically in their body.

While the surface language requires just these very few type annotations, MLF also comes with an internal language, called xMLF, where all type manipulations become explicit so that they can traced during program transformations and symbolic evaluation. The internal language is described in a journal paper [13].
Gabriel Scherer has maintained and improved a prototype implementation of MLF including the elaboration of MLF into xMLF and an extension to higher-order types.

6.2.4. First-class module systems

Participants: Benoît Montagu [University of Pennsylvania], Didier Rémy, Gabriel Scherer.

Singleton kinds are used to handle type definitions in modules. They accurately model the propagation of type definitions through higher-order functor applications. However, type equivalence in the presence of singleton types is hard to formalize and to implement. In his PhD dissertation [48], Benoît Montagu has proposed a new way of checking equivalence in the presence of singleton types, based on expansors. Expansors are eta-expansion constants that are inserted in the source program in such a way that equivalence of two programs becomes equality of their normal forms after insertion of expansors. This approach was described in an article to be submitted to a conference.

Since October, Gabriel Scherer has been working on mixin modules. Mixin modules are an attractive generalization of modules with horizontal composition, a mechanism that allows more flexible construction of modules.

Gabriel Scherer has been studying whether the use of open existential types introduced earlier by Benoît Montagu for first-class modules can be used to simplify the presentation of mixin modules, hoping that they could be given a direct semantics, instead of one by means of elaboration into another language with recursive modules.

6.2.5. Coercion abstraction

Participants: Julien Cretin, Didier Rémy.

Expressive type systems often allow non trivial conversions between types, which may lead to complex, challenging, and sometimes ad hoc type systems. Such examples are the extension of System F with type equalities to model GADT and type families of Haskell, or the extension of System F with explicit contracts. A useful technique to simplify the meta-theoretical studies of such systems is to make type conversions explicit in terms using “coercions”.

We studied $F^\#_\iota$, a language where all type transformations are represented as coercions. This language provides polymorphism as in System F, (upper) bounded polymorphism as in $F_{<:}$, lower bounded polymorphism as in MLF, and $\eta$-expansion as in $F_\eta$. Hence, $F^\#_\iota$ unifies these four languages in a generic framework.

We showed that $F^\#_\iota$ has a type erasing semantics by bisimulation with the lambda calculus. This means that coercions can be dropped before evaluation without changing the meaning of programs.

This work is described in a paper to be presented at the POPL 2012 conference [21] and in a technical report [33].

6.2.6. Kind-level typing in Haskell

Participants: Julien Cretin, Brent Yorgey [University of Pennsylvania], Stephanie Weirich [University of Pennsylvania], José Pedro Magalhães [Utrecht University], Simon Peyton Jones [Microsoft Research Cambridge], Dimitrios Vytiniotis [Microsoft Research Cambridge].

Haskell is a functional programming language with a rich static type system. Programmers use advanced type features to enforce invariants over data structures. This quickly leads to the need for computation in types. Until now, computation at the type level was untyped in Haskell and therefore prone to errors and hard to debug.

We extended the kind level of Haskell with two features already present at the type level: data types and polymorphism. These features are already well-known at the type level, and should remain easy to understand for programmers at the kind level.

Kind polymorphism is now implemented and used in the core language of the
Glasgow Haskell Compiler (GHC). Promotion of data-types is implemented in a branch of GHC. Both extensions are described in a paper to be presented at the TLDI 2012 workshop [31].

6.3. Software specification and verification

6.3.1. Proved time complexity bounds for program components

Participants: Sylvain Dailler, François Pottier.

During a six-month master internship (M2), Sylvain Dailler extended Arthur Charguéraud’s CFML tool with a notion of “time credit”. This allows CFML to be used to prove not only that an algorithm (or a data structure, or a library) is correct, but also that it meets a desired worst-case asymptotic complexity bound. Because CFML is hosted within Coq, these proofs are machine-checked. Sylvain Dailler was able to establish the functional correctness and the time complexity of a library that implements “bags” as circular doubly-linked lists [35].

6.3.2. Hybrid contract checking via symbolic simplification

Participant: Na Xu.

Program errors are hard to detect or prove absent. Allowing programmers to write formal and precise specifications, especially in the form of contracts, is one popular approach to program verification and error discovery. Na Xu formalized and implemented a hybrid contract checker for a subset of OCaml. The key technique is the use of symbolic simplification, which makes integrating static and dynamic contract checking easy and effective. This technique statically verifies that a function satisfies its contract or blames the function violating the contract. When a contract satisfaction is undecidable, it leaves residual code for dynamic contract checking. A paper describing this result will be presented at the PEPM’2012 conference [30]. A technical report version is also available [34].

6.3.3. Tools for TLA+

Participants: Damien Doligez, Leslie Lamport [Microsoft Research], Stephan Merz [EPI VeriDis], Denis Cousineau [Microsoft Research-INRIA Joint Centre], Markus Kuppe [Microsoft Research-INRIA Joint Centre], Hernán Vanzetto [Microsoft Research-INRIA Joint Centre].

Damien Doligez is head of the “Tools for Proofs” team in the Microsoft-INRIA Joint Centre. The aim of this team is to extend the TLA+ language with a formal language for hierarchical proofs, formalizing the ideas in [45], and to build tools for writing TLA+ specifications and mechanically checking the corresponding formal proofs.

This year, the TLA+ project prepared the release of the third version of the TLA+ tools: the GUI-based TLA Toolbox and the TLA+ Proof System, an environment for writing and checking TLA+ proofs. This new release will add many improvements in terms of efficiency, notably with a system of fingerprints to support incremental development of proofs. It will also bring support for new back-ends based on SMT provers (CVC3, Z3, Yices, VeriIT). This extends the range of proof obligations that the system can discharge automatically.

Web site: http://tlaplus.net/.

6.3.4. The Zenon automatic theorem prover

Participant: Damien Doligez.

Damien Doligez continued the development of Zenon, a tableau-based prover for first-order logic with equality and theory-specific extensions. This year, a refactoring of the prover’s architecture was started.

6.4. The Caml language and system

6.4.1. The OCaml system

Participants: Xavier Clerc [team SED], Damien Doligez, Alain Frisch [Lexifi SAS], Jacques Garrigue [University of Nagoya], Fabrice Le Fessant [EPI Asap and OCamlPro start-up company], Jacques Le Normand [Lexifi SAS], Xavier Leroy, Nicolas Pouillard, Pierre Weis [EPI Estime].
This year, we released version 3.12.1 of the OCaml system. This is a minor release that fixes 65 reported bugs and 9 unreported bugs, and introduces 11 small extensions. Damien Doligez acted as release manager for this version.

In parallel, we have been working on the next major release of OCaml. The major innovation is support for generalized algebraic datatypes (GADTs). These non-uniform datatype definitions enable programmers to express some invariants over data structures, and the OCaml type-checker to enforce these invariants. They also support interesting ways of reflecting types into run-time values. GADTs are found in proof assistants such as Coq and in functional languages such as Agda and Haskell. Their integration in OCaml raised delicate issues of partial type inference and principality of inferred types, to which Jacques Garrigue and Jacques Le Normand provided original solutions [43].

Other features in preparation for the next major release include:

- More lightweight first-class modules. Signature annotations over first-class modules can now be omitted when they are determined by the context.
- Better reporting of type errors: shorter but more relevant context is shown; improved tracking of source code locations in modules.
- Improvements in native-code generation, for instance in the case of partial function applications.
- Improvements in the generic hashing primitive and the standard library for hash tables.

6.4.2. Customizable unmarshaling for OCaml

Participants: Pascal Cuq [CEA LIST], Damien Doligez, Julien Signoles [CEA LIST].

In collaboration with members of the CEA LIST laboratory, Damien Doligez developed a Caml library for treating marshaled data by applying user-specified on-the-fly transformations during the unmarshaling process. This library is used in CEA's Frama-C software to support marshaling of hash-consed data. The library was presented at the ML workshop [32].

6.5. Meta-programming

Participants: Nicolas Pouillard, François Pottier.

In an effort to improve meta-programming support (the ability to write programs that manipulate other programs) in programming languages, we have focused first on the issue of binders. Programming with data structures containing binders occurs frequently: from compilers and static analysis tools to theorem provers and code generators, it is necessary to manipulate abstract syntax trees, type expressions, logical formulae, proof terms, etc. All these data structures contain variables and binding constructs.

Nicolas Pouillard, under the supervision of François Pottier, investigated the design of a programming interface for names and binders where the representations of these two types are kept abstract. This interface is sufficiently general to enable a large body of program transformations.

This year, the de Bruijn indices approach has been investigated more in-depth, resulting in a programming interface specialized to safe programming with de Bruijn indices and providing much more precise results than those published in 2010. This work was published at the ICFP 2011 conference [27].

6.6. Formal management of package dependencies

Participants: Roberto Di Cosmo, Ralf Treinen [University Paris Diderot], Jaap Boender [University Paris Diderot], Pietro Abate [University Paris Diderot], Jerôme Vouillon [University Paris Diderot], Stefano Zacchiroli [University Paris Diderot].

Roberto Di Cosmo’s current main line of research is the study and analysis of large component-based software repositories, in particular GNU/Linux-based distributions. These distributions consists of collections of dozens of thousands of software packages, together with metadata, installation and configuration tools, and a variety of different production processes, involving quality assurance at several levels.
Ensuring quality of software assemblies built using these components is a challenging issue: the simple question of knowing whether a single component can or not be deployed turns out to be NP-complete, and yet industry needs to deploy components all the time.

The research currently conducted within the Mancoosi FP7 european project, coordinated by Roberto Di Cosmo, addresses some of the relevant issues, by elaborating sophisticated deployment algorithms and designing specialised installation and configuration languages targeted at enabling transactional capabilities in the tools used to maintain software assemblies built out of GNU/Linux based distributions.

The results of this project are available at http://www.mancoosi.org/ and include four publications this year: one at the CBSE conference [18], which received an ACM distinguished paper award; one at the FSE conference [23], which received an ACM distinguished artifact award; one in the Science of Computer Programming journal [12]; and one at the workshop on Logics for Component Configuration [22].
6. New Results

6.1. Mesh Generation and Geometry Processing

6.1.1. Isotropic 2D Quadrangle Meshing with Size and Orientation Control

Participants: Pierre Alliez, Bertrand Pellenard.

In collaboration with Jean-Marie Morvan from University of Lyon.

We propose an approach for automatically generating isotropic 2D quadrangle meshes from arbitrary domains with a fine control over sizing and orientation of the elements. At the heart of our algorithm is an optimization procedure that, from a coarse initial tiling of the 2D domain, enforces each of the desirable mesh quality criteria (size, shape, orientation, degree, regularity) one at a time, in an order designed not to undo previous enhancements. Our experiments demonstrate how well our resulting quadrangle meshes conform to a wide range of input sizing and orientation fields. [31].

Figure 1. The algorithm takes as input a 2D domain, a sizing field and a cross field (not shown). It then operates on a triangle background mesh: The initialization clusters background mesh triangles so that the tiling roughly meets the size and shape criteria; A relaxation then improves the tiling for shape and orientation while preserving size; A conforming relaxation improves the degree of the tiles and the regularity of the tiling; A series of local parameterizations further improves the degrees and regularity; Barycentric subdivision generates a pure quadrangle mesh; Smoothing finally improves the shape of the quadrangles. We depict the conformance both to the sizing and to the cross field.
6.1.2. An Optimal Transport Approach to Robust Reconstruction and Simplification of 2D Shapes

**Participants:** Pierre Alliez, David Cohen-Steiner.

*In collaboration with Fernando de Goes and Mathieu Desbrun from Caltech.*

We propose a robust 2D shape reconstruction and simplification algorithm which takes as input a defect-laden point set with noise and outliers. We introduce an optimal-transport driven approach where the input point set, considered as a sum of Dirac measures, is approximated by a simplicial complex considered as a sum of uniform measures on 0- and 1-simplices. A fine-to-coarse scheme is devised to construct the resulting simplicial complex through greedy decimation of a Delaunay triangulation of the input point set. Our method performs well on a variety of examples ranging from line drawings to grayscale images, with or without noise, features, and boundaries. [25].

![Figure 2. Robustness to noise and outliers. The input shape (3K points) has sharp corners subtending small angles as well as boundaries. Our reconstruction is perfect for a noise-free input (left); as noise is added (middle, 2% and 2.5% of bounding box), the output degrades gracefully, still capturing most of the sharp angles; even after adding 4K or 4.5K outliers and 2% of noise (right), the reconstruction remains of quality, although artifacts start appearing in this regime.](image)

6.1.3. Anisotropic Delaunay Mesh Generation

**Participants:** Jean-Daniel Boissonnat, Mariette Yvinec.

*In collaboration with Camille Wormser from Google.*

Anisotropic meshes are triangulations of a given domain in the plane or in higher dimensions, with elements elongated along prescribed directions. Anisotropic triangulations are known to be well suited for interpolation of functions or solving PDEs. Assuming that the anisotropic shape requirements for mesh elements are given through a metric field varying over the domain, we propose a new approach to anisotropic mesh generation, relying on the notion of anisotropic Delaunay meshes. An anisotropic Delaunay mesh is defined as a mesh in which the star of each vertex $v$ consists of simplices that are Delaunay for the metric associated to vertex $v$. This definition works in any dimension and allows to define a simple refinement algorithm. The algorithm takes as input a domain and a metric field and provides, after completion, an anisotropic mesh whose elements are shaped according to the metric field. [46]

6.1.4. Triangulating Smooth Submanifolds with Light Scaffolding

**Participants:** Jean-Daniel Boissonnat, Arijit Ghosh.
We propose an algorithm to sample and mesh a $k$-submanifold $\mathcal{M}$ of positive reach embedded in $\mathbb{R}^d$ [45]. The algorithm first constructs a crude sample of $\mathcal{M}$. It then refines the sample according to a prescribed parameter $\epsilon$, and builds a mesh that approximates $\mathcal{M}$. Differently from most algorithms that have been developed for meshing surfaces of $\mathbb{R}^3$, the refinement phase does not rely on a subdivision of $\mathbb{R}^d$ (such as a grid or a triangulation of the sample points) since the size of such scaffoldings depends exponentially on the ambient dimension $d$. Instead, we only compute local stars consisting of $k$-dimensional simplices around each sample point. By refining the sample, we can ensure that all stars become coherent leading to a $k$-dimensional triangulated manifold $\hat{\mathcal{M}}$. The algorithm uses only simple numerical operations. We show that the size of the sample is $O(\epsilon^{-k})$ and that $\hat{\mathcal{M}}$ is a good triangulation of $\mathcal{M}$. More specifically, we show that $\mathcal{M}$ and $\hat{\mathcal{M}}$ are isotopic, that their Hausdorff distance is $O(\epsilon^2)$ and that the maximum angle between their tangent bundles is $O(\epsilon)$. The asymptotic complexity of the algorithm is $T(\epsilon) = O(\epsilon^{-k^2-k})$ (for fixed $\mathcal{M}$, $d$ and $k$).

6.2. Topological and Geometric Inference

6.2.1. Metric graph reconstruction from noisy data

**Participants:** Frédéric Chazal, Marc Glisse.

*In collaboration with Mridul Aanjaneya, Daniel Chen, Leonidas J. Guibas and Dmitriy Morozov.*

Many real-world data sets can be viewed as noisy samples of special types of metric spaces called metric graphs. Building on the notions of correspondence and Gromov-Hausdorff distance in metric geometry, we describe a model for such data sets as an approximation of an underlying metric graph. We present a novel algorithm that takes as an input such a data set, and outputs the underlying metric graph with guarantees. We also implement the algorithm, and evaluate its performance on a variety of real world data sets [26].

6.2.2. Persistence-Based Clustering in Riemannian Manifolds

**Participants:** Frédéric Chazal, Steve Oudot.

*In collaboration with Leonidas J. Guibas and Primoz Skraba.*
We introduce a clustering scheme that combines a mode-seeking phase with a cluster merging phase in the corresponding density map. While mode detection is done by a standard graph-based hill-climbing scheme, the novelty of our approach resides in its use of topological persistence to guide the merging of clusters. Our algorithm provides additional feedback in the form of a set of points in the plane, called a persistence diagram (PD), which provably reflects the prominences of the modes of the density. In practice, this feedback enables the user to choose relevant parameter values, so that under mild sampling conditions the algorithm will output the correct number of clusters, a notion that can be made formally sound within persistence theory.

The algorithm only requires rough estimates of the density at the data points, and knowledge of (approximate) pairwise distances between them. It is therefore applicable in any metric space. Meanwhile, its complexity remains practical: although the size of the input distance matrix may be up to quadratic in the number of data points, a careful implementation only uses a linear amount of memory and takes barely more time to run than to read through the input. [29].

6.2.3. Data-driven trajectory smoothing

**Participant:** Frédéric Chazal.

*In collaboration with Daniel Chen, Leonidas J. Guibas, Xiaoye Jiang and Christian Sommer*

Motivated by the increasing availability of large collections of noisy GPS traces, we present a new data-driven framework for smoothing trajectory data. The framework, which can be viewed of as a generalization of the classical moving average technique, naturally leads to efficient algorithms for various smoothing objectives. We analyze an algorithm based on this framework and provide connections to previous smoothing techniques. We implement a variation of the algorithm to smooth an entire collection of trajectories and show that it performs well on both synthetic data and massive collections of GPS traces. [28].

6.2.4. A Weighted \( k \)-Nearest Neighbor Density Estimate for Geometric Inference

**Participants:** Frédéric Chazal, David Cohen-Steiner.
Motivated by a broad range of potential applications in topological and geometric inference, we introduce a weighted version of the $k$-nearest neighbor density estimate. Various pointwise consistency results of this estimate are established. We present a general central limit theorem under the lightest possible conditions. In addition, a strong approximation result is obtained and the choice of the optimal set of weights is discussed. In particular, the classical $k$-nearest neighbor estimate is not optimal in a sense described in the manuscript. The proposed method has been implemented to recover level sets in both simulated and real-life data.\[12\].

6.2.5. **Deconvolution for the Wasserstein metric and geometric inference**

*Participants:* Frédéric Chazal, Claire Caillerie.

In collaboration with Jérôme Dedecker and Bertrand Michel

Recently, \[17\], \[13\] have defined a distance function to measures to answer geometric inference problems in a probabilistic setting. According to their result, the topological properties of a shape can be recovered by using the distance to a known measure $\nu$, if $\nu$ is close enough to a measure $\mu$ concentrated on this shape. Here, close enough means that the Wasserstein distance $W_2$ between $\mu$ and $\nu$ is sufficiently small. Given a point cloud, a natural candidate for $\nu$ is the empirical measure $\mu_n$. Nevertheless, in many situations the data points are not located on the geometric shape but in the neighborhood of it, and $\mu_n$ can be too far from $\mu$. In a deconvolution framework, we consider a slight modification of the classical kernel deconvolution estimator, and we give a consistency result and rates of convergence for this estimator. Some simulated experiments illustrate the deconvolution method and its application to geometric inference on various shapes and with various noise distributions.\[14\].

6.2.6. **Manifold Reconstruction Using Tangential Delaunay Complexes**

*Participants:* Jean-Daniel Boissonnat, Arijit Ghosh.

We give a new provably correct algorithm to reconstruct a $k$-dimensional manifold embedded in $d$-dimensional Euclidean space \[44\]. The input to our algorithm is a point sample coming from an unknown manifold. Our approach is based on two main ideas: the notion of tangential Delaunay complex and the technique of sliver removal by weighting the sample points. Differently from previous methods, we do not construct any subdivision of the $d$-dimensional ambient space. As a result, the running time of our algorithm depends only linearly on the extrinsic dimension $d$ while it depends quadratically on the size of the input sample, and exponentially on the intrinsic dimension $k$. This is the first certified algorithm for manifold reconstruction whose complexity depends linearly on the ambient dimension. We also prove that for a dense enough sample the output of our algorithm is ambient isotopic to the manifold and a close geometric approximation of the manifold.

6.2.7. **Equating the witness and restricted Delaunay complexes**

*Participants:* Jean-Daniel Boissonnat, Ramsay Dyer, Arijit Ghosh, Steve Oudot.

It is a well-known fact that the restricted Delaunay and witness complexes may differ when the landmark and witness sets are located on submanifolds of $\mathbb{R}^d$ of dimension 3 or more. Currently, the only known way of overcoming this issue consists of building some crude superset of the witness complex, and applying a greedy sliver exudation technique on this superset. Unfortunately, the construction time of the superset depends exponentially on the ambient dimension, which makes the witness complex based approach to manifold reconstruction impractical. This work \[43\] provides an analysis of the reasons why the restricted Delaunay and witness complexes fail to include each other. From this a new set of conditions naturally arises under which the two complexes are equal.

6.2.8. **Reconstructing 3D compact sets**

*Participant:* David Cohen-Steiner.

In collaboration with Frédéric Cazals.
Reconstructing a 3D shape from sample points is a central problem faced in medical applications, reverse engineering, natural sciences, cultural heritage projects, etc. While these applications motivated intense research on 3D surface reconstruction, the problem of reconstructing more general shapes hardly received any attention. This paper develops a reconstruction algorithm changing the 3D reconstruction paradigm as follows.

First, the algorithm handles general shapes i.e. compact sets as opposed to surfaces. Under mild assumptions on the sampling of the compact set, the reconstruction is proved to be correct in terms of homotopy type. Second, the algorithm does not output a single reconstruction but a nested sequence of plausible reconstructions. Third, the algorithm accommodates topological persistence so as to select the most stable features only. Finally, in case of reconstruction failure, it allows the identification of under-sampled areas, so as to possibly fix the sampling.

These key features are illustrated by experimental results on challenging datasets (see Figure 5), and should prove instrumental in enhancing the processing of such datasets in the aforementioned applications. [16].

Figure 5. Reconstruction obtained on a data set with heterogeneous dimensions.

6.3. Data Structures and Robust Geometric Computation

6.3.1. Explicit array-based compact data structures for triangulations

Participants: Olivier Devillers.

In collaboration with Luca Castelli Aleardi (LIX, Palaiseau).

We consider the problem of designing space efficient solutions for representing triangle meshes. Our main result is a new explicit data structure for compactly representing planar triangulations: if one is allowed to permute input vertices, then a triangulation with \( n \) vertices requires at most \( 4n \) references (\( 5n \) references if vertex permutations are not allowed). Our solution combines existing techniques from mesh encoding with a novel use of minimal Schnyder woods. Our approach extends to higher genus triangulations and could be
applied to other families of meshes (such as quadrangular or polygonal meshes). As far as we know, our solution provides the most parsimonious data structures for triangulations, allowing constant time navigation in the worst case. Our data structures require linear construction time, and all space bounds hold in the worst case. We have implemented and tested our results, and experiments confirm the practical interest of compact data structures [47], [35].

6.3.2. Hyperbolic Delaunay triangulations and Voronoi diagrams made practical

Participants: Mikhail Bogdanov, Olivier Devillers, Monique Teillaud.

We show how to compute Delaunay triangulations and Voronoi diagrams of a set of points in hyperbolic space in a very simple way. The algorithm is implemented in an exact and efficient way [34] (see Figure 6).

6.4. Applications

6.4.1. Study of the cosmic web

Participant: Monique Teillaud.

In collaboration with many coauthors: members of the OrbiCG Associate Team (Section 8.3.1.3), Herbert Edelsbrunner (IST Austria, Duke University, and Geomagic Inc.), and others

We introduce a new descriptor of the weblike pattern in the distribution of galaxies and matter: the scale dependent Betti numbers which formalize the topological information content of the cosmic mass distribution (see Figure 7). While the Betti numbers do not fully quantify topology, they extend the information beyond conventional cosmological studies of topology in terms of genus and Euler characteristic used in earlier analyses of cosmological models. The richer information content of Betti numbers goes along with the availability of fast algorithms to compute them. When measured as a function of scale they provide a “Betti signature” for a point distribution that is a sensitive yet robust discriminator of structure. The signature is highly effective in revealing differences in structure arising in different cosmological models, and is exploited towards distinguishing between different dark energy models and may likewise be used to trace primordial non-Gaussianities. In this study we demonstrate the potential of Betti numbers by studying their behaviour in simulations of cosmologies differing in the nature of their dark energy [48], [41]. This work uses previous results obtained in GEOMETRICA [49], [50].
Figure 7. Four $\alpha$-shapes of a Voronoi filament model realization. It concerns a sample of 200000 particles in a periodic box of 50 $h^3$Mpc size with 8 Voronoi cells. From top left to bottom right:

$\alpha = 0.5 \times 10^4, 1.0 \times 10^4, 2 \times 10^4$ and $4.0 \times 10^4$. 
6.5. Software

6.5.1. CGAL

Two major new releases of CGAL, versions 3.8 and 3.9, have been made available in 2011. These releases contain the following new features, involving GEOMETRICA researchers:

- **Generator.** In release 3.8, the package Generator has been extended to provide various point set generators in dimensions higher than 3. It can generate random point sets in/on a sphere, in a cube, and points on a grid [40].

- **Spatial sorting.** Spatial sorting allows to order a set of points to improve the efficiency of incremental randomized algorithms. The spatial sorting package was existing in previous releases, and has been extended to dimensions higher than 3 in release 3.9 [39].

- **3D Mesh Generation.** The mesh generation package was introduced in CGAL 3.5. From release CGAL 3.6, the package offers, after Delaunay refinement, an optional optimization step to either improve the global mesh quality or get rid of slivers. Release CGAL 3.7 includes an interactive demo based on Qt and the code has been optimized for efficiency. Release 3.8 and further [38] offer the possibility to preserve sharp features such as creases and corners when provided in the description of the input domain.

The new release also contains new packages implemented by our CGAL partners and improvements to some existing packages: a detailed list can be found on the CGAL web site.

Two one-week CGAL developers meetings take place each year. The last one, organized in September at INRIA Sophia Antipolis by Mariette Yvinec, gathered 20 participants.
6. New Results

6.1. Discrete logarithms

Participant: Andreas Enge.

In [10], we presented for the first time an algorithm for the discrete logarithm problem in certain algebraic curves that runs in subexponential time less than $L^{1/2}$, namely, $L^{1/3 + \varepsilon}$ for any $\varepsilon > 0$. In [13], we lower this complexity to $L^{1/3}$, showing that the corresponding algebraic curves (essentially $C_{ab}$ curves of genus $g$ growing at least quadratically with the logarithmic size of the finite field of definition, $\log q$) result in cryptosystems that are as easily attacked as RSA or traditional cryptosystems based on discrete logarithms in finite fields. We provide a complete classification of all the curves to which the attack applies.

6.2. Class groups and other invariants of number fields

Participants: Jean-François Biasse, Jean-Paul Cerri, Pierre Lezowski.

J.-F. Biasse has determined a class of number fields for which the ideal class group, the regulator, and a system of fundamental units of the maximal order can be computed in subexponential time $L^{1/3, O(1)}$ (whereas the best previously known algorithms have complexity $L^{1/2, O(1)}$). This class of number fields is analogous to the class of curves described in [13], cf. ref ssec:dlog. The article [18] has been submitted to Mathematics of Computation.

Using new theoretical ideas and his novel algorithmic approach, J.-P. Cerri has discovered examples of generalised Euclidean number fields and of 2-stage norm-Euclidean number fields in degree greater than 2 [11]. These notions, extending the link between usual Euclideanity and principality of the ring of integers of a number field had already received much attention before; however, examples were only known for quadratic fields.

P. Lezowski extended J.-P. Cerri’s algorithm, which was restricted to totally real number fields, to decide whether a generic number field is norm-Euclidean. His procedure allowed to find principal and non norm-Euclidean number fields of various signatures and degrees up to 8, but also to give further insight about the norm-Euclideanity of some cyclotomic fields. Besides, many new examples of generalised Euclidean and 2-stage Euclidean number fields were obtained. The article [25] has been submitted to Mathematics of Computation.

In another direction, norm-Euclidean ideal classes have been studied. They generalise the notion of norm-Euclideanity to non principal number fields. Very few such number fields were known before. A modification of the algorithm provided many new examples and allowed to complete the study of pure cubic fields equipped with a norm-Euclidean ideal class. The article [26] has been submitted to International Journal of Number Theory.

With E. Hallouin, J.-M. Couveignes has studied descent obstructions for varieties [21]. Such obstructions play an important role when one studies families of varieties (e.g. curves of a given genus). Obstructions are often measured by elements in groups like class groups. The theory of stacks provides a more general treatment for these obstructions. Couveignes and Hallouin give the first example of a global obstruction for a variety (that is an obstruction that vanishes locally at every place).

6.3. Number and function field enumeration

Participants: Henri Cohen, Anna Morra, Pieter Rozenhart.
In joint work with R. Scheidler and M. Jacobson, P. Rozenhart has generalized Belabas’s algorithm for tabulating cubic number fields to cubic function fields [30]. This generalization required function field analogues of the Davenport-Heilbronn Theorem and of the reduction theory of binary cubic and quadratic forms. As an additional application, they have modified the tabulation algorithm to compute 3-ranks of quadratic function fields by way of a generalisation of a theorem due to Hasse. The algorithm, whose complexity is quasi-linear in the number of reduced binary cubic forms up to some upper bound \(X\), works very well in practice. A follow-up article [29] describes how to use these results to compute 3-ranks of quadratic function fields, in particular yielding examples of unusually high 3-rank.

H. Cohen and A. Morra [12] have obtained an explicit expression for the Dirichlet generating function associated to cubic extensions of an arbitrary number field with a fixed quadratic resolvent. As a corollary, they have proved refinements of Malle’s conjecture in this context.

### 6.4. Complex multiplication and modularity

**Participants:** Jean-Marc Couveignes, Andreas Enge, Damien Robert.

The book [16] edited by J.-M. Couveignes and B. Edixhoven, with contributions by J.-M. Couveignes, B. Edixhoven, R. de Jong, F. Merkl and J. Bosman, describes the first polynomial time algorithms for computing Galois representations and coefficients of modular forms. Modular forms are tremendously important in various areas of mathematics, from number theory and algebraic geometry to combinatorics and lattices. Their Fourier coefficients, with Ramanujan’s \( \tau \)-function as a typical example, have deep arithmetic significance. Prior to this book, the fastest known algorithms for computing these Fourier coefficients took exponential time, except in some special cases. The case of elliptic curves (Schoof’s algorithm) was at the birth of elliptic curve cryptography around 1985. This book gives an algorithm for computing coefficients of modular forms of level one in polynomial time. For example, Ramanujan’s \( \tau \) of a prime number \( p \) can be computed in time bounded by a fixed power of the logarithm of \( p \). Such fast computation of Fourier coefficients is itself based on the main result of the book: the computation, in polynomial time, of Galois representations over finite fields attached to modular forms by the Langlands programme. Because these Galois representations typically have a nonsolvable image, this result is a major step forward from explicit class field theory, and it could be described as the start of the explicit Langlands programme.

The computation of the Galois representations uses their realisation, following Shimura and Deligne, in the torsion subgroup of Jacobian varieties of modular curves. The main challenge is then to perform the necessary computations in time polynomial in the dimension of these highly nonlinear algebraic varieties. Exact computations involving systems of polynomial equations in many variables take exponential time. This is avoided by numerical approximations with a precision that suffices to derive exact results from them. Bounds for the required precision – in other words, bounds for the height of the rational numbers that describe the Galois representation to be computed – are obtained from Arakelov theory. Two types of approximations are treated: one using complex uniformisation and another one using geometry over finite fields.

With F. Morain, A. Enge has determined exhaustively under which conditions “generalised Weber functions”, that is, simple quotients of \( \eta \) functions of not necessarily prime transformation level and not necessarily of genus 1, yield class invariants [24]. The result is a new infinite family of generators for ring class fields, usable to determine complex multiplication curves. We examine in detail which lower powers of the functions are applicable, thus saving a factor of up to 12 in the size of the class polynomials, and describe the cases in which the polynomials have integral rational instead of integral quadratic coefficients.

With J.-C. Faugère and D. Lubicz, D. Robert has given an explicit construction for a modular correspondance between abelian varieties [14]. This correspondance describes the algebraic relations of ThetaNullWerte of different levels on isogenous abelian varieties. With R. Cosset, D. Robert has then given an algorithm explaining how to construct the corresponding isogeny, when we are given its (maximally isotropic) kernel [20]. This usse a formula by Koizumi for changing the level of the ThetaNullWerte. This is the first algorithm allowing to compute in polynomial time an isogeny between abelian varieties, and a public implementation is available in AVI_SOGENIES.
With K. Lauter, D. Robert has worked on improving the computation of class polynomials in genus 2 by the CRT method. This involves some improvements to detect if the curve is maximal, a better sieving of the primes used, and the use of the CRT over the real quadratic field rather than over $\mathbb{Q}$ for the case of dihedral CM fields.

The main improvements come from using the above isogeny computation, both in order to be able to find a maximal curve from a curve in the correct isogeny class, and in order to find all others maximal curves from one. A preprint describing these improvements is being written, some details are described in the talk http://www.normalesup.org/~robert/pro/publications/slides/2011-04-C2.pdf.

With Reynald Lercier, J.-M. Couveignes has given in [23] a quasi-linear time randomised algorithm that on input a finite field $\mathbb{F}_q$ with $q$ elements and a positive integer $d$ outputs a degree $d$ irreducible polynomial in $\mathbb{F}_q[x]$. The running time is $d^{1+o(1)} \times (\log q)^{5+o(1)}$ elementary operations. The $o(1)$ in $d^{1+o(1)}$ is a function of $d$ that tends to zero when $d$ tends to infinity. And the $o(1)$ in $(\log q)^{5+o(1)}$ is a function of $q$ that tends to zero when $q$ tends to infinity. The fastest previously known algorithm for this purpose was quadratic in the degree. The algorithm relies on the geometry of elliptic curves over finite fields (complex multiplication) and on a recent algorithm by Kedlaya and Umans for fast composition of polynomials.

6.5. Elliptic curve cryptology

Participants: Jean-Marc Couveignes, Vincent Verneuil.

In joint work with C. Clavier, B. Feix, G. Gagnerot and M. Roussellet, V. Verneuil has presented in [15] new side-channel analysis results on the AES. They propose improvements on collision-correlation attacks which require less power traces than classical second-order power analysis techniques. In particular, two new methods are presented and are shown to be efficient in practice on two first-order protected AES implementations. They also mention that other symmetric embedded algorithms can be targeted by these new techniques.

With the same coauthors, V. Verneuil has presented new exponentiation algorithms for embedded implementations in [19]. Embedded exponentiation techniques have become a key concern for security and efficiency in hardware devices using public key cryptography. An exponentiation is basically a sequence of multiplications and squarings, but this sequence may reveal exponent bits to an attacker on an unprotected implementation. Although this subject has been covered for years, they present new exponentiation algorithms based on trading multiplications for squarings. This method circumvents attacks aimed at distinguishing squarings from multiplications at a lower cost than other countermeasures. Finally, they present new algorithms using two parallel squaring blocks which provide one of the fastest exponentiation algorithms.

Together with D. Lubicz, D. Robert has extended their algorithm to compute pairings on abelian varieties using theta functions (published at ANTS 2010) to the case of the ate and optimal ate pairings. This involves a description of the Miller functions in term of theta coordinates and an extension of the addition law using more general Riemann relations in order to compute them. The case of theta functions of level 2 has been optimised by introducing a way to compute “compatible” additions without the need for a square roots. A preprint describing these results is being written, and some details can be found in the talk http://www.normalesup.org/~robert/pro/publications/slides/2011-06-Geocrypt.pdf.

With J.-G. Kammerer, J.-M. Couveignes has given in [22] an appropriate geometric method for studying and classifying encodings into elliptic curves in a cryptographic context. Such encodings were first proposed by Icart in 2009, and later on by Farashahi, Kammerer, Lercier, and Renault. But it was a little bit disappointing to see that it was no more than an application of Tartaglia’s result without any geometrical explanations for the existence of such “parameterisations” of elliptic curves. Couveignes and Kammerer have filled this gap by giving exactly what can be expected from geometry: a clear explanation. Moreover, they unify all the recent “parameterisations” of elliptic curves under the same geometric point of view. The approach described in this article uses dual curves with some results coming from intersection theory. The main originality of this work is that these geometrical tools are employed to explain symbolic computations used in cryptography, that is, encoding on elliptic curves.
5. New Results

5.1. Liability issues in software engineering

Software contracts usually include strong liability limitations or even exemptions of the providers for damages caused by their products. This situation does not favour the development of high quality software because software editors do not have sufficient economic incentives to apply stringent development and verification methods. Indeed, experience shows that products tend to be of higher quality and more secure when the actors in position to influence their development are also the actors bearing the liability for their defects. The usual argument to justify this lack of liability is the fact that software products are too complex and versatile objects whose expected features (and potential defects) cannot be characterised precisely, and which thus cannot be treated as traditional (tangible) goods. Taking up this challenge is one of our objectives [12]: we study liability issues both from the legal and the technical points of view with the aim to put forward a formal framework to (1) define liability in a precise and unambiguous way and (2) establish such liability in case of incident.

Obviously, specifying all liabilities in a formal framework is neither possible nor desirable. Usually, the parties wish to express as precisely as possible certain aspects which are of prime importance for them and prefer to state other aspects less precisely (either because it is impossible to foresee at contracting time all the events that may occur or because they do not want to be bound by too precise commitments). Taking this requirement into account, we provide a set of tools and methods to be used on a need basis in the contract drafting process (as opposed to a monolithic, “all or nothing” approach). Our model is based on execution traces which are abstractions of the log files of the system. In a nutshell, liability is specified as a function taking as parameters a claim and an execution trace and returning a set of “responsible” actors. This set of actors (ideally a singleton) depends on the claim and the errors occurring in the trace. Both errors and claims are expressed as trace properties. The liability function can be made as precise or detailed as necessary by choosing the claims and errors relevant for a given situation [5].

In order to provide a more generic way to define liabilities, we have also introduced a concept of “logical causality” [11]. Causality has been studied for a long time in computer science, but with quite different perspectives and goals. In the distributed systems community, causality is seen essentially as a temporal property. We have defined several variants of logical causality allowing us to express the fact that an event \( e_2 \) (e.g. a failure) would not have occurred if another event \( e_1 \) had not occurred (“necessary causality”) or the fact that \( e_2 \) could not have been avoided as soon as \( e_1 \) had occurred (“sufficient causality”). We have applied these technical definitions of causality to real case studies and related them to the legal views of causality.

As far as legal issues are concerned, we have studied the legal validity of the technical solutions proposed in the project both in terms of legal evidence and allocation of liabilities [6]. Contract templates have been defined in collaboration with lawyers to allow the parties to effectively integrate our results in a legal contract [6].

5.2. Privacy

Despite apparently strong legal protections, many citizens feel that information technologies have invaded so much of their lives that they no longer have suitable guarantees about their privacy. As a matter of fact, many aspects of new information technologies render privacy protection difficult to put into practice. A lot of data communications already take place nowadays on the Internet without the users’ notice and the situation is going to get worse with the advent of “ambient intelligence” or “pervasive computing” [19]. One of the most challenging privacy issues in this context is to reconcile this continuous flow of data with privacy protection. One possible option to improve the situation when data has to be disclosed (or when it is practically impossible to object to its disclosure) is to enhance the obligations of the controllers and enforce more stringent rules on the use of personal data. We have followed this approach, considering both
• technical means to define and enforce obligations and
• possible evolutions of data protection regulations to avoid discriminations based on the use of personal data.

**Technical means: specification and a posteriori verification of obligations**

A major challenge for the formalization of privacy policies is the integration of deontic and temporal operators. Deontic operators are required because privacy policies are typically expressed in terms of obligations and interdictions. Temporal operators are necessary because obligations and interdictions usually come with deadlines: for example, the controller must inform the data subject before forwarding his data to a third party or must delete the data within a given period of time. On the theoretical side, the limitations of Standard Deontic Logic (SDL) have constantly been pointed out, almost since its introduction. However, no other unified mathematical formalization of this logic has been proposed so far. Instead, many specialized logics have been put forward, each aimed at addressing one particular issue. To address this challenge, we have proposed a language called FLAVOR (Formal Language for A posteriori Verification Of legal Rules) for the expression of privacy policies and, more generally, obligations to be fulfilled by organizations. Indeed, organizations have to comply with a growing number of legal rules stemming from law, regulations, corporate policies or contractual agreements. Generally speaking, the actions to be monitored can be checked either a priori or a posteriori. A priori checks are stronger in the sense that they make it possible to ensure that no breach will occur. However, they are too constraining, if not inapplicable, in many situations. Even when they could be implemented, a priori checks are not desirable in situations in which it could be legitimate to bypass the rules. For instance, it is necessary to provide emergency procedures to access personal health records when human lives are at stake, even if the medical practitioner on duty does not have sufficient permissions. The essential features provided by FLAVOR are the possibility to express “contrary to duty” obligations (substitute obligations to be fulfilled in case of breach of the primary obligation), obligations with deadlines and contextual obligations. We have defined a strength ordering between obligations and illustrated the language with typical privacy policy rules [10]. We have also considered the delegation of obligations between actors in [8] and studied the impact of delegation on different types of responsibilities (causal, functional, legal).

**Legal means: privacy and non discrimination**

In order to address the new threats to individual rights that are made possible by the progress of information technologies, we have proposed to distinguish two very different types of data collection [9]:

1. The collection of data as part of formal procedures with clearly identified parties or in the course of clearly identified events, recognized as such by the subjects (e.g. when submitting a file, filling a questionnaire, using a smart card or providing one’s fingerprint to get access to a building).
2. The apparently insignificant and almost continuous collection of data that will become more and more common in the digital society (digital audit trails, audio and video recordings, etc.). This collection may be more or less perceived or suspected by the subject or remain completely invisible and unsuspected. Another worrying phenomenon is the automatic generation of new knowledge using data mining and knowledge inference techniques. In this kind of situation, the subject may ignore not only the process but also the generated knowledge itself, even if this knowledge is about him and could be used to take actions affecting him (e.g. not offering him a job or an insurance contract or adjusting the price of a service up to the level he would be prepared to pay).

The regulations on personal data protection were originally designed to address the first type of situation. Efforts are made to adapt them to the complex issues raised by the second type of data collection but they tend to be increasingly ineffective in these situations. The main cause of this ineffectiveness is their underlying philosophy of a priori and procedural controls. Starting from this observation, we have argued that a possible option is to strengthen a posteriori controls on the use of personal data and to ensure that the victims of data misuses can get compensations which are significant enough to represent a deterrence for data controllers. We have also argued that the consequences of such misuses of personal data often take the form of unfair discriminations and this trend is likely to increase with the generalization of the use of profiles. For this reason, we advocate the establishment of stronger connections between anti-discrimination and data protection laws.
in particular to ensure that any data processing resulting in unfair differences of treatments between individuals is prohibited and is subject to effective compensations and sanctions [9].
6. New Results

6.1. Type theory and formalization of mathematics

6.1.1. Foundational aspects of mechanized proofs

Participants: José Grimm, Loïc Pottier.

We attempt to prove all theorems in the “Theory of Sets” of Bourbaki. The first chapter describes Formal Mathematics, and we show that it can be interpreted in the Coq language, thanks to a bunch of axioms introduced by Carlos Simpson (CNRS, Nice), modulo some modifications. This work that was started in 2009, when J. Grimm was in the Apics project-team. A new formulation of this work using ssreflect has proved more efficient than the initial formulation relying on standard Coq.

The second chapter of Bourbaki covers the theory of sets, per se. It defines ordered pairs, correspondences, unions, intersections and products of a family of sets, as well as equivalence relations. The work of formalizing this chapter comprises 15000 lines of Coq script and is described in a technical report and a paper for the journal of formal reasoning published in 2010.

The third chapter of Bourbaki covers the theory of ordered sets, well-ordered sets, equipotent sets, cardinals, natural integers, and infinite sets; its implementation in Coq is described in [21]. This chapter is longer (22000 lines of code), and there are more exercises (18000 lines of code for about half of the exercises currently implemented).

We also looked at the univalent foundation proposed by V. Voevodsky to provide a new model for equality in type theory and simplified the proof that he proposed to derive extensionality from the univalence axiom.

6.1.2. Group theory (Character theory)

Participants: Georges Gonthier [Microsoft Research], Laurence Rideau, Laurent Théry.

We participate in the collaborative research agreement “Mathematical Components” with Microsoft Research. This project aims at evaluating the applicability of a new approach to mathematical proofs called “small-scale reflection”, especially in the domain of finite group theory [4].

This year, we have initiated the formalisation of the second book of the proof of Feit-Thompson’s theorem. The basic properties of character theories are now covered. This lets us formalised the first 4 chapters of the second book, “Character theory for the Odd Order Theorem” by Peterfalvi.

6.1.3. Proofs in geometry

Participants: Tuan Minh Pham, Yves Bertot.

The work on elementary (synthetic) geometry has been completed. A publication on the topic has also been presented at a conference [19]. This work was also the main content of Tuan Minh Pham’s thesis which was defended in November [5].

6.1.4. Towards constructive algebraic topology

Participants: Laurence Rideau, Maxime Dénès, Yves Bertot.

We have participated in the formalization of a complete chain of computation from an image (as a bitmap) to the corresponding Betti numbers and homology groups. In particular, we improved the formalization of “incidence simplicial matrices” in ssreflect. This work was described in conference article [17].

6.1.5. Computing with polynomials and matrices

Participants: Maxime Dénès, Yves Bertot.
The libraries of the project "Mathematical Components" propose a rather complete formalisation of polynomials and matrices. Unfortunately, these objects cannot be used directly for computing.

We have continued our study of executable algorithms to compute with matrices and polynomials inside Coq. In collaboration with other members of the European project Formath, we have looked at implementation of Strassen-Winograd and Karatsuba for fast matrix multiplication and other algorithms for various kinds of matrix normal forms: Smith normal form, Frobenius, and Jordan normal forms. This work is described in an article that has been submitted for publication.

6.1.6. Regularity of interval matrices

Participants: Guillaume Cano, Yves Bertot.

As part of our work on the regularity of interval matrices, we still needed to formalize the Perron-Frobenius theorem. This year we concentrated on an important lemma for this formalization, the Bolzano-Weierstrass theorem, which requires a usable formalization of general topology, in particular the concept of compact.

6.1.7. Type-based termination

Participants: Jorge Luis Sacchini, Benjamin Grégoire.

The work on this topic has been completed and is described in Jorge-Luis Sacchini’s Ph.D thesis, which was defended in June 2011 [6].

6.1.8. Native compilation of terms with primitive structures

Participants: Mathieu Boespflug [McGill University, Canada], Maxime Dénès, Benjamin Grégoire.

We kept working on the integration of the native compiler of the Ocaml language into a scheme for the efficient reduction of terms in the calculus of inductive constructions. This work is described in a publication at the conference CPP11 in Taiwan [14].

6.2. Proving tools

6.2.1. Connecting an SMT prover and Coq

Participants: Michaël Armand, Germain Faure [project-team Typical], Benjamin Grégoire, Chantal Keller [project-team Typical], Laurent Théry.

Our previous work on integrating SAT technology has been used as a basis to obtain SMT automation within Coq. We are now capable of replaying traces produced by the SMT prover VERIT that deal with conjunctive normal forms, congruence closures, and linear arithmetic. We are actively working on adding quantified formulae. This work is supported by the ANR Decert project. A preliminary version [10] of this work has been presented at the workshop PSATTT’11, a full version [9] at the conference CPP11. The generic exchange proof format [13] for SMT has been presented at the workshop PXT’11.

6.2.2. Geometric Algebras and Automatic Theorem Proving

Participants: Laurent Fuchs [Université de Poitiers], Laurent Théry.

We have completed our work on Grassman-Cayley algebras. This has been published in the post-proceedings of the ADG’10 conference. We are now working on the natural continuation of this work: Clifford’s algebras. We have very encouraging premilary results.

6.2.3. Taylor models in Coq

Participants: Erik Martin-Dorel [project-team Arénaire], Ioana Paşca [project-team Arénaire], Micaela Mayero [Université Paris XIII], Laurence Rideau, Laurent Théry.
Taylor models are a very effective way to approximate real functions with polynomials. We have started a formalisation of these models in the Coq prover. In a first step, we have concentrated our efforts in having a computational version of these models within Coq using native computations, certified floating point and interval arithmetics. Since our first evaluations show that they behave well computationally, we are now working on completing this work with the corresponding correctness proofs. This work is supported by the ANR Tamadi.

6.2.4. Tactics on polynomial equalities: nsatz

Participant: Loïc Pottier.

We started describing in the Coq programming language an efficient algorithm to compute Gröbner bases, similar to the one written in ocaml for the nsatz tactic. We hope to prove it correct and to use it for proofs by reflexion in commutative algebra.

6.2.5. D-Modules

Participant: Loïc Pottier.

We studied normalization of non-commutative polynomials ad exponentials in the Weyl algebra. The normal forms we found are similar with the one described found by Blasiak and Flajolet for graph models.

6.3. Formal study of cryptography

6.3.1. Certicrypt

Participants: Gilles Barthe, Benjamin Grégoire, Sylvain Heraud, Santiago Zanella.

CertiCrypt is a general framework to certify the security of cryptographic primitives in the Coq proof assistant. We completed a machine-checked proof of the security of OAEP (a widely public-key encryption scheme based on trapdoor permutations) against adaptive chosen ciphertext attacks under the assumption that the underlying permutation is partial-domain one-way. This work has been described in a publication at the conference CT-RSA 2011 in San Francisco [12].

6.3.1.1. Easycrypt

Participants: Gilles Barthe [IMDEA], Benjamin Grégoire, Sylvain Heraud, Anne Pacalet, Santiago Zanella.

Based on our experience with Certicrypt, we started last year the development of the tool Easycrypt. The goal of this work is to provide a friendly tool easily usable by cryptographers without knowledge of formal proof assistants. The idea is to use the techniques formally proved in Certicrypt and to call SMT-provers instead of using Coq. We have applied Easycrypt on a variety of academic examples and one bigger example: the proof of IND-CCA security of the Cramer-Shoup cryptosystem. The drawback of this tool is that it provide less guarantees than Certicrypt for the correctness of the proof. To fill this gap we are now able to generate Coq files (based on Certicrypt) allowing to check the validity of Easycrypt proofs. This work has been described in a publication at the conference CRYPTO 2011 in Santa Barbara and has obtained the best paper Award [11].
6. New Results

6.1. Verification of Concurrent Recursive Programs

We have introduced a mathematical model to capture the behavior of concurrent and recursive systems. We have also identified typical properties of these systems that programmers may want to verify. We have come up with a specification language which is powerful enough to express such properties. In fact, we give a framework by which programmers can define their own specification language depending on the specific application as long as the semantics of the operators can be defined in monadic second-order logic. We have shown that checking whether a specified property is satisfiable or whether a given system satisfies a property specified in such a language is decidable with a manageable complexity (double exponential time). The proof technique is so general that it captures the results for various other well studied models. Our results were presented at MFCS’11 [50].

6.2. Product form for Petri nets

While product-form Petri nets have been intensively studied some important questions were left open. In [52], we have solved most of the open problems. We have provided a sound and complete set of rules to synthetise product form Petri nets. We have characterized the complexity class of standard problems (liveness, reachability and cover ability). At last we have proposed a large subclass of product form Petri nets for which the normalising constant (a key quantity) can be efficiently computed. This paper has obtained the outstanding paper award of the ATPN’2011 conference.

6.3. Statistical model checking

We have designed a logic HASL for stochastic systems that can express elaborated performance indices related to path behaviours [46]. We have shown how it can be integrated in the design process of flexible manufacturing systems [47]. We have developed a tool COSMOS for statistical model checking of HASL formula over a stochastic Petri net with general distributions [45]. In parallel, we have developed the first importance sampling method for rare event that still produces a confidence interval (rather than an estimated value) and we have integrated this method in COSMOS [48].

6.4. Symmetry and stochastic systems

We have developed a framework to efficiently solve large Markov decision problems specified by high-level Petri nets [42]. Such a specification allows to decrease the size of the MDP by the analysis of the symmetries of the model. In a purely probabilistic context, we have designed two complementary methods for handling partial symmetries in stochastic high level Petri nets and studied their efficiency on several relevant case studies [41].

6.5. Automata for Data Words

We studied data words, i.e, strings where each position carries both a label from a finite alphabet and some values from an infinite domain. Data words are suitable to model the behavior of concurrent systems with dynamic process creation, as the infinite alphabet can be used to represent an unbounded number of process identifiers. A variety of formalisms, including logic and automata, have been studied to specify sets of data words in the context of verification. However, logic and automata that capture dynamic communicating systems were missing. We closed this gap and developed a quite general logical and automata-theoretic framework for the specification and implementation of sets of data words. On the specification side, we considered a fragment of monadic second-order (MSO) logic, which comes with a predicate to test two word
positions for data equality. As a model of an implementation, we introduced class register automata. Our model combines the well known models of register automata and class memory automata, and it indeed captures dynamic communicating automata, whose semantics can be described as a set of message sequence charts. We studied the realizability problem and show that every formula from the existential fragment of MSO logic can be effectively translated into a class register automaton. These results were presented at CONCUR’11 [49].

6.6. Weighted Logics with Navigation for Trees

We continued our study of the verification of quantitative properties and applications to queries over XML documents. Verification of quantitative systems follow a classical scheme in three steps: specification, modeling, and algorithmics. Hence, we started by exhibiting a specification language. To describe natural qualitative properties, we chose to use, as a fragment, boolean logic like first-order logic or monadic second-order logic. We then encapsulate this properties into the quantitative formalism, allowing sums and products computations in a specified general semiring. In the word case, we obtained very strong results relating this kind of specification/computation languages with the well-known weighted finite automata, and the new weighted pebble automata, which permit to model several interesting quantitative computations over words. We extended these results to trees, and in particular, finite unranked trees or nested words, which are a natural model for XML documents. We published preliminary results in a research report [57], and we have worked on a submission of these results to several conferences. Our next goal is to tackle some of the algorithmic questions that naturally arise in this context, like satisfiability or model checking.

6.7. Contextual Petri nets

Contextual nets are an extension of Petri nets that – unlike ordinary Petri nets – faithfully models concurrent read accesses to shared resources. This is not only interesting from a semantic but also from an algorithmic point of view, as the analysis of such nets can better exploit the fact that concurrent reads are independent and concurrent.

In particular, the unfolding of a contextual net may be up to exponentially smaller in certain situations. While the theoretical foundations of contextual unfoldings were established in [66] and [6], it remained unclear whether the approach could be useful in practice.

Recent work on the theoretical foundations, as well as appropriate data structures and algorithms, has closed this gap. This underlying work has been presented at Concur’11 [53] and has resulted in an efficient tool [59]. More details can be found in a technical report [58].

We are currently exploring applications of these techniques in the areas of verification, diagnosis, and planning. Some preliminary steps have been presented in [54].

6.8. Occurrence net Synthesis

Occurrence nets are a well known partial order model for the concurrent behavior of Petri nets. The causality and conflict relations between events, which are explicitly represented in occurrence nets, induce logical dependencies between event occurrences: the occurrence of an event e in a run implies that all its causal predecessors also occur, and that no event in conflict with e occurs. But these structural relations do not express all the logical dependencies between event occurrences in maximal runs: in particular, the occurrence of e in any maximal run may imply the occurrence of another event that is not a causal predecessor of e, in that run. The reveals relation had been introduced in [33] to express this dependency between two events. In this work, presented at ACSD 2011 [44], we extend the theory in two ways: first, we generalize the reveals relation to express more general dependencies, involving more than two events, and we introduce ERL logic to express these dependencies as boolean formulas. Secondly, we solve the synthesis problem that arises: given an ERL formula \( \phi \), is there a finite occurrence net \( N \) such that \( \phi \) describes exactly the dependencies between the events of \( N \)? The resulting method requires only two synthesis rules.
6.9. Computing Reveals

The *reveals* relation has been introduced in [33] between events in occurrence nets. Essentially, event $a$ is said to *reveal* event $b$ if in any maximal run that contains $a$, $b$ must also occur, be it before, after, or concurrently with $a$, and even if $a$ and $b$ are not causally related. Information of this kind is useful for diagnosis; a sensor for event $a$ may render a sensor for $b$ unnecessary.

In [33], the reveals relation was shown to be decidable for occurrence nets that represent unfoldings of safe Petri nets. However, the upper bound was prohibitive for computing the relation in practice. In [51] we address this problem and drastically reduce the upper bound. We also propose efficient algorithms to actually compute the relation on a given occurrence net.
6. New Results

6.1. Weak Memory Models, Litmus-PPC-WMM tools

Participants: Jade Alglave, Luc Maranget, Pankaj Pawan, Susmit Sarkar [U. of Cambridge], Peter Sewell [U. of Cambridge], Francesco Zappa Nardelli.

Shared memory multiprocessors typically expose subtle, poorly understood and poorly specified relaxed-memory semantics to programmers. To understand them, and to develop formal models to use in program verification, we find it essential to take an empirical approach, testing what results parallel programs can actually produce when executed on the hardware. We describe a key ingredient of our approach, our litmus tool, which takes small “litmus test” programs and runs them for many iterations to find interesting behaviour. It embodies various techniques for making such interesting behaviour appear more frequently. We presented a tool, litmus, to run “litmus tests”. on real hardware at TACAS’11 [31].

An automated exploration of machine memory models (on the dont tool) is submitted to TACAS’12.

During a two month long internship, Pankaj Pawan (IIT Kanpur, India) ported the PPCMEM application from OCaml to JavaScript, and developed a suitable web-interface. This enabled a wide dissemination (including at IBM) of the PPCMEM tool. The tool is available online at: http://www.cl.cam.ac.uk/~pes20/ppcmem/help.html.

6.2. Operational semantics for the memory model of Power (IBM) machines

Participants: Jade Alglave, Luc Maranget, Susmit Sarkar [U. of Cambridge], Peter Sewell [U. of Cambridge], Derek Williams [IBM, Austin].

Exploiting today’s multiprocessors requires high-performance and correct concurrent systems code (optimising compilers, language runtimes, OS kernels, etc.), which in turn requires a good understanding of the observable processor behaviour that can be relied on. Unfortunately this critical hardware/software interface is not at all clear for several current multiprocessors.

We characterise the behaviour of IBM POWER multiprocessors, which have a subtle and highly relaxed memory model (ARM multiprocessors have a very similar architecture in this respect). We have conducted extensive experiments on several generations of processors: POWER G5, 5, 6, and 7. Based on these, on published details of the microarchitectures, and on discussions with IBM staff, we give an abstract-machine semantics that abstracts from most of the implementation detail but explains the behaviour of a range of subtle examples. Our semantics is explained in prose but defined in rigorous machine-processed mathematics; we also confirm that it captures the observable processor behaviour, or the architectural intent, for our examples with an executable checker. While not officially sanctioned by the vendor, we believe that this model gives a reasonable basis for reasoning about current POWER multiprocessors. Our work should bring new clarity to concurrent systems programming for these architectures, and is a necessary precondition for any analysis or verification. It should also inform the design of languages such as C and C++, where the language memory model is constrained by what can be efficiently compiled to such multiprocessors.

This work was presented at PLDI’11 [26]. This operational model is now being enriched for Power machine with loads and link/store conditionals (the Power primitives to write locks), and connected to C++ semantics. It is submitted to PLDI’2012 with many co-authors from Cambridge.

6.3. Restoring Sequential Consistency for x86 and Power machines

Participants: Jade Alglave, Luc Maranget.
Concurrent programs running on weak memory models exhibit relaxed behaviours, making them hard to understand and to debug. To use standard verification techniques on such programs, we can force them to behave as if running on a Sequentially Consistent (SC) model. Thus, we examine how to constrain the behaviour of such programs via synchronisation to ensure what we call their stability, i.e. that they behave as if they were running on a stronger model than the actual one, e.g. SC. First, we define sufficient conditions ensuring stability to a program, and show that Power’s locks and read-modify-write primitives meet them. Second, we minimise the amount of required synchronisation by characterising which parts of a given execution should be synchronised. Third, we characterise the programs stable from a weak architecture to SC. Finally, we present our offence tool which places either lock-based or lock-free synchronisation in a x86 or Power program to ensure its stability.

This work was presented at CAV’11 [30]

6.4. Relaxed-Memory Concurrency and Verified Compilation


We studied the semantic design and verified compilation of a C-like programming language for concurrent shared-memory computation above x86 multiprocessors. The design of such a language is made surprisingly subtle by several factors: the relaxed-memory behaviour of the hardware, the effects of compiler optimisation on concurrent code, the need to support high-performance concurrent algorithms, and the desire for a reasonably simple programming model. In turn, this complexity makes verified (or verifying) compilation both essential and challenging.

This project started in 2010, when we defined a concurrent relaxed-memory semantics for ClightTSO, an extension of CompCert’s Clight in which the processor memory model is exposed for high-performance code, and, building on CompCert, we implemented a certifying compiler from ClightTSO to x86, and proved in Coq several compilation phases. A paper describing our approach has been accepted in POPL [29] 2011.

During 2011 we completed this project by developing correctness proofs for all the compilation phases, and made a public distribution of the compiler, available from http://www.cl.cam.ac.uk/~pes20/CompCertTSO .

In 2011 Zappa Nardelli and Vafeiadis investigated the soundness of fence elimination optimisations for x86TSO. They implemented and proved correct two optimisations that remove redundant fence instructions as compiler passes over RTL in CompCertTSO. Despite an apparent simplicity, these optimisations generate almost optimal code for several standard concurrent algorithms (CompCertTSO does not implement escape analysis, which would enhance the effectiveness of the optimisations), and since they only perform data-flow analysis over the code of each thread (without analysing the full-system thread interactions) they do not suffer form the finite-state and finite control limitation of other approaches. The proof of correctness of one optimisation was challenging has required some for prophecy variable simulation. This work has been published in SAS 2011 [28] and the code is part of CompCertTSO.

A journal version, describing the correctness proof of all the phases of CompCertTSO (including the fence eliminations) as been submitted to the Journal of the ACM [35].

6.5. Compiling C/C++ Concurrency: from C++11 to POWER

Participants: Kayvan Memarian, Francesco Zappa Nardelli.

The upcoming C and C++ revised standards add concurrency to the languages, for the first time, in the form of a subtle relaxed memory model (the C++11 model). This aims to permit compiler optimisation and to accommodate the differing relaxed-memory behaviours of mainstream multiprocessors, combining simple semantics for most code with high-performance low-level atomics for concurrency libraries.

We studied the correctness of two proposed compilation schemes for the C++11 load and store concurrency primitives to Power assembly, having noted that an earlier proposal was flawed. (The main ideas apply also to ARM, which has a similar relaxed memory architecture.)
This should inform the ongoing development of production compilers for C++11 and C1x, clarifies what properties of the machine architecture are required, and builds confidence in the C++11 and Power semantics. A paper describing this work will appear in POPL 2012 [22].

6.6. F*: Secure Distributed Programming with Value-Dependent Types

Participants: Nikhil Swamy [MSR Redmond], Juan Chen [MSR Redmond], Cédric Fournet [MSR Cambridge], Pierre-Yves Strub [MSR-INRIA], Karthikeyan Bhargavan [correspondent], Jean Yang [MIT].

Distributed applications are difficult to program reliably and securely. Dependently typed functional languages promise to prevent broad classes of errors and vulnerabilities, and to enable program verification to proceed side-by-side with development. However, as recursion, effects, and rich libraries are added, using types to reason about programs, specifications, and proofs becomes challenging.

We present F*, a full-fledged design and implementation of a new dependently typed language for secure distributed programming. Unlike prior languages, F* provides arbitrary recursion while maintaining a logically consistent core; it enables modular reasoning about state and other effects using affine types; and it supports proofs of refinement properties using a mixture of cryptographic evidence and logical proof terms. The key mechanism is a new kind system that tracks several sub-languages within F* and controls their interaction. F* subsumes two previous languages, F7 and Fine. We prove type soundness (with proofs mechanized in Coq) and logical consistency for F*.

We have implemented a compiler that translates F* to .NET bytecode, based on a prototype for Fine. F* provides access to libraries for concurrency, networking, cryptography, and interoperability with C#, F#, and the other .NET languages. The compiler produces verifiable binaries with 60x size overhead for proofs and types, as much as a 45x improvement over the Fine compiler, while still enabling efficient bytecode verification.

To date, we have programmed and verified more than 20,000 lines of F* including (1) new schemes for multiparty sessions; (2) a zero-knowledge privacy-preserving payment protocol; (3) a provenance-aware curated database; (4) a suite of 17 web-browser extensions verified for authorization properties; and (5) a cloud-hosted multi-tier web application with a verified reference monitor.

This paper was published at ICFP 2011 [27].

6.7. Security by Typing for Cryptographic Protocol Implementations

Participants: Karthikeyan Bhargavan [correspondent], Cédric Fournet [MSR Cambridge], Andrew D. Gordon [MSR Cambridge], Alfredo Pironti.

We propose to use refinement typing to verify the security of cryptographic protocol implementations. Our method is based on declaring and enforcing invariants on the usage of cryptography. We develop cryptographic libraries that embed a logic model of their cryptographic structures and that specify preconditions and postconditions on their functions so as to maintain their invariants.

We implement the method for protocols coded in F# and verified using F7, our SMT-based typechecker for refinement types, that is, types carrying formulas to record invariants. As illustrated by a series of programming examples, our method can flexibly deal with a range of different cryptographic constructions and protocols [24].

We are currently evaluating this method on a fully-fledged implementation of TLS. While previous uses of typing for cryptographic protocol implementations focused on the symbolic model of cryptography, we use a new technique by Fournet et al to develop computational proofs for our implementations. Our TLS implementation consists of 6000 lines of code. We have currently annotated and verified about half of this implementation.

We recently published a tutorial on our verification method as part of the proceedings of FOSAD 2010, and a journal paper on our type system at TOPLAS [20].
6.8. Authorization for the Social Web: from Formal Analysis to Concrete Attacks

Participants: Chetan Bansal [BITS Goa], Karthikeyan Bhargavan [correspondant], Sergio Maffeis [Imperial College].

Social sign-on and social sharing are becoming an ever more popular feature of web applications. This success is caused in part by the APIs support offered by leading websites such as Facebook, Twitter and Google, and by the openness of standards such as OAuth 2.0. A formal analysis of such protocols must account for malicious websites and their JavaScript, and common website vulnerabilities, such as cross site request forgery and open redirectors.

We present a formal model for web application protocols called WebSpi, implemented as a library for the protocol verification tool ProVerif. We use WebSpi to model and verify several configurations of the OAuth 2.0 protocol. We show that our formal analysis can be used to reconstruct concrete website attacks. Our approach is validated by finding dozens of previously unknown vulnerabilities in popular websites such as Yahoo and Wordpress, when they connect to social networks such as Twitter and Facebook.

We are in discussion with Facebook, Twitter and other websites to address the vulnerabilites found by our analysis. We have submitted a paper describing this work, and plan to release the WebSpi library in 2012.

6.9. Verified Android Cryptographic Applications

Participants: Karthikeyan Bhargavan [correspondant], Quentin Lefebvre [MPRI].

With the emergence of application markets for smartphones, hundreds of third-party applications now use cryptography to protect sensitive user data before storing it on disk or sending it out on the network. However, using cryptographic mechanisms correctly to fulfill a desired security goal is challenging and error-prone, even for experts. Our goal is to build verification tools that developers may use to develop security proofs for their applications.

We show how to verify the security of third-party cryptographic applications written in Java for the Android platform. We first develop symbolic security specifications for classes in the JCA. We can then verify that applications that use these libraries satisfy their security goals, even in the presence of a Dolev-Yao adversary who controls the network, the disk, and potentially other applications on the device. We report preliminary verification results using the Krakatoa verification tool for Java programs.

We presented this work at the ASA workshop 2011 [23].

6.10. Logical Auditing of JavaScript Programs for Security

Participants: Karthikeyan Bhargavan [correspondant], Sergio Maffeis [Imperial College], Ravinder Shankesi [UIUC].

Client side web applications are error-prone and hard to secure, as proven by frequent vulnerability reports. We experiment with using logical annotations as a means to specify inlined security policies for web pages, and we implement a run-time monitoring system that generates a logical trace of the program execution. Feeding the logical trace to external theorem provers, it is possible to detect, post-facto, violations of the security policies, helping the on-line debugging of web applications.

We present JSTY a browser-based logical auditing framework for JavaScript programs. We show how first-order logic contracts can be used to express cryptographic assumptions and security goals for JavaScript programs that use cryptography. We demonstrate our approach on realistic examples, including browser extensions for password management. We find security vulnerabilities in commercial products by logical auditing. This work is currently unpublished.

6.11. Secure Interpreters for Sessions

Participants: James Leifer, Jean Pichon [intern from ENST Telecom ParisTech].
We developed an interpreter for decentralised multi-party sessions. The interpreter takes a graph-based description of a session and provides a high-level interface for sending and receiving messages permissible in the session. The interpreter protects the integrity of session execution in a realistic security setting where an adversary has the ability to: (1) control the network to capture and reinject messages at will, and read and forge messages using leaked cryptographic keys; (2) compromise arbitrary session participants. By producing and verifying cryptographic signatures, the interpreter ensures that all non-compromised participants have consistent views of the session’s execution history.

We previously worked on a session compiler. The compiler took as input “local graphs” described in process-calculus fashion. It produced one F# library for each role, and an F7 interface/specification for this library. If the library typechecked (with the F7 type-checker) against its interface, then it was secure. However, the production of the library was not itself verified, and could fail.

By contrast, this present work is concerned with a session interpreter. This interpreter works like an ML functor (i.e. a compile-time function from modules to modules). It takes as input a module describing a global graph by exposing a specific interface. The application of the interpreter functor to a graph module yields an F$\#$ module that contains the interface/specification. The interpreter functor, being checked against an abstract description of a graph, is typechecked “once and for all”. Because the interpreter functor typechecks against its interface (that contains specifications), it is secure (through refinements). The interpreter functor can then be applied to a concrete session graph to be used.

The interpreter is written in F$\#$, a dialect of ML enriched with refinement types, and its correctness is proven by type annotations.

The interpreter consists of approximately 3000 lines of code.

6.12. Models of Audit Logs

Participants: Karthikeyan Bhargavan, Cédric Fournet [MSR Cambridge], Nataliya Guts, Francesco Zappa Nardelli.

This line of research was accurately described in last year activity report of Moscova. Here we just mention that Nataliya Guts defended her PhD [19] on "Auditability for security protocols" on January 11th, 2011.
PAREO Project-Team

6. New Results

6.1. Improvement of theoretical foundations

6.1.1. Algebraic rewriting

Participant: Yves Guiraud.

With P. Malbos (Institut Camille Jordan, Univ. Lyon 1), in [13], we have used rewriting to give a theoretical setting and concrete formal methods to formalise and give constructive proofs of coherence theorems. The first one is Mac Lane’s classical result on monoidal categories: the new proof we give is a direct application of [7]. Then, cases like symmetric monoidal categories are a first step towards a “rewriting modulo” version of the same work. Finally, we give a new understanding and a constructive proof of the result for cases like braided monoidal categories.

With P. Malbos, in [14], we have generalised the work of [7] to any dimension. We have introduced a notion of polygraphic resolution that generalises both usual algebraic resolutions, in combinatorial algebra, but also, more surprisingly, normalisation strategies, as used in rewriting theory and, in particular, in rule-based programming languages. Thus, a functional program can be mathematically defined as a complete cellular model of the functions it computes. This gives a strong mathematical background to the notion of program, together with a constructive way to build resolutions from convergent polygraphs. This work has been presented during an invited conference at the International Congress on Operads and Universal Algebra, held in Tianjin, China, in July 2010. Those results will be further explored to give a mathematical description of the strategies used in Tom, in order to develop methods from algebraic topology to study their computational properties, like termination and complexity.

With S. Gaussent (Institut Élie Cartan, Univ. Nancy 1) and P. Malbos, in [26], we have applied higher-dimensional rewriting methods to actions of monoids on categories. The objective was to compute, starting from a presentation of a monoid by generators and relations, a “homotopy basis” that generates all the relations between the relations: this is exactly the piece of data one needs to use a presentation to give a practical definition of action of the monoid on categories. We show that methods from rewriting theory (Squier’s theorem, Knuth-Bendix completion procedure), adapted to Burroni’s polygraphs, can be used to compute that homotopy basis. In particular, we get a new, algebraic and constructive proof of a result by Deligne on actions of braid groups.

6.1.2. Certification of induction proofs

Participant: Sorin Stratulat.

We have defined a methodology for validating implicit induction proofs. In collaboration with Vincent Demange, we gave evidence of the possibility to perform implicit induction-based proofs inside certified reasoning environments, as that provided by the Coq proof assistant. This is the first step of a long term project focused on 1) mechanically certifying implicit induction proofs generated by automated provers like Spike, and 2) narrowing the gap between automated and interactive proof techniques by devising powerful proof strategies inside proof assistants that aim to perform automatically multiple induction steps and to deal with mutual induction more conveniently. Contrary to the current approaches of reconstructing implicit induction proofs into scripts based on explicit induction tactics that integrate the usual proof assistants, our checking methodology is simpler and fits better for automation. The underlying implicit induction principles are separated and validated independently from the proof scripts that consist in a bunch of one-to-one translations of implicit induction proof steps. The translated steps can be checked independently, too, so the validation process fits well for parallelisation and for the management of large proof scripts. Moreover, our approach is more general; any kind of implicit induction proof can be considered because the limitations imposed by the proof reconstruction techniques no longer exist. This result has been firstly presented at the Poster session of 2010 Grande Region Security and Reliability Day, Saarbrucken.
Based on the previous result, an implementation that integrates automatic translators for generating fully checkable Coq scripts from Spike proofs is reported in [55]. The induction ordering underlying the Spike induction principle was defined using COCCINELLE [34], a Coq library well suited for modelling mathematical notions needed for rewriting, such as term algebras and RPO. COCCINELLE formalises RPO in a generic way using a precedence and a status (multiset/lexicographic) for each function symbol. Spike automatically generates a term algebra starting from Coq function symbols which preserve the precedence of the original Spike symbols. Many useful properties about the RPO orderings have been already provided by COCCINELLE. On the other hand, the induction ordering was modelled as a multiset extension of RPO and only few properties about it were provided by COCCINELLE. We have proved useful lemmas about it and added them to COCCINELLE, for example, the multiset extensions of RPO is stable under substitutions. Finally, every single inference step derived with a restricted version of Spike can be automatically translated into equivalent Coq script. The restricted inference system was powerful enough to prove properties about specifications involving mutually defined functions and to validate a sorting algorithm. The scripts resulted from the translation of these proofs were successfully validated by Coq.

Another improvement of the COCCINELLE library was the redefinition of the RPO ordering in order to consider precedencies that take into account equivalent function symbols. The new release of CoLoR (http://color.inria.fr) that compiles with the new version 8.3 of Coq includes the updated version of COCCINELLE. This improvement allowed Rainbow, a program developed within the CoLoR project that uses COCCINELLE’s formalisation for RPO, to certify more than 30 additional proofs from the set of CPF files generated during the 2009 annual termination competition (http://termination-portal.org/wiki/Termination_Competition).

In [21], we reported new improvements in order to certify implicit induction proofs concerning industrial-size applications, in particular the validation proof of a conformance algorithm for the ABR protocol [51]. An interactive proof using PVS [54] was firstly presented in [52], then it has been shown in [53] that more than a third of the user interactions can be avoided using implicit induction techniques, Spike succeeding to prove 60% of the user-provided lemmas automatically. Now, a simpler but more restrictive version of the Spike inference system has been shown powerful enough to prove 2/3 out of these lemmas. Moreover, any generated proof has been automatically translated into a Coq script, then automatically certified by Coq. We stressed the importance of the automatic feature since the proof scripts are in many cases big and hard to manipulate by the users. The bottom-line is that these improvements allowed us to certify big proof scripts in a reasonable time, 20 times faster than in [55].

6.2. Integration of formal methods in programming languages

6.2.1. Formal islands and Tom

Participants: Jean-Christophe Bach, Horatiu Cirstea, Pierre-Etienne Moreau, Claudia Tavares.

In [1] we have proposed a framework which makes possible the integration of formally defined constructs into an existing language. The Tom system is an instance of this principle: terms, first-order signatures, rewrite rules, strategies and matching constructs are integrated into Java and C for instance. The high level programming features provided by this approach are presented in [48]. The Tom system is documented in [27]. A general overview of the research problem raised by this approach are presented in [47].

One interest of Tom is to make the compilation process independent of the considered data-structure. Given any existing data-structure, using a formal anchor definition, it becomes possible to match and to apply transformation rules on the considered data-structures.

During the internship of Maxime Gabut, we have developed a new approach for parsing island based languages. We have clearly separated the grammars of the two considered languages (host and island languages). At present, there is one scanner for each language, but this is not powerful enough to handle the general case. We are currently studying another approach based on a scanner able to recognized tokens from both languages.
During the internship of Alexandre Papin, we have developed a new backend for ADA 2005. The mid-term project is to use tom-ADA to develop new optimisation phases of the GNAT-ADA compiler.

We have defined a new type system for Tom which allows to declare first-order signatures with subtyping. This is particularly useful to encode inheritance relations into algebraic terms. Considering Java as the host language, in [25] we present this type system with subtyping for Tom, that is compatible with Java’s type system, and that performs both type checking and type inference. We propose an algorithm that checks if all patterns of a Tom program are well-typed. In addition, we propose an algorithm based on equality and subtyping constraints that infers types of variables occurring in a pattern. Both algorithms are exemplified and the proposed type system is showed to be sound and complete.

A first application consists in defining algebraic mappings for implementations of meta-models generated by the Eclipse Modeling Framework.

6.3. Practical applications

6.3.1. Security policies specification and analysis

Participants: Tony Bourdier, Horatiu Cirstea, Hélène Kirchner, Pierre-Etienne Moreau.

Access control policies, a particular case of security policies should guarantee that information can be accessed only by authorized users and thus prevent all information leakage. We proposed [18] a framework where the security policies and the systems they are applied on are specified separately but using a common formalism. This separation allows not only some analysis of the policy independently of the target system but also the application of a given policy on different systems. In this framework, we propose a method to check properties like confidentiality, integrity or confinement over secure systems based on different policy specifications.

We also propose an approach [12] where the specification of a security policy is split into two distinct elements: a security model and a configuration. The security model (expressed as an equational problem) describes how authorization requests must be evaluated depending on security information. The configuration (expressed as a rewriting system) assigns values to security information. We show that this separation eases the formal analysis of security policies and makes it possible to automatically convert a given policy with respect to an alternative security model.

6.3.2. Symbolic analysis of network security policies

Participants: Tony Bourdier, Horatiu Cirstea.

In computer networks, security policies are generally implemented by firewalls. We propose in [16] an original framework based on tree automata which can be used to specify firewalls and which takes into account the network address translation functionality. We show that this framework allows us to perform structural analysis as well as query analysis and comparisons over firewall policies.

We have extended the above formalism to take into account the composition of firewalls. In our approach [17] all the components of a firewall, i.e. filtering and translation rules, are specified as rewrite systems. The same formalism can be used to formally describe the composition of firewalls (including routing) in order to build a whole network security policy. The properties of the obtained rewrite systems are strongly related to the properties of the specified networks and thus, classical theoretical and practical tools can be used to obtain relevant properties of the security policies. We showed that the proposed specifications allow us to handle usual problems such as comparison, structural analysis, and query analysis over complete networks.

6.3.3. Model transformation using rewriting

Participants: Jean-Christophe Bach, Pierre-Etienne Moreau.

New development chains of critical systems rely on Domain Specific Modeling Languages (DSML) and on qualifiable transformations (insurance that a transformation preserves interesting properties). To specify and to make such transformations we have started to extend Tom.
A first part of this extension is an EMF\textsuperscript{1} mapping generator which allows to use Tom with EMF. The idea of this tool is to generate Tom mappings (i.e. an algebraic view) by introspecting EMF generated Java code. These mappings can then be used to describe transformations of models that have been created in Eclipse. Tom-EMF is documented and available in the Tom source distribution.

The second part of this extension consists in the addition of new Tom language constructs to express transformations of models. Studying several use cases\textsuperscript{2}, we have already handwritten the code that should be generated. We are currently considering abstraction of the code in order to make the generation of this one automatic in a near future.

### 6.3.4. A constraint language for algebraic terms

**Participants:** Horatiu Cirstea, Pierre-Etienne Moreau, François Prugniel.

With the development of model transformation formalisms and tools, a need for checking the result of transformations appears. For UML, the Object Management Group developed OCL\textsuperscript{3} but there are no equivalent formalisms for the algebraic views of models used in Tom-EMF. Starting from the OCL experiences, we have proposed an extension of Tom with a constraint language for algebraic signatures [23].

We show that the constructs of this new constraint language inspired from OCL and XPath can be naturally described using strategic rewriting and consequently, constraint checking can be performed by executing Tom programs. This kind of language is a major asset for debugging tools and a great step to obtain trustworthy compilers. Indeed, the first important application for this language will be the Tom compiler itself.

---

\textsuperscript{1}Eclipse Modeling Framework, http://www.eclipse.org/modeling/emf/

\textsuperscript{2}available on the Quarteft subversion repository: https://gforge.enseeiht.fr/projects/quarteft/

\textsuperscript{3}Object Constraint Language, http://www.omg.org/spec/OCL/2.2/
6. New Results

6.1. Compilation techniques for synchronous languages

- The paper *Modular Static Scheduling of Synchronous Data-flow Programs* by M. Pouzet and P. Raymond (VERIMAG Grenoble) has been selected among the two best papers at EMSOFT 2009. An extended version is published in a special issue of the *Journal of Design Automation for Embedded Systems*, in 2010. This work solves a 20 years problem raised by P. Raymond in 88 [59].

- Significant extensions have been provided to the Heptagon language. C. Pasteur contributed a memory optimization pass, combining a static analysis with user-given annotations (a paper on this subject is submitted to LCTES’2012), and the integration of discrete controller synthesis, developed and used by the SARDES team at INRIA Rhones-Alpes. Other work-in-progress extensions and experiments conducted inside the PARKAS team include the generation of VHDL code, parallel code generation, and n-synchronous code generation from Lucy-n.

6.2. Semantics and Implementation of Hybrid System Modelers

Hybrid systems modelers have become the corner stone of embedded system development, with Simulink a de facto standard and Modelica a new player. They allow both discrete controllers and their continuous environments to be expressed in a single language. Despite the availability of such tools, there remain a number of issues related to the lack of reproducibility of simulations and to the separation of the continuous part, which has to be exercised by a numerical solver, from the discrete part, which must be guaranteed not to evolve during a step. Such tools still raise a number of issues that, we believe, require more fundamental understanding.

In collaboration with Albert Benveniste and Benoit Caillaud (INRIA Rennes) we have proposed using nonstandard analysis as a semantic domain for hybrid systems. Non standard analysis is an extension of classical analysis in which infinitesimals can be manipulated as first class citizens. This allows us to provide a denotational semantics and a constructive semantics for hybrid systems, thus establishing simulation engines on a firm mathematical basis. In passing, we cleanly separate the job of the numerical analyst (solving differential equations) from that of the computer scientist (generating execution schemes).

- In late 2010, we presented in 49th Conference on Design and Control in 2010 the use of nonstandard semantics as a semantical ground for a hybrid synchronous language.

- Since the, we have extended this work in the following directions: 1/ a synchronous Kahn semantics for hybrid programs. Programs are viewed as synchronous ones running on an infinitely fast discrete base clock of the form $BaseClock(\partial) = \{n\partial | n \in \mathbb{N}\}$, with infinitesimal step $\partial$ and $\mathbb{N}$ the non-standard extension of $\mathbb{N}$. 2/ the definition of a standardization principle that gives sufficient conditions for a hybrid program to be standardizable. Under these conditions, the semantics corresponds to the semantics using super-dense time [44], [40], [42] for hybrid systems defined in [10]. 3/ a large amount of experimentations with Simulink to illustrate some of its pitfalls concerning in particular the treatment of zero-crossing cascades. This work is detailed in a long paper appearing in the *Journal of Computer Science and Systems* [1], in 2011.

---

12 https://modelica.org/
• Starting from a minimal, yet full-featured, Lustre-like synchronous language, we have proposed a conservative extension where data-flow equations can be mixed with ordinary differential equations (ODEs) with possible reset. A type system is proposed to statically distinguish discrete computations from continuous ones and to ensure that signals are used in their proper domains. The extended data-flow language is realized through a source-to-source transformation into a synchronous subset, which can then be compiled using existing tools into routines that are both efficient and bounded in their use of memory. These routines are orchestrated with a single off-the-shelf numerical solver using a simple but precise algorithm which treats causally-related cascades of zero-crossings. We have validated the viability of the approach through experiments with the SUNDIALS library. The basis of this work has been presented at the ACM International Conference on Languages, Compilers, Theory of Embedded Systems, 2011 [3].

• This work shows that it is possible to define a language which combines both the expressiveness of synchronous a synchronous language and that of ODEs where continuous solvers are approximated by a black-box solver. The most noticeable result was to recycle several techniques developed for synchronous languages: Kahn semantics, compilation techniques, static analysis. During year 2011, we extended the basic language with with hierarchical automata. This work has been presented at the ACM International Conference on Embedded Software, 2011 [2].

• In parallel with these theoretical works, M. Pouzet and T. Bourke have developed during year 2011 a new synchronous language and its compiler. The language, called ZELUS, extends a synchronous language with ODEs. It is first-order, functional and which mixes continuous-time and discrete-time signals. The expressiveness is that of (the first-order subset of) Lucid Synchrone (e.g., type inference and polymorphism, mix of data-flow and hierarchical automata) and ODEs with possible reset. Continuous trajectories are computed by a black-box numerical solver and we made our experiments with SUNDIALS.

\[ \text{6.3. N-Synchronous Languages} \]

The n-synchronous model introduced a way to compose streams which have almost the same clock and can be synchronized through the use of a finite buffer.

We have designed the language Lucy-n to program in this model of computation [4], [48]. This language is similar to the first order synchronous data-flow language Lustre in which a buffer operator is added. A dedicated type system allows to check that programs can be executed in bounded memory and to compute the buffers sizes needed. Technically it is done through the introduction of a subtyping constraint at each bufferization point.

To solve the subtyping constraints we have defined an algorithm that uses an improved version of the state-of-the-art abstraction. We have proved the correctness properties of this new abstraction in Coq.

We also worked on new typing algorithms that do not use clock abstraction and thus allows to model Latency Insensitive Design [5] in Lucy-n [48].

• A. Guatto, together with L. Mandel and M. Pouzet, worked on the code generation for Lucy-n. They investigated two approaches. The first one was to use the schedules and buffer sizes computed by the compiler to generate a classical Lustre program. The second approach was to define a dynamic scheduling protocol similar to the ones used in latency insensitive designs.

• L. Mandel, in collaboration with F. Plateau (Prove&Run), developed a new resolution constraint algorithm for the clocking of Lucy-n programs [4]. Even if the new algorithm is less efficient that the one using abstraction, it has the advantage to be more precise and thus to accept more programs.

• L. Mandel, F. Plateau and M. Pouzet have extended the Lucy-n language with a new operator to be able to model Latency Insensitive Designs. Thanks to the new resolution constraint algorithm, the Lucy-n compiler is able to compute static schedules for such designs [5].

\[ \text{13 https://computation.llnl.gov/casc/sundials/main.html} \]
6.4. Synchronous Circuits

- Followed up on J. Vuillemin’s result that the XOR variant of non-deterministic automata can be efficiently minimized [65], we explore this newly opened branch of computational automata theory. One contribution is a Decision Diagram for Boolean functions which has minimal dimension: this is appealing for both the verification and synthesis of memory-less circuits. Parts of this recent work were presented at the Boole ANR cooperation [8] and Synchron 2011.

- An extension of Boolean Decision Diagrams to integer representation and operations is given in [66]: we pursue software experimentations with arithmetics on such gigantic (yet sparse) numbers.

- The paper [32] was translated to English from the 1974 original and published in honor of Gilles Kahn.

6.5. Reactive Programming

ReactiveML is an extension of OCaml with synchronous concurrency, based on synchronous parallel composition and broadcast of signals. The goal is to provide a general model of deterministic concurrency inside a general purpose functional language to program reactive systems. It is particularly suited to program discrete simulations, for instance of sensor networks. The current focus of the research is being able to simulate huge systems, composed of millions of agents, by extending the current purely sequential implementation in order to be able to take advantage of multi-core and distributed architectures.

A first experiment consisted in creating a parallel runtime without any modification to the language. As the OCaml language on which ReactiveML is based (the ReactiveML compiler generates OCaml code) does not allow to create parallel programs communicating via shared memory, the new runtime was written in the F# language, part of Microsoft .Net environment. As the language is almost source to source compatible with OCaml, the ReactiveML compiler was left untouched. Several parallel runtimes were written, using traditional task scheduling techniques like work stealing or directly using light task mechanisms available in F#. This experiment demonstrated many speedup opportunities by parallelizing the runtime but also highlighted several problems and limitations of the language. Although this experiment was very useful in understanding the stakes of parallelizing ReactiveML, the performance gap between OCaml and F# (OCaml generates sequential code that is about 10 times slower) makes this version of the runtime of little practical use. We then proposed an extension of ReactiveML called clock domains. It consists in creating local notions of instants that are invisible from the outside. This extension should solve most of the problems raised by the previous experiment and help the parallelization of the language. The sequential runtime was adapted to this extension and a distributed version, using processes communicating via message passing, is currently being developed.

In collaboration with P. Attar (INRIA), F. Boussinot (INRIA) and J.-F. Susini (CNAM), L. Mandel worked on the design of DSL [9], a script language for the orchestration of concurrent programs. L. Mandel developed in ReactiveML and JoCaml an interpreter for the DSL language.

6.6. Polyhedral Compilation

L.-N. Pouchet presented fundamental advances in the construction and linear optimization of multidimensional affine transformation spaces at the POPL 2011 conference, in collaboration with A. Cohen and colleagues from Paris Sud University, Louisiana State University and Ohio State University.

Our team is actively integrating the polyhedral optimization framework in two production compilers: the Graphite framework in GCC and the Polly framework in LLVM. We are also working towards using the polyhedral framework to target GPGPU and manycore architectures, and to generate aggressively optimized code starting from high level languages. New isl-based version of Graphite and Polly have been contributed, enabling state-of-the-art affine transformations in GCC and LLVM, respectively. Dramatic performance improvements are expected in 2012, impacting the upcoming GCC 4.8 and LLVM 3.1, respectively.

K. Trifunovic, F. Li and A. Cohen, in collaboration with R. Ladelsky from IBM Research Haifa, presented a paper about some of these progresses at the GROW 2011 workshop [7].
Among the challenges that arise when adapting the polyhedral framework to production compilers, Riyadh Baghdadi has been working on memory-based dependences. Part of it is a practical compiler construction issue, where upstream passes such as the transformation to three-address code and PRE/CSE introduce new scalar variables leading to additional memory-based dependences. The other difficulty is to identify a profitable tradeoff between memory expansion (privatization, renaming) and parallelism. Memory-based dependences not only increase the complexity of the optimization but most importantly, they reduce the degree of freedom available to express effective loop nest transformations, limiting the overall effectiveness of the polyhedral framework. We designed and implemented a technique that solves this problem by allowing a compiler to relax the constraint of memory-based dependences on loop nest transformations and that does not incur the memory footprint overhead of scalar and array expansion. The proposed technique is based on the concept of polyhedral live range interval interference. While previous polyhedral optimization techniques could not achieve any speedups in benchmarks with scalar variables. This technique enabled a speedup of up to $16\times$ on numerical kernels from the Polybench benchmark suite (on a 24-core machine).

### 6.7. Parallel Data-Flow Programming

A. Pop defended his PhD thesis in September at MINES ParisTech, on a data-flow, streaming extension of OpenMP. Semantical, compilation and runtime system aspects have been covered in depth. Early results were published at HiPEAC 2011 and presented by A. Pop, in collaboration with A. Cohen. Follow-up work include the maturation of the proposed semantics and language extensions, with a thorough implementation and experimentation.

In parallel, F. Li, in collaboration with A. Pop and A. Cohen, explores the automatic compilation of SSA programs into dynamic data-flow parallelism, and the integration of streaming dependences to extend the method to non-scalar data flow. A paper has been published at WIR 2011 (workshop associated to CGO 2011), and a comprehensive, modular compilation method for arbitrary control flow, will be presented at MULTIPROG 2012 (associated with HiPEAC 2012).

Classical compilation techniques, found in Lustre, Scade, Lucid Synchrone, and all the dataflow synchronous languages, generate very efficient sequential code. Thus our main goal is to allow parallel code generation without changing the generation of the sequential code. To this matter, we introduced in the dataflow synchronous setting the famous asynchronous calls bundled with futures, which date back to MultiLisp designed by R. Halstead in the early 1980. It allows to separate the request of a computation from the actual use of its result. This approach has two main advantages. First, the compilation of these asynchronous calls is implemented by a simple wrapper encapsulating the called sequential code. It permits full compatibility with existing generated code. The futures are treated like usual values, so except for the asynchronous calls, we use the known sequential code generation. Second, this asynchronous calls and futures are only annotations and may be fully erased without changing the semantics of the program.

L. Gérard implements this proposition in our Heptagon compiler. The first backend was done in Java, as a proof of concept, using the threads and futures of the Java language. More efficient back-ends are being explored, using OpenMP stream-computing extensions and the TStar data-flow primitives.
6. New Results

6.1. The Focused Calculus of Structures

Participants: Kaustuv Chaudhuri, Nicolas Guenot, Lutz Straßburger.

The *sequent calculus* is a proof system for logic that has many nice properties from a proof search perspective, the most famous being the subformula property that is essential to tame the search space. In recent years, the *focusing* property of sequent systems has become another useful property, both for shrinking the search space and to improve the representation of proofs. However, the sequent calculus does have some limitations. Primarily, not all logics have *analytic* (i.e., cut-free) proof systems, which are the *sine qua non* of proof search. A less obvious but equally bothersome limitation is that cut-free sequent proofs tend to contain large repeated sub-proofs.

To remedy these deficiencies, one can use the *calculus of structures*, a proof system that allows inferences anywhere inside a formula. This system can represent many more logics than the sequent calculus and can produce better (i.e., usually smaller) proofs because it can avoid sharing large subformulas. Nevertheless, because the rules of the calculus of structures have finer granularity than sequent rules, it has more nondeterminism during search.

In this work, we show how to transplant the focusing result from the sequent calculus to the calculus of structures [19]. We thus improve the search capabilities of the calculus of structures, including the ability to go back and forth between focused sequent proofs and focused calculus of structures proofs, but we retain all the distinguishing features of the calculus of structures.

In particular, we preserve the ability to permute contractions below all other rules (first observed for the calculus of structures in [55], [31]). This permutation enables a two-stage normal form of proofs. The first stage contains only contractions, which increases the complexity of the formulas and is therefore a potential source of unbounded search; this phase needs to be recorded in order to reconstruct proofs by bounded search. The second stage that contains the remaining logical rules (except contraction) is strictly bounded and finite—hence decidable—and can be reconstructed if omitted from the proof object. Thus, we have the potential of obtaining very simple proof objects, recording only the first phase, for focused proofs; moreover, because of the bidirectional link, we can reconstruct focused sequent proofs from such proof objects.

Both the search and the representational aspects of focused calculus of structures proofs are being investigated in the Profound tool (see section 5.1).

6.2. Nested Sequents for Intuitionistic Modal Logics

Participant: Lutz Straßburger.

We show how the recent results for treating classical modal logics in the modal cube under S5 via nested sequents [32], [33] can be carried to intuitionistic modal logics. Thus, we present cut-free nested sequent systems for all intuitionistic modal logics in the modal cube up to IS5, and we show how this can be done in a modular way, i.e., to each of the axioms d, t, b, 4, and 5, we assign a set of inference rules, such that for each subset of d,t,b,4,5 the corresponding set of rules is sound and complete for the defined logic. This work has been presented in an invited talk at the IMLA workshop in Nancy [25].

6.3. The logic of nominal abstraction

Participant: Dale Miller.
In the paper [12], Gacek, Miller, and Nadathur have developed a strong logic that allows strong forms of induction and co-induction in the presence of the $\nabla$ quantifier. This quantifier was introduced in earlier work by Miller and Tiu [50] where it was shown to provide declarative and flexible operational semantic specifications for a number of systems such as the $\lambda$-calculus and the $\pi$-calculus. The paper [12] introduces a generalization form of equality, called nominal abstraction, that permits natural specifications of predicates such as freshness. This paper provides the necessary meta theory (cut-elimination) for this new logic.

6.4. A framework for focusing and cut-elimination  
**Participant:** Dale Miller.

Liang and Miller provide in [14] a general setting for specifying proofs in intuitionistic and classical logic. In this setting, it is possible to guarantee cut-elimination and initial elimination results simply by checking how certain simple side-conditions are used within a focused proof system. This setting treats both intuitionistic and classical logics as well allowing certain hybridizations of these two logics. This work helped lead the authors to finding a way to truly mix in one logic and one proof system both classical and intuitionistic logic, as described in [22].

6.5. Synthetic connectives and their proof system  
**Participant:** Dale Miller.

In recent years, focused proof systems have being used to expand our understanding of how introduction rules and structural rule relate to each other. In these proof systems, inference rules and logical connectives are polarized as negative or positive in such a way that the invertible inference rules all belong to the negative polarity. Groups of negative connectives can then be grouped into one negative synthetic connective: similarly, positive connectives can be grouped into a positive synthetic connective. Such synthetic connectives admit cut-elimination. Remarkably, focused proof systems for classical and intuitionistic logics can be organized so that negative formulas are, in fact, treated linearly. That is, if weakening or contraction is applied to a formula, that formula is positive.

Focused proof systems can be used to design richly varying collections of synthetic connectives. These proof systems also provide for new means of describing parallelism within proofs and mixing computation and deduction. The ability to treat negative formulas linearly provides important information for the design of automated theorem provers. Synthetic connectives and their associated inference rules will also allow for the design of broad spectrum proof certificates that theorem provers will be able to print and simple proof checkers will be able to validate. Miller’s conference paper [20] develops this approach to proof certificates.

6.6. Automated reasoning and SMT solving  
**Participants:** Mahfuza Farooque, Stéphane Lengrand.

Automated reasoning uses a broad range of techniques whose soundness and completeness relate to the existence of proofs. The research programme of the ANR PSI project at PariCal is to build a finer-grained connection by specifying automated reasoning techniques as the step-by-step construction of proofs, as we know it from proof theory and logic programming. The goal is to do this in a unifying framework, namely proof-search in a classical polarized sequent calculus. One of the advantages of this is to combine those techniques more easily. Another one is to envisage extending those techniques.

For instance, the algorithm at the heart of SMT-solving (SAT-modulo-Theory) is DPLL(T), whose theory does not treat existential variables (although SMT-provers often implement incomplete ad hoc techniques for them). We have first encoded DPLL(T) as the step-by-step construction of a proof tree in a classical polarized sequent calculus extended with calls to a decision procedure for $T$ (to be published). This proof-theoretic view now allows us to envisage how to extend the algorithm with existential variables.
Another range of techniques that we are addressing is the handling of equality (superposition / paramodulation calculi).

### 6.7. Towards a Stochastic Linear Logic for Biological Computation

**Participants:** Kaustuv Chaudhuri, Joëlle Despeyroux.

In previous work [35], Joëlle Despeyroux and Kaustuv Chaudhuri have given an encoding of the synchronous stochastic $\pi$-calculus in a hybrid extension of intuitionistic linear logic (called HyLL). Precisely, they have shown that focused partial sequent derivations in the encoding are in bijection with stochastic traces. The modal worlds are used to represent the rates of stochastic interactions, and the connectives of hybrid logic are used to represent the constraints in the stochastic transition rules.

One of the most successful applications of the stochastic $\pi$-calculus has been in representing signal transduction networks in cellular biology. An interesting application of this work would therefore be the direct representations of biological processes in HyLL, the original motivation for this line of investigation.

This year, we have worked on specifying some simple examples of regulatory gene networks, together with basic properties of them—such as stability or oscillation—in HyLL. This is ongoing work in collaboration with Gilles Bernot’s team at Nice-Sophia university.
6. New Results

6.1. Proof-theoretical investigations


6.1.1. Sequent calculus and Computational duality

Thunks and duality. Guillaume Munch-Maccagnoni investigated a notion dual to the thunks of call-by-value lambda-calculus which allows to simulate call-by-value in call-by-name. He started to investigate how this structure arises in many models of call-by-name lambda-calculus, how it could explain various phenomena such as storage operators, and how it could relate to features of actual programming languages.

Categorical semantics. Guillaume Munch-Maccagnoni started a collaboration with Marcelo Fiore with the aim of understanding structures behind sequent calculus (of system L in particular) and polarisation (as found in Girard’s classical logic). The goal is to draw links with algebraic and/or computing structures in the unifying language of category theory. This work is also in collaboration with Pierre-Louis Curien.

Noam Zeilberger has continued to work with Paul-André Melliès (PPS) on developing a categorical framework for better understanding contexts and inference rules in proof theory and type theory, with the aim of achieving an integration with the theory of side-effects in programming languages. He presented some results from this work in March at the European Workshop on Computational Effects.

Polarised Peano arithmetic. Guillaume Munch-Maccagnoni extended polarised classical logic and polarised classical realisability to predicate calculus and to Peano arithmetic. This decomposes and simplifies technical artefacts found in call-by-name classical realisability, and sheds a new light on witness extraction from proofs of $\Sigma$ formulae.

(Co)Inductive Types in Sequent Calculus. Hugo Herbelin and Jeffrey Sarnat continued their work towards a sequent calculus presentation of a simply-typed fragment of CIC that has inductive and coinductive types, as formalized using recursion operators and a guard condition. Some progress was made on the formalization of the guard condition and normalization proof, but both remained unfinished as of the end of Sarnat’s postdoc in June.

Classical call-by-need and the duality of computation. In a collaboration with Zena Ariola (and especially after a 3-week visit by Alexis Saurin in Oregon early 2011), Zena Ariola, Hugo Herbelin and Alexis Saurin have presented the call-by-need strategy in the framework of the duality of computation, that is a sequent calculus approach to call-by-need, and they extended call-by-need from minimal logic to classical logic which allowed to integrate smoothly control operators, resulting in particular in a call-by-need $\lambda\mu$-calculus. Moreover, the duality principles involved in such a framework unveiled a new calculus, dual to the usual call-by-need but which is distinct from call-by-name, call-by-value as well as the usual call-by-need. These results were presented at TLCA 2011.

Pursuing the previous collaboration, the three previous authors, together with Paul Downen and Keiko Nakata, studied abstract machines for this classical call-by-need calculus.

6.1.2. Linear dependent types

Arnaud Spiwack started investigating dependent types variants of linear sequent calculus based on Curien & Herbelin’s $\mu\tilde{\mu}$. The goal is to study what kind of set theory arises from such a linear type theory when following the tradition of intuitionistic type theory of defining a set as a type equipped with a relation (this construction is also known as setoid). Sets defined in this manner in Coq gives rise to a quasi-topos (as proved in Arnaud Spiwack’s PhD thesis) which makes for a reasonable approximation of usual mathematics, however
“linear sets” should be quite different and may support some unorthodox mathematics. To have an appropriate theory of linear sets seems to require a fairly rich linear type theory in particular one that supports so-called strong elimination (the type theoretic equivalent to induction). So far, if extending $\mu\eta$ to dependently typed linear logic has been achieved, strong elimination proves to be harder to coin.

Pierre-Marie Pérot just started his PhD on this same general topic. While linear logic appeared as an operational decomposition of intuitionistic logic, dependent types are conversely an essential enrichment of the latter, as they permit to constructively formalize important parts of mathematics. Even though it is nowadays pervasively thought that we should mix them, actually it does seem that nobody seriously tried hitherto. This subject is quite vast, and both sides may mutually enhance the understanding of one another. From the linear side, linear logic is seriously lacking any satisfying syntax at all; and worse, it seems more prone to solely describe computational behavior instead of truly formalizing mathematics, for its very lack of richer types. From the dependent side, usual dependent type systems are based upon PTS, which, being a plain enrichment of basic lambda-calculus, are intrinsically call-by-name structures. Hence, a linear decomposition inspired from polarization techniques may permit a better analysis of the inner, yet-to-be discovered gears of PTS. One may even hope to include non-intuitionistic effects therein. Furthermore, practical systems used today (Coq, for example) come bundled with additional constructs, such as inductives, whose understanding with respect to models is still highly incomplete. One could expect linear logic to shed a new light upon these issues. This thesis stems from preliminary work of Pierre-Marie Pérot (during his M2 internship) inspired by models of GoI and other closely related results from Girard, which suggest a natural way to integrate dependency into them.

6.1.3. Proving with side-effects

Axiom of dependent choice. Hugo Herbelin showed that classical arithmetic in finite types extended with strong elimination of existential quantification proves the axiom of dependent choice. To get classical logic and choice together without being inconsistent is made possible first by constraining strong elimination of existential quantification to proofs that are essentially intuitionistic and second to turn countable universal quantification into an infinite conjunction of classical proofs evaluated along a call-by-need evaluation strategy so as to extract of them intuitionistic contents that complies to the intuitionistic constraint put on strong elimination of existential quantification. This work has been presented at the TYPES conference.

Memory assignment, forcing and delimited control. Hugo Herbelin investigated how to extend his work on intuitionistically proving Markov’s principle [35] and the work of Danko Ilik on intuitionistically proving the double negation shift (i.e. $\forall x \neg \neg A \rightarrow \neg \neg \forall x A$) [38] to other kind of effects. In particular, memory assignment is related to Cohen’s forcing as emphasized by Krivine [40] and as the observation that Cohen’s translation of formula $P$ into $\forall y \leq x \exists z \leq y P(z)$ is similar to a state-passing-style transformation of type $P$ into $S \rightarrow S \times P$.

Hugo Herbelin then designed a logical formalism with memory assignment that allows to prove in direct-style any statement provable using the forcing method, the same way as logic extended with control operators allows to support direct-style classical reasoning. Thanks to the use of delimiters over “small” formulas similar to the notation of $\Sigma^0_1$-formulas in arithmetic, the whole framework remains intuitionistic, in the sense that it satisfies the disjunction and existence property.

Two typical applications of proving with side-effects are global-memory proofs of the axiom of countable choice and an enumeration-free proof of Gödel’s completeness theorem.

The main ideas of this research program have been presented at the Geocal-Lac meeting of the GDR IM.

6.1.4. Delimited control

Delimited control and infinitary/stream calculi

During his summer internship, Paul Downen studied with Alexis Saurin some infinitary $\lambda$-calculi and infinitary rewriting and in particular a proposal by Ketema, Blom, Aoto and Simonsen which allows to consider transfinitely deep terms. The proposed calculus presented several defects on which Paul Downen studies focused. Some of these defects were corrected but the work is still on-going.
In a collaboration with Marco Gaboardi and Koji Nakazawa, Alexis Saurin has been studying how to turn the \(\Lambda\mu\)-calculus into a truly stream-based calculus. This involved enlarging the syntactical category for streams, defining a type system and comparing with other proposals for computing on streams.

Alexis Saurin also developed previous results on \(\Lambda\mu\)-calculus in a paper currently under final revision for publication in TCS.

**PTS and delimited control** Following Danvy-Filinski’ simply-typed system for a \(\lambda\)-calculus with delimited control, Hugo Herbelin and Pierre Boutillier have defined a set of rules for pure type systems with control operator. The work relies on the CPS used to encode them in standard Pure Type Systems and involves extra type annotations. Nevertheless, it seems to be more general than previous attempts to build classical PTS [20]. It has been presented at the workshop TPDC (Theory and Practice of Delimited Continuations) in Novi Sad and an article is under preparation.

### 6.1.5. Interactive Realizability

Thanks to the Curry-Howard correspondence for classical logic, it is possible to extract programs from classical proofs. These programs use control operators as a way to implement backtracking and processes of intelligent learning by trial and error. Unfortunately, such programs are often hard to write, difficult to understand and are very inefficient: every time a program backtracks, it forgets way too much information. This state of thing is due to a poor understanding and control of the backtracking mechanism that interprets classical proofs. In order to write down more efficient programs, it is necessary to describe exactly: a) what the programs learn, b) how the knowledge of programs varies during the execution.

A first step towards this goal is the theory of Interactive Realizability, a semantics for intuitionistic Arithmetic with excluded middle over semi-decidable predicates. It is based on a notion of state, which describes the knowledge of programs coming from a classical proof, and explains how the knowledge evolves during computation.

Federico Aschieri is working in two directions. First he is extending this realizability semantics to second-order intuitionistic Arithmetic with the same excluded middle over semi-decidable predicates. He has also discovered a new state passing style transformation, which allows to implement in System F efficient programs, which backtrack at the right point and do not forget anything when backtracking. He is also investigating an interesting relation with the forcing semantics: it seems that his transformation is a very direct, new constructive formulation of forcing.

Secondly, in collaboration with Berardi, he is also extending Interactive Realizability to first-order Arithmetic, with full excluded middle. This work promises to provide a significantly finer description of the learning mechanism that interprets classical proofs.

### 6.1.6. Substitutions and isomorphisms

Pierre-Louis Curien extended his collaboration with Martin Hofmann (Univ. of Munich) and Richard Garner (MacQuarie University, Sydney), started in 2010, to the point where the picture sought and announced in the report of last year turned out to be a bit less idyllic than expected. Let us recall the question addressed. We wanted to compare precisely two ways of giving a categorical interpretations of Martin-Löf type theory, both overcoming the following mismatch: syntax has exact substitutions, while their categorical interpretation, in terms of pullbacks or fibrations, “implements” substitutions only up to isomorphism. One can then either change the model (strictification) [36], or modify the syntax (by introducing explicit substitutions and more importantly explicit coercions between types that are now only isomorphic) [2]. These approaches turn out to be related through adjunctions in a suitable 2-categorical framework that has a conceptual interest of its own. But these adjunction do not fit entirely together, as we found out early 2011: One is base-dependent, and the other not. So we cannot put them directly aside to get the nice conceptual picture we hoped for. Still, our initial goal of expressing the interpretations in terms of on another can be attained, but this remains to be worked out in detail.
6.1.7. Miscellanea
During his three months visit in Beijing, Pierre-Louis Curien worked on the relations between Rewriting theory and the theory of Gröbner bases and other bases like Janet bases or involutive bases that have been introduced in computer algebra. These comparisons shed some light on the classification of various completion techniques for rewriting systems (a completion turns a rewriting system into an equivalent locally confluent one). This is work in progress.

6.2. Metatheory of Coq and beyond
Participants: Pierre Boutillier, Hugo Herbelin, Yann Régis-Gianas, Jeffrey Sarnat, Vincent Siles, Matthieu Sozeau, Noam Zeilberger.
From the work he has done last year on the Coq termination checker, Pierre Boutillier wrote an article to describe formally the exact new Coq implementation of a structural guard condition that handles commutative cuts.

6.2.1. Calculus of inductive constructions and typed equality
The work of Hugo Herbelin and Vincent Siles on the equivalence of Pure Type Systems with typed or untyped equality has been accepted for publication [11].

6.2.2. Proofs of higher-order programs
Jeffrey Sarnat and Noam Zeilberger continued to investigate the classical program transformations of continuation-passing-style translation and defunctionalisation [46], from the point-of-view of their effect on the termination proofs of higher-order programs. The practical aim of these investigations is to develop a more systematic understanding of termination proofs, which eventually could result in a compiler from proof assistants with higher-order reasoning (such as Coq) to ones with only first-order reasoning (such as Twelf).

6.2.3. Unification
Matthieu Sozeau and Hugo Herbelin worked on improving the unification algorithm of Coq, making it more predictable and resolving important soundness issues (e.g. type-checking, dealing with universes). In collaboration with Beta Ziliani at (PhD student at MPI Sarbrucken) and Aleksandar Nanevski (Researcher at IMDEA Madrid), Matthieu Sozeau started a project to formalize (on paper) the improved unification algorithm, taking into account advanced features such as canonical structures and type classes. This will give a detailed view of the system to power users like [33] and improve the system to handle the delicate usage patterns developed in the Mathematical Components team at MSR to scale Coq to large formalizations.

6.3. Coq as a functional programming language
Participants: Stéphane Glondu, Pierre Letouzey, Matthias Puech, Matthieu Sozeau.

6.3.1. Type-classes and libraries
Matthieu Sozeau has worked with members of the Foundations team at Nijmegen on enhancing the implementation of type-classes to suit the needs of the development of the MathClasses formalization of abstract algebra (part of the ForMath EU project). This gave rise to an experimental implementation of forward reasoning for instance resolution. Concurrently, he started a collaboration with Jael Kriener (Univ. of Canterbury, who visited for a week in November) to adapt techniques from logic programming to the system, including a determinacy analysis that would give better control of the system when building large hierarchies of structures (as in the MathClasses library).

Pierre Castéran from INRIA Bordeaux and Matthieu Sozeau are developing a tutorial on the use of type-classes that will be the basis of an invited tutorial given by M. Sozeau at the JFLA conference in February 2012. It will be published as part of the new version of the Coq’Art book.
6.3.2. Equations

Matthieu Sozeau has continued the development of the Equations package to build and reason on dependently-typed programs. He reworked the internals of the tool for a more efficient implementation and extended it to handle recursion on arbitrary inductive families. He developed a sizable example of the use of the tool for proving the metatheory of the LF system. The plugin will be released with the upcoming 8.4 version of Coq. A journal article presenting the tool and its usage is in preparation.

6.3.3. Recursive definitions in proof assistants

Together with Ana Bove and Alexander Krauss, Matthieu Sozeau has written a survey article on tools and methods to build and reason on recursive functions in proof assistants based on type theory. The survey compares the relative strengths and weaknesses of various formalization methods available in constructive type theories like Agda or Coq and classical systems like HOL. The article is still under review.

6.3.4. Dependent pattern-matching

Hugo Herbelin and Pierre Boutillier worked on a new simulation of Agda’s style dependent pattern-matching [24] in the Calculus of Inductive Constructions which, on the contrary of Goguen et al’s simulation [32], does not rely on any explicit notion of dependent equality. The simulation relies on a systematic re-generalization of a given matching over the next pattern to matching and over a notion of matching narrowed by type constraints formalized by Pierre Boutillier [14]. The former itself relies on Pierre Boutillier’s master thesis on making the guard condition traversing blocked commutative cuts. The simulation has been implemented in Coq 8.4 but without narrowing on subpatterns yet.

6.3.5. Modularized arithmetical libraries

Pierre Letouzey has integrated in Coq a deep reform of some parts of its Standard Library, mainly the Numbers library of generic / efficient arithmetic. This reform is part of the the version Coq 8.4. The idea is to take advantage of recent improvements of the Coq system in terms of modularity (Type Classes by Sozeau and better Modules by Soubiran) for providing more uniformity in the functions and properties about integers provided in the Standard Library. We now have a base of functions and lemmas which is statically guaranteed to be coherent from one numerical datatype representation to the other, allowing the user to choose more easily between these representations, according to the user’s need in term of simplicity or efficiency. These modernized libraries are also one of the first large-scale experimentations with many recent features of modules and type-classes, and have helped maturing them and established new usage guidelines. The transition from previous versions of these libraries should be relatively transparent thanks to a compatibility layer.

6.3.6. Certified extraction

Stéphane Glondu continued his work on extraction in the Coq-in-Coq formalisation. He proved that reductions in the source language can be simulated in the target language (the converse had been proven two years before). This proof highlighted a critical bug in the actual implementation of Coq that has been solved by Pierre Letouzey.

The formalised extraction is only a step of the actual implementation of Coq: replacing logical subterms by an inert constant. Stéphane Glondu considers this work done, even though some parts have been admitted. He worked on how the formalisation could be redesigned in order to avoid currently admitted lemmas and to allow a better integration with the existing Coq system.

6.3.7. Incrementality in proof languages

Matthias Puech and Yann Régis-Gianas worked on incremental type checking. This preliminary work will be presented during a contributed talk at TLDI 2012 [15]. It sets the grounds of an incremental proof development and checking system, by means of a representation language for repositories of proofs and proof changes.
The traditional interaction with a proof-checker is a batch process. Coq (among others) refines this model by providing a read-eval print loop with a global undo system, implemented in an ad-hoc way. A more general approach to incrementality is being developed by means of a finer-grained analysis of dependencies. The approach developed is adaptable to any typed formal language: the language is specified in a meta-language close to the Edinburgh Logical Framework, in which subsequent versions of a development can be type-checked incrementally. Applications of this framework are: proof language for proof assistants, integrated development environments for proof or programming languages, typed version control systems.

6.3.8. Proofs of programs in Coq

As part of the CerCo european project, in collaboration with Roberto Amadio (PPS, Paris Diderot University), Nicolas Ayache and Yann Régis-Gianas maintained a prototype compiler for a large subset of the C language whose specificity is to annotate input source programs with information about the worst-case execution cost of their machine code. Yann Régis-Gianas started a mechanized version of the proof technique used to prove the correctness of such an annotating compiler.

Yann Régis-Gianas developed another compiler for Core ML that uses a generalization of CerCo technique in order to obtain certified worst case execution time bounds on functional programs. This compiler produces proof obligations in Coq. The corresponding paper will be published in January 2012 in the proceedings of FOPARA 2011. It is available as a technical report [16].
6. New Results

6.1. Dependable Distributed Real-time Embedded Systems

Participants: Pascal Fradet, Alain Girault [contact person], Emil Dumitrescu.

6.1.1. The TSH multi-criteria scheduling heuristic

For autonomous critical real-time embedded systems (e.g., satellite), guaranteeing a very high level of reliability is as important as keeping the power consumption as low as possible. We have designed an off-line scheduling heuristics which, from a given software application graph and a given multiprocessor architecture (homogeneous and fully connected), produces a static multiprocessor schedule that optimizes three criteria: its length (crucial for real-time systems), its reliability (crucial for dependable systems), and its power consumption (crucial for autonomous systems). Our tricriteria scheduling heuristics, TSH, uses the active replication of the operations and the data-dependencies to increase the reliability, and uses dynamic voltage and frequency scaling to lower the power consumption [17]. By running TSH on a single problem instance, we are able to provide the Pareto front for this instance in 3D, therefore exposing the user to several tradeoffs between the power consumption, the reliability and the execution time. Thanks to extensive simulation results, we have shown how TSH behaves in practice. Firstly, we have compared TSH versus an optimal Mixed Linear Integer Program on small instances; the experimental results show that TSH behaves very well compared to the ILP. Secondly, we have compared TSH versus the ECS heuristic (Energy-Conscious Scheduling [84]); the experimental results show that TSH performs systematically better than ECS.

This is a joint work with Ismail Assayad (U. Casablanca, Morocco) and Hamoudi Kalla (U. Batna, Algeria), who both visit the team regularly.

6.1.2. Automating the Addition of Fault Tolerance with Discrete Controller Synthesis

In collaboration with Emil Dumitrescu (INSA Lyon), Hervé Marchand (VERTECS team from Rennes), and Eric Rutten (SARDES team from Grenoble), we have defined a complete framework for the automatic design of fault tolerant embedded systems, based on discrete controller synthesis (DCS) [88]. Its interest lies in the ability to obtain automatically systems satisfying by construction formal properties specified a priori. Our aim is to demonstrate the feasibility of this approach for fault tolerance. We start with a fault intolerant program, modeled as the synchronous parallel composition of finite labeled transition systems. We specify formally a fault hypothesis, state fault tolerance requirements and use DCS to obtain automatically a program having the same behavior as the initial fault intolerant one in the absence of faults, and satisfying the fault tolerance requirements under the fault hypothesis. Our original contribution resides in the demonstration that DCS can be elegantly used to design fault tolerant systems, with guarantees on key properties of the obtained system, such as the fault tolerance level, the satisfaction of quantitative constraints, and so on. We have shown with numerous examples taken from case studies that our method can address different kinds of failures (crash, value, or Byzantine) affecting different kinds of hardware components (processors, communication links, actuators, or sensors). Besides, we have shown that our method also offers an optimality criterion very useful to synthesize fault tolerant systems compliant to the constraints of embedded systems, like power consumption or execution times. In summary, our framework for fault tolerance has the following advantages [67]:

- The automation, because DCS produces automatically a fault tolerant system from an initial fault intolerant one.
- The separation of concerns, because the fault intolerant system can be designed independently from the fault tolerance requirements.
- The flexibility, because, once the system is entirely modeled, it is easy to try several fault hypotheses, several environment models, several fault tolerance goals, several degraded modes, and so on.
• The safety, because, in case of positive result obtained by DCS, the specified fault tolerance properties are guaranteed by construction on the controlled system.

• The optimality when optimal synthesis is used, modulo the potential numerical equalities (hence a non strict optimality). We consider weights cumulated along bounded-length paths. We have adapted our models in order to take into account the additive costs of, e.g., execution time or power consumption, and adapting synthesis algorithms in order to support the association of costs with transitions, and the handling of these new cost functions in the optimal synthesis [59].

We therefore combine, on the one hand, guarantees on the safety of the execution by tolerating faults, and on the other hand, guarantees on the worst cumulated consumption of the resulting dynamically reconfiguring fault tolerant system. Recently, we have incorporated multi-criteria optimization results in this work, to take into account several weight functions: for instance the execution costs of several tasks, the execution of which must be controlled thanks to DCS. We therefore propose several synthesis algorithms, to aggregate the costs into a single cost function, to hierarchize the costs (e.g., to reflect the priorities of the tasks), or to compute the Pareto front of non-dominated solutions.

6.2. Controller Synthesis for the Safe Design of Embedded Systems

Participants: Gwenaël Delaval [contact person], Gregor Goessler, Sebti Mouelhi.

6.2.1. Synthesis of Switching Controllers using Approximately Bisimilar Multiscale Abstractions

The use of discrete abstractions for continuous dynamics has become standard in hybrid systems design (see e.g. [92] and the references therein). The main advantage of this approach is that it offers the possibility to leverage controller synthesis techniques developed in the areas of supervisory control of discrete-event systems [88]. The first attempts to compute discrete abstractions for hybrid systems were based on traditional systems behavioral relationships such as simulation or bisimulation, initially proposed for discrete systems most notably in the area of formal methods. These notions require inclusion or equivalence of observed behaviors which is often too restrictive when dealing with systems observed over metric spaces. For such systems, a more natural abstraction requirement is to ask for closeness of observed behaviors. This leads to the notions of approximate simulation and bisimulation introduced in [63].

These notions enabled the computation of approximately equivalent discrete abstractions for several classes of dynamical systems, including nonlinear control systems with or without disturbances, and switched systems. These approaches are based on sampling of time and space where the sampling parameters must satisfy some relation in order to obtain abstractions of a prescribed precision. In particular, the smaller the time sampling parameter, the finer the lattice used for approximating the state-space; this may result in abstractions with a very large number of states when the sampling period is small. However, there are a number of applications where sampling has to be fast; though this is generally necessary only on a small part of the state-space.

In [22] we have presented a novel class of multiscale discrete abstractions for incrementally stable switched systems that allows us to deal with fast switching while keeping the number of states in the abstraction at a reasonable level. We assume that the controller of the switched system has to decide the control input and the time period during which it will be applied before the controller executes again. In this context, it is natural to consider abstractions where transitions have various durations. For transitions of longer duration, it is sufficient to consider abstract states on a coarse lattice. For transitions of shorter duration, it becomes necessary to use finer lattices. These finer lattices are effectively used only on a restricted area of the state-space where the fast switching occurs.

These abstractions allow us to use multiscale iterative approaches for controller synthesis as follows. An initial controller is synthesized based on the dynamics of the abstraction at the coarsest scale where only transitions of longer duration are enabled. An analysis of this initial controller allows us to identify regions of the state-space where transitions of shorter duration may be useful (e.g. to improve the performance of the controller). Then, the controller is refined by enabling transitions of shorter duration in the identified regions. The last two steps can be repeated until we are satisfied with the obtained controller.
In [21] we propose a technique for the synthesis of safety controllers for switched systems using multi-scale abstractions. We present a synthesis algorithm that exploits the specificities of multi-scale abstractions. The finest scales of the abstraction are effectively explored only when fast switching is needed, that is when the system approaches the unsafe set. We provide experimental results that show drastic improvements of the complexity of controller synthesis using multi-scale abstractions instead of uniform abstractions.

6.2.2. Modular Discrete Controller Synthesis

Discrete controller synthesis (DCS) [88] allows to design programs in a mixed imperative/declarative way. From a program with some freedom degrees left by the programmer (e.g., free controllable variables), and a temporal property to enforce which is not a priori verified by the initial program, DCS tools compute off-line automatically a controller which will constrain the program (by e.g., giving values to controllable variables) such that, whatever the values of inputs from the environment, the controlled program satisfies the temporal property.

Our motivation w.r.t. DCS concerns its modular application, improving the scalability of the technique by using contract enforcement and abstraction of components. Moreover, our aim is to integrate DCS into a compilation chain, and thereby improve its usability by programmers, not experts in discrete control. This work has been implemented into the HEPTAGON/BZR language and compiler [57]. This work is done in collaboration with Hervé Marchand (VERTECS team from Rennes) and Éric Rutten (SARDES team from Grenoble).

The implemented tool allows the generation of the synthesized controller under the form of an HEPTAGON node, which can in turn be analyzed and compiled, together with the HEPTAGON source from which it has been generated. This full integration allows this method to aim different target languages (currently C, JAVA or VHDL), and its integrated use in different contexts.

A formal semantics of BZR has been defined, taking into account its underlying nondeterminism related to the presence of controllable variables.

This language has been used in different contexts. In [15], BZR is used for the generation of discrete handlers of real-time continuous control tasks, in the framework of the ORCCAD tool. BZR has also been used in a case-study of a Fractal designed HTTP server [19]. The purpose of the synthesized controller is to control the automatic reconfigurations of the system (e.g., start of new components, migrations of some components from one computing element to another), in order to preserve some properties (either functional, e.g., exclusivity of activities of two components, or non-functional, e.g., bounded overall load of the system).

6.3. Automatic Distribution of Synchronous Programs

Participants: Gwenaël Delaval [contact person], Alain Girault, Gregor Goessler, Xavier Nicollin, Gideon Smeding.

6.3.1. Modular Distribution

Synchronous programming languages describe functionally centralized systems, where every value, input, output, or function is always directly available for every operation. However, most embedded systems are nowadays composed of several computing resources. The aim of this work is to provide a language-oriented solution to describe functionally distributed reactive systems. This research is conducted within the INRIA large scale action SYNCHRONE and is a joint work with Marc Pouzet (ENS, PARKAS team from Rocquencourt) and Xavier Nicollin (Grenoble INP, VERIMAG lab).

We are working on type systems to formalize, in a uniform way, both the clock calculus and the location calculus of a synchronous data-flow programming language (the HEPTAGON language, inspired from LUCID SYNCHRONE [45]). On one hand, the clock calculus infers the clock of each variable in the program and checks the clock consistency: e.g., a time-homogeneous function, like +, should be applied to variables with identical clocks. On the other hand, the location calculus infers the spatial distribution of computations and checks the spatial consistency: e.g., a centralized operator, like +, should be applied to variables located at

---

31 Open Robot Controller Computer-Aided Design
the same location. Compared to the PhD of Gwenaël Delaval [55], [56], the goal is to achieve modular distribution. By modular, we mean that we want to compile each function of the program into a single function capable of running on any computing location. We make use of our uniform type system to express the computing locations as first-class abstract types, exactly like clocks, which allows us to compile a typed variable (typed by both the clock and the location calculi) into if ... then ... else ... structures, whose conditions will be valuations of the clock and location variables.

We currently work on an example of software-defined radio. We have shown on this example how to use a modified clock calculus to describe the localisation of values as clocks, and the architecture as clocks (for the computing resources) and their relations (for communication links).

### 6.3.2. Distribution of Synchronous Programs under Real-Time Constraints

With the objective to distribute synchronous data-flow programs (e.g. LUSTRE) over GALS architectures, preserving only explicitly specified properties, we have developed a quantitative clock calculus to (1) describe timing properties of the architecture’s clock domain, and (2) describe the properties of the synchronous program to be preserved. The clock calculus is inspired by the network calculus [83], with the difference that clocks are described only with respect to one-another, not with respect to real-time.

As a first result, we have applied our clock calculus to analyze the properties of periodic synchronous data-flow programs executed on a network of processors. Because our clock calculus is relational, it can model and preserve correlated variations of streams. In particular, the common case of a data-flow system that splits a stream for separate treatment, and joins them afterwards, this analysis yields more precise result than comparable methods.

We aim to extend the analysis to account for shared resources and synchronization protocols, so as to distribute synchronous programs preserving specified properties.

### 6.4. New Programming Languages for Embedded Systems

**Participants:** Alain Girault [contact person], Pascal Fradet, Petro Poplavko, Vagelis Bebelis, Bertrand Jeannet, Peter Schrammel.

#### 6.4.1. The DSystemJ programming language

In collaboration with Avinash Malik (IBM Watson) and Zoran Salcic (University of Auckland), we have designed the SYSTEMJ programming language [9], which implements the Globally Asynchronous Locally Synchronous (GALS) Model of Computation (MoC) over JAVA. In a nutshell, SYSTEMJ uses the notion of clock domains (CD) to design portions of the system that must operate at unrelated clocks. CDs communicate with each other via asynchronous rendez-vous. Then, a CD consists of one or several reactions, which react synchronously in lock-step and communicate with each other via synchronous broadcast of signals. Finally, all the data computations are implemented in JAVA.

We have further extended SYSTEMJ to allow programmers to design dynamic GALS systems: this is the new language DSYSTEMJ [27], [12], aimed at dynamic distributed systems that use socket based communication protocols for communicating between components. DSYSTEMJ allows the creation and control at runtime of CDs, their mobility on a distributed execution platform, as well as the runtime reconfiguration of the system’s functionality and topology. We have defined the formal semantics of DSYSTEMJ, based on the Dynamic GALS MoC: it offers very safe mechanisms for implementation of distributed systems, as well as potential for their formal verification. The runtime support is implemented in the SYSTEMJ language, which can as such be considered as a static subset of DSYSTEMJ.

This work has been done within the AFMES associated team with the Electric and Computer Engineering Department of the University of Auckland.
6.4.2. The PRET-C programming language for time-predictable systems

Typical safety critical embedded applications, ranging from complex aircraft flight controllers to embedded health devices require worst case guarantees on their timing behavior. The problem is that general-purpose processors, being highly speculative, are intrinsically non-deterministic, and thus are not ideally suited for implementing such systems: either the computed worst-case execution time is highly pessimistic, or heroic efforts are required to accurately model the caches, pipeline, and speculative execution [93]. For similar reasons, using an RTOS to guarantee the determinism of a program’s behavior, along with temporal guarantees, is not feasible. The ability to analyze temporal bounds is dependent on the selected programming language, compiler tool chain, operating system, and the target hardware.

To alleviate these problems, we have defined a synchronous variant of C called PRET-C, together with Sidharta Andalam and Partha Roop (University of Auckland). PRET-C offers constructs for reactive inputs/outputs; it supports a notion of logical time, synchronous concurrency, and preemption [40]. We have also designed the ARPRET architecture for efficient and predictable execution of PRET-C. ARPRET inherits from the long lasting research effort on reactive processors conducted at the University of Auckland. Finally, all timing constraints are precisely verified using a Worst Case Reaction Time (WCRT) analyzer. While there has been a considerable body of work on the timing analysis of procedural programs [93], such analysis for synchronous programs has received less attention. Current state-of-the-art analyses for synchronous programs use integer linear programming (ILP) combined with path pruning techniques to achieve tight results. These approaches first convert a concurrent synchronous program into a sequential program. ILP constraints are then derived from this sequential program to compute the longest tick length. For PRET-C, we have proposed an alternative approach based on model checking [16]. Unlike conventional programs, synchronous programs are concurrent and state-space oriented, making them ideal for model checking based analysis. Our analysis of the abstracted state-space of the program is combined with expressive data-flow information, to facilitate effective path pruning. We have demonstrated through extensive experimentation that the proposed approach is both scalable and about 67% tighter compared to the existing approaches (namely Protothreads [60] and SC [94]).

This overall framework provides an ideal platform for designing and verifying precision timed real-time systems. It has been conducted within the AFMES associated team with the Electric and Computer Engineering Department of the University of Auckland, and is the topic of the PhD of Sidharta Andalam.

6.4.3. Analysis and Scheduling of Parametric Data-Flow Models

Recent data-flow programming environments support applications whose behavior is characterized by dynamic variations in resource requirements. The high expressive power of the underlying models (e.g., Kahn Process Networks, the CAL actor language) makes it challenging to ensure predictable behavior. In particular, checking liveness (i.e., no part of the system will deadlock) and boundedness (i.e., the system can be executed in finite memory) is known to be hard or even undecidable for such models. This situation is troublesome for the design of high-quality embedded systems.

We have introduced the schedulable parametric data-flow (SPDF) model of computation (MoC) for dynamic streaming applications [23], [32], [36], [34], [35]. SPDF extends the standard data flow model by allowing rates to be parametric (e.g., of the form $2xy$). SPDF was designed to be statically analyzable while retaining sufficient expressive power. We formulated sufficient and general static criteria for boundedness and liveness. In SPDF, parameters can be changed dynamically even within iterations. The safety of dynamic parameter changes can be checked and their implementation made explicit in the graph. These different analyses are made possible using well-defined static operations on symbolic expressions. The same holds for quasi-static scheduling which is the first step towards code generation for multi-core systems.

We are now considering other kinds of analyses for this new data-flow MoC. The objective of these analyses is to generate distributed schedules optimizing both the power consumption and the execution time of applications. The targeted hardware is P2012, a new embedded many-core platform designed by STMicroelectronics consisting of several clusters (9 in the current implementation) interconnected through a 2D mesh asynchronous NoC. Each cluster comprises 16 identical computing cores and is equipped with
a hardware mechanism for DVFS (dynamic voltage and frequency scaling). As a first step, we have studied energy efficient scheduling of simple data-flow graphs for that platform [81]. The next step is to extend the approach to SPDF.

This line of research will be followed in the PhD thesis of Vagelis Bebelis which has just started. It will be conducted in collaboration with STMicroelectronics.

6.4.4. Translating Hybrid Data-Flow Languages to Hybrid Automata

Hybrid systems are used to model embedded computing systems interacting with their physical environment. There is a conceptual mismatch between high-level hybrid system languages like SIMULINK, which are used for simulation, and hybrid automata, the most suitable representation for safety verification. Indeed, in simulation languages the interaction between discrete and continuous execution steps is specified using the concept of zero-crossings, whereas hybrid automata exploit the notion of staying conditions.

In the context of the INRIA large scale action SYNCHRONICS (see § 8.1.4), we studied how to translate the ZELUS hybrid data-flow language [43] developed in this project into logico-numerical hybrid automata by carefully pointing out this issue. We investigated various zero-crossing semantics, proposed a sound translation, and discussed to which extent the original semantics is preserved. This work has been accepted to the conference HSCC’2012 (Hybrid Systems: Computation and Control).

This work is part of the PhD thesis of Peter Schrammel.

6.5. Static Analysis and Abstract Interpretation

Participants: Alain Girault, Bertrand Jeannet [contact person], Lies Lakhdar-Chaouch, Peter Schrammel, Pascal Sotin.

6.5.1. Numerical and logico-numerical abstract acceleration

Acceleration methods are used for computing precisely the effects of loops in the reachability analysis of counter machine models. Applying these methods to synchronous data-flow programs with Boolean and numerical variables, e.g., LUSTRE programs, firstly requires the enumeration of the Boolean states in order to obtain a control graph with numerical variables only. Secondly, acceleration methods have to deal with the non-determinism introduced by numerical input variables.

Concerning the latter problem, we pushed further the work presented in [90] that extended the concept of abstract acceleration of Gonnord et al. [69], [68] to numerical input variables, and we wrote a journal version [13]. The original contributions of [13] compared to [91] is abstract backward acceleration (for backward analysis) and a detailed comparison of the abstract acceleration approach with the derivative closure approach of [39], which is related to methods based on transitive closures of relations.

We then worked more on the first point, which is to apply acceleration techniques to data-flow programs without resorting to an exhaustive enumeration of Boolean states. To this end, we introduced (1) logico-numerical abstract acceleration methods for CFGs with Boolean and numerical variables and (2) partitioning techniques that make logical-numerical abstract acceleration effective. Experimental results showed that incorporating these methods in a verification tool based on abstract interpretation provides not only significant advantage in terms of accuracy, but also a gain in performance in comparison to standard techniques. This work was published in [28].

This line of work is part of the PhD thesis of Peter Schrammel.

6.5.2. Improving dynamic approximations in static analysis

Abstract interpretation [51] formalizes two kind of approximations that can be done in the static analysis of programs:
Static approximations, defined by the choice of an abstract domain of abstract properties (for instance, intervals or convex polyhedra that approximates set of points in numerical spaces), and the definition of sound approximations in this domain of concrete operations (variable assignments, tests, ...). These abstract properties and operations are substitutes to the concrete properties and operations defined by the semantics of the analyzed program. This stage results into a abstract fixpoint equation \( Y = G(Y) \), \( Y \in A \), where \( A \) is the abstract domain. The best (least) solution of this equation can be obtained by Kleene iteration, which consists in computing the sequence \( Y_0 = \bot_A, Y_{n+1} = G(Y_n) \), where \( \bot_A \) is the least element of the lattice \( A \).

Dynamic approximations, that makes the Kleene iteration sequence converge in finite time by applying an extrapolation operator called widening and denoted with \( \nabla \). This results in a sequence \( Z_0 = \bot_A, Z_{n+1} = Z_n \nabla G(Z_n) \) that converges to a post-fixpoint \( Z_\infty \supseteq G(Z_\infty) \). For instance, for many numerical abstract domains (like octagons \( [86] \) or convex polyhedra \( [75] \)) the “standard” widening \( \nabla : A \times A \rightarrow A \) consists in keeping in the result \( R = P \nabla Q \) the numerical constraints of \( P \) that are still satisfied by \( Q \).

The problem addressed here is that the extrapolation performed by widening often loses crucial information for the analysis goal.

6.5.2.1. Widening with thresholds.

A classical technique for improving the precision is “widening with thresholds”, which bounds the extrapolation. The idea is to parameterize \( \nabla \) with a finite set \( \mathcal{C} \) of threshold constraints, and to keep in the result \( R = P \nabla\mathcal{C}Q \) those constraints \( c \in \mathcal{C} \) that are still satisfied by \( Q \): \( P \nabla\mathcal{C}Q = (P \nabla Q) \cap \{ c \in \mathcal{C} | Q \models c \} \). In practice, one extrapolates up to some threshold; in the next iteration, either the threshold is still satisfied and the result is better than with the standard widening, or it is violated and one extrapolates up to the remaining thresholds.

The benefit of this refinement strongly depends on the choice of relevant thresholds. In \([33],[26]\) we proposed a semantic-based technique for automatically inferring such thresholds, which applies to any control graph, be it intraprocedural, interprocedural or concurrent, without specific assumptions on the abstract domain. Despite its technical simplicity, we showed that our technique is able to infer the relevant thresholds in many practical cases.

6.5.2.2. Policy Iteration.

Another direction we investigated for solving the fix-point equation \( Y = G(Y) \), \( Y \in A \) is the use of Policy Iteration, which is a method for the exact solving of optimization and game theory problems, formulated as equations on min max affine expressions. In this context, a policy is a strategy for the min-player, which gives rise to a simplified equation \( X = F^\pi(X), F^\pi \geq F, X \in \mathbb{R}^n \) which is easier to solve that the initial equation \( X = F(X), X \in \mathbb{R}^n \). Policy iteration iterates on policies rather than iterating the application of \( F \) (as in Kleene iteration), using the property that the least fixpoint of \( F \) corresponds to the least fixpoint of \( F^\pi \) for some \( \pi \).

\([50]\) showed that the problem of finding the least fixpoint of semantic equations on some abstract domains can be reduced to such equations on min max affine expressions, that can then be solved using Policy Iteration instead of the traditional Kleene iteration with widening described above.

We first investigated the integration of the concept of Policy Iteration in a generic way into existing numerical abstract domains. We implemented it in the APRON library (see module \( 5.4 \)). This allows the applicability of Policy Iteration in static analysis to be considerably extended.

In particular we considered the verification of programs manipulating Boolean and numerical variables, and we provided an efficient method to integrate the concept of policy in the logico-numerical abstract domain \( \text{BDDAPRON} \) that mixes Boolean and numerical properties (see module \( 5.4 \)). This enabled the application of the policy iteration solving method to much more complex programs, that are not purely numerical any more. This work was published in \([30]\).
6.5.3. Analysis of imperative programs

We also studied the analysis of imperative programs. Even if it is preferable to analyze embedded systems
described in higher-level languages such as synchronous languages, it is also useful to be able to analyze C
programs. Moreover, it enables a wider diffusion of the analysis techniques developed in the team.

6.5.3.1. Inferring Effective Types for Static Analysis of C Programs

This work is a step in the project of connecting the C language to our analysis tool INTERPROC/CONCURINTERPROC (see section 5.5.4). The starting point is the connection made by the industrial partner EADS-IW in the context of the ANR project ASOPT (§ 8.1.2) from a subset of the C language to INTERPROC. This translation uses the NEWSPEAK intermediate language promoted by EADS [77].

```c
typedef struct {
    int n;
} t;

int main()
{
    t x; t* y;
    int *p,*q;
    y = alloc(t); p = &(y->n);
    y = &x; q = &(y->n);
    *p = 1; *q = 2; *p = *p < 1;
    return *p;
}
```

```c
typedef enum {
    10=0,11=1,12=2
} e;

typedef struct {
    e n;
} t;

int main()
{
    t x; t* y;
    e *p,*q;
    y = alloc(t); p = &(y->n);
    y = &x; q = &(y->n);
    *p = l1; *q = l2; *p = (*p==l0)?l1:l0;
    return *p;
}
```

Initial program. Transformed program.

Figure 3. Inferring finite types in C programs

The problem addressed here is that the C language does not have a specific Boolean type: Boolean values are encoded with integers. This is also true for enumerated types, that may be freely and silently cast to and from integers. On the other hand, our verification tool INTERPROC that infers the possible values of variables at each program point may benefit from the information that some integer variables are used solely as Boolean or as enumerated type variables, or more generally as finite type variables with a small domain. Indeed, specialized and efficient symbolic representations such as BDDs are used for representing properties on such variables, whereas approximated representations like intervals and octagons are used for larger domain integers and floating-points variables.

Driven by this motivation, we proposed in [25] a static analysis for inferring more precise types for the variables of a C program, corresponding to their effective use. The analysis addresses a subset of the C99 language, including pointers, structures and dynamic allocation. The principle of the method is very different from type inference techniques used in functional programming languages such as ML, where the types are inferred from the context of use. Instead, our analysis can be seen as a simple points-to analysis, followed by a disjunction version of a constant propagation analysis, and terminated by a program transformation that generates a strongly typed program. Fig. 3 illustrates this process. On this example, we discover that the program is a finite-state one, to which exact analysis technique can be applied.

6.5.3.2. Interprocedural analysis with pointers to the stack

This work addressed the problem of interprocedural analysis when side-effect are performed on the stack containing local variables. Indeed, in any language with procedures calls and pointers as parameters (C, Ada) an instruction can modify memory locations anywhere in the call-stack. The presence of such side effects
breaks most generic interprocedural analysis methods, which assume that only the top of the stack may be modified. In [29] we presented a method that addresses this issue, based on the definition of an equivalent local semantics in which writing through pointers has a local effect on the stack. Our second contribution in this context is an adequate representation of summary functions that models the effect of a procedure, not only on the values of its scalar and pointer variables, but also on the values contained in pointed memory locations. Our implementation in the interprocedural analyzer PINTERPROC (see §5.5.4) results in a verification tool that infers relational properties on the value of Boolean, numerical, and pointer variables.

6.6. Component-Based Construction

Participants: Lacramioara Astefanoaei, Alain Girault, Gregor Goessler [contact person], Roopak Sinha, Gideon Smeding.

6.6.1. Incremental converter synthesis

We have proposed and implemented a formal incremental converter-generation algorithm for system-on-chip (SoC) designs. The approach generates a converter, if one exists, to control the interaction between multiple intellectual property (IP) protocols with possible control and data mismatches, and allows pre-converted systems to be re-converted with additional IPs in the future. IP protocols are represented using labeled transition systems (LTS), a simple but elegant abstraction framework which can be extracted from and converted to standard IP description languages such as VHDL. The user can provide control properties, each stated as an LTS with accepting states, to describe desired aspects of the converted system, including fairness and liveness. Furthermore, data specifications can be provided to bound data channels between interacting IPs such that they do not over/under flow. The approach takes into account the uncontrollable environment of a system by allowing users to identify signals exchanged between the SoC and the environment, which the converter can neither suppress nor generate.

Given these inputs, the conversion algorithm first computes the reachable state-space of a maximal non-deterministic converter that ensures (i) the satisfaction of the given data specifications and (ii) the trace equivalence with the given control specifications, using a greatest fix-point computation. It then checks, using the standard algorithm for Büchi games, whether the converter can ensure the satisfaction of the given control specifications (reachability of accepting states) regardless of how the environment behaves. If this is found to be true, deterministic converters can be automatically generated from the maximal non-deterministic converter generated during the first step. The algorithm is proven to be sound and complete, with a polynomial complexity in the state-space sizes of given IP protocols and specifications. It is also shown that it can be used for incremental design of SoCs, where IPs and specifications are added to an SoC in steps. Incremental design allows to constrain the combinatorial explosion of the explored state-space in each step, and also reduces on-chip wire congestion by decentralizing the conversion process.

A Java implementation has been created, and experimental results show that the algorithm can handle complex IP mismatches and specifications in medium to large AMBA based SoC systems. Future work involves creating a library of commonly-encountered specifications in SoC design such as sharing of control signals between interacting IPs using buffers, signal lifespans, and the generation of optimal converters based on quantitative criteria such as minimal power usage.

This work has been done within the AFMES associated team with the Electric and Computer Engineering Department of the University of Auckland.

6.6.2. Causality Analysis in Contract Violation

Establishing liabilities in case of litigation is generally a delicate matter. It becomes even more challenging when IT systems are involved. Generally speaking, a party can be declared liable for a damage if a fault can be attributed to that party and this fault has caused the damage. The two key issues are thus to establish convincing evidence with respect to (1) the occurrence of the fault and (2) the causality relation between the fault and the damage. The first issue concerns the technique used to log the relevant events of the system and to ensure that the logs can be produced (and have some value) in court. The second issue is especially complex when several
faults are detected in the logs and the impact of these faults on the occurrence of the failure has to be assessed. In [6] we have focused on this second issue and proposed a formal framework for reasoning about causality. A system based on this framework could be used to provide relevant information to the expert, the judge, or the parties themselves (in case of amicable settlement) to analyze the origin of the failure of an IT system.

The notion of causality has been studied for a long time in computer science, but with very different perspectives and goals. In the distributed systems community, causality (following Lamport’s seminal paper [82]) is seen essentially as a temporal property. In our context, the temporal ordering contributes to the analysis, but it is obviously not sufficient to establish the logical causality required to rule on a matter of liability: the fact that an event \( e_1 \) has occurred before an event \( e_2 \) does not imply that \( e_1 \) was the cause for \( e_2 \) (or that \( e_2 \) would not have occurred if \( e_1 \) had not occurred).

Our formal model is based on components interacting according to well identified interaction models [5]. Each component is associated with an individual contract which specifies its expected behavior. The system itself is associated with a global contract which is assumed to be implied by the composition of the individual contracts.

In [6] we have defined several variants of logical causality. The first variant, necessary causality, characterizes cases when the global contract would not have been violated if the local contract had been fulfilled. The second variant, sufficient causality, characterizes cases when the global contract would have been violated even if all the other components had fulfilled their contracts. In other words, the violation of its contract by a single component was sufficient to violate the global contract.

We are currently extending to framework to other models of computation and communication, in particular, to timed automata.

### 6.6.3. Realizability of Choreographies for Services Interacting Asynchronously

Choreography specification languages describe from a global point of view interactions among a set of services in a system to be designed. Given a choreography specification, the goal is to obtain a distributed implementation of the choreography as a system of communicating peers. These peers can be given as input (e.g., obtained using discovery techniques) or automatically generated by projection from the choreography. Checking whether some set of peers implements a choreography specification is called realizability. This check is in general undecidable if asynchronous communication is considered, that is, services interact through message buffers.

In [24] we consider conversation protocols as a choreography specification language, and leverage a recent decidability result [54] to check automatically the realizability of these specifications by a set of peers under an asynchronous communication model with a priori unbounded buffers.

### 6.6.4. A Theory of Fault Recovery for Component-Based Models

In [18] we have introduced a theory of fault recovery for component-based models. A model is specified in terms of a set of atomic components that are incrementally composed and synchronized by a set of glue operators. We define what it means for such models to provide a recovery mechanism, so that the model converges to its normal behavior in the presence of faults. We identify corrector (atomic or composite) components whose presence in a model is essential to guarantee recovery after the occurrence of faults. We also formalize component based models that effectively separate recovery from functional concerns.

### 6.7. Aspect-Oriented Programming

**Participants:** Henri-Charles Blondeel, Pascal Fradet [contact person], Alain Girault, Marnes Hoff.

The goal of Aspect-Oriented Programming (AOP) is to isolate aspects (such as security, synchronization, or error handling) which cross-cut the program basic functionality and whose implementation usually yields tangled code. In AOP, such aspects are specified separately and integrated into the program by an automatic transformation process called weaving.
Although this paradigm has great practical potential, it still lacks formalization and undisciplined uses make reasoning on programs very difficult. Our work on AOP addresses these issues by studying foundational issues (semantics, analysis, verification) and by considering domain-specific aspects (availability, fault tolerance or refinement aspects) as formal properties.

### 6.7.1. Aspects Preserving Properties

Aspect Oriented Programming can arbitrarily distort the semantics of programs. In particular, weaving can invalidate crucial safety and liveness properties of the base program.

We have identified categories of aspects that preserve some classes of properties [10]. Our categories of aspects comprise, among others, observers, aborters, and confiners. For example, observers do not modify the base program’s state and control-flow (e.g., persistence, profiling, and debugging aspects). These categories are defined formally based on a language independent abstract semantic framework. The classes of properties are defined as subsets of LTL for deterministic programs and CTL* for non-deterministic ones. We have formally proved that, for any program, the weaving of any aspect in a category preserves any property in the related class.

In a second step, we have designed for each aspect category a specialized aspect language which ensures that any aspect written in that language belongs to the corresponding category. These languages preserve the corresponding classes of properties by construction.

This work was conducted in collaboration with Rémi Douence from the ASCOLAINRIA team at École des Mines de Nantes.

### 6.7.2. Fault Tolerance Aspects

In the recent years, we have studied the implementation of specific fault tolerance techniques in real-time embedded systems using program transformation [1]. We are now investigating the use of fault-tolerance aspects in digital circuits. To this aim, we consider program transformations for hardware description languages (HDL). Our goal is to design an aspect language allowing users to specify and tune a wide range of fault tolerance techniques, while ensuring that the woven HDL program remains synthesizable. The advantage would be to produce fault-tolerant circuits by specifying fault-tolerant strategies separately from the functional specifications.

We have reviewed the different fault tolerant techniques used in integrated circuits: concurrent error detection, error detecting and correcting codes (Hamming, Berger codes, ...), spatial and time redundancy. We have designed a simple hardware description language inspired from Lustre and Lucid Synchrone. It is a core functional language manipulating synchronous boolean streams. Faults are represented by bit flips and we take into account all fault models of the form “at most k faults within n clock signals”. The language semantics as well as the fault model have been formalized in Coq. The next step is to express standard fault tolerance techniques as program transformations and prove that they allow to tolerate all faults of a given model.

### 6.7.3. Refinement Aspects

Chemical programming describes computation in terms of a chemical solution in which molecules (representing data) interact freely according to reaction rules (representing the program). Solutions are represented by multisets of elements and reactions by rewrite rules which consume and produce new elements according to conditions. This paradigm makes it possible to express programs without artificial sequentiality in a very abstract way. It bridges the gap between specification and implementation languages.

A drawback of chemical languages is that their very high-level nature usually leads to very inefficient programs. We have proposed a refinement oriented approach where the basic functionality is expressed as a chemical program whereas efficiency is achieved separately by:

- structuring the multiset with a data type defining neighborhood relations;
- describing the selection of elements according to their neighborhood;
- specifying the evaluation strategy (i.e., the application of rules and termination).
Using these three implementation aspects (data structure, selection and strategy), the chemical program can then be refined automatically into an efficient low-level program. The crucial methodological advantage is that logical issues are decoupled from efficiency issues.

This research, that takes place within the AUTOCHEM project (see Section 8.1.1), is done in collaboration with Jean-Louis Giavitto (Ircam, Paris). It is the subject matter of Marnes Hoff’s PhD thesis.
6. New Results

6.1. Models of Programming

- P. Herms, together with C. Marché and B. Monate (CEA List), developed a certified VC generator, using Coq. The program for VC calculus and its specifications are both written in Coq, but the code is crafted so that it can be extracted automatically into a stand-alone executable. It is also designed in a way that allows the use of arbitrary first-order theorem provers to discharge the generated obligations [37]. This is a first step towards a certified VC generator for C programs annotated with ACSL [54].

- Until now, we only considered the proof of sequential programs. However the rely/guarantee approach give a natural way to extend deductive verification to concurrent programs. During his internship supervised by C. Paulin, N. Gaspar explored this idea. He formalised in Coq the axiomatic semantics of a simple concurrent language using the rely-guarantee approach [43] and proposed a translation of this language into Why programs in order to automatically generate the proof obligations.

- W. Urribarri, together with C. Paulin, proposed an extension of the Why language with modules and functors together with a refinement calculus in order to organise large developments in a structured and abstract way. She built a prototype implementation of this calculus.

- D. Baelde, in cooperation with P. Courtieu (CNAM), D. Gross-Amblard (U. Bourgogne and Rennes), C. Paulin and X. Urbain proposed a formal proof of security for two watermarking algorithms. The proof uses a reduction of an arbitrary attack unmarking the data to the discovery of a secret key. It has been fully formalized in Coq using the ALEA library. This work has been presented at the conference TYPES 2011 and a paper is in preparation.

- Generating multimedia streams, such as in a netradio, is a task which is complex and difficult to adapt to every users’ needs. D. Baelde, in cooperation with R. Beauxis (Tulane University, LA, USA) and S. Mimram (CEA List) introduce a novel approach, based on a dedicated high-level functional programming language, called Liquidsoap, for generating, manipulating and broadcasting multimedia streams [20]. Unlike traditional approaches, which are based on configuration files or static graphical interfaces, it also allows the user to build complex and highly customized systems. This language is based on a model for streams and contains operators and constructions, which make it adapted to the generation of streams. The interpreter of the language also ensures many properties concerning the good execution of the stream generation.

6.2. Proofs of Imperative Programs

- The Why3 reimplementation of the Why platform, started in 2010, was publicly released in 2011 [35]. The main developers are A. Paskevich, J.-C. Filliâtre, F. Bobot, and C. Marché. The language of Why, both programming and annotation parts, was significantly extended: algebraic types, records, pattern matching, recursive logical definitions are now supported. These logical declarations are structured in modules (a.k.a. theories). The module language comes with an original mechanism for reusing theories in specialized contexts using partial instantiations. These new features have been presented at the first international workshop on intermediate verification languages (BOOGIE 2011) [21] and will be presented at VSTTE 2012 [69].
• A. Tafat and C. Marché used the Why3 system to perform a complete proof of the “Binary Heaps” challenge \[41\] from the VACID-0 international collection \[76\]. Solving this challenge is a case study for a general approach of abstract interfaces for C programs, currently under development by A. Tafat, based on initial ideas described together with S. Boulmé which were published in a 2011 special issue \[29\].

• In 2011, we set up a web gallery to publicly expose the programs that we specified and proved. This is the so-called ProVal collection of verified programs, and available at URL http://proval.lri.fr/gallery/index.en.html.

• K. Krishnamani and C. Marché proposed a technique for automatically generating loop invariants in C programs. It is based on the well-known predicate abstraction approach, which is adapted to take into account pre-existing specifications, and to proceed modularly, that is each function of a program is processed independently, with its own sets of predicates. The approach is also extended in order to generate universally quantified invariants \[38\]. The prototype is available as a Frama-C plugin at URL http://proval.lri.fr/agen.

• C. Dross, together with Y. Moy (Adacore) and J.-C. Filliâtre, addressed the problem of verifying programs involving containers. Containers such as lists, vectors, sets or maps are an attractive alternative to ad-hoc data structures based on pointers. C. Dross gave a definition of containers whose aim is to facilitate their use in certified software, using modern proof technology and novel specification languages. Correct usage of containers and user-provided correctness properties can be checked either by execution during testing or by formal proof with an automatic prover. It relies on a formal semantics for containers and an axiomatization of this semantics targeted at automatic provers. C. Dross proved in Coq that the formal semantics is consistent and that the axiomatization thereof is correct. This work was presented at TAP 2011 \[26\].

• Proving that pointer programs preserve data invariants, in a modular way, is known to be a challenge. R. Bardou and C. Marché designed a new approach based on advanced static typing systems (based on memory regions and permissions) to control memory updates and ownership of data \[11\]. A prototype implementation is built, called Capucine, available for download at http://romain.bardou.fr/capucine. To demonstrate the ability of this approach, the challenge “sparse arrays” of the VACID-0 benchmarks \[76\] has been certified \[30\].

• Separation logic has shown to be an elegant way to deal with programs which use data-structures with pointers. However it requires a specific logical language, provers, and specific reasoning techniques. In his PhD, F. Bobot introduced a technique to express ideas from separation logic in the traditional framework of deductive verification \[12\]. He proposed to derive “separation predicates” from user-supplied inductive definitions. These predicates come with suitable axiomatization, including frame rules, expressed in usual first-order logic. This translation takes special care to ensure the best use of automated theorem provers.

6.3. Automated Deduction

• In his thesis \[14\], S. Lescuyer formalized and designed purely reflexive tactics for automated deduction in Coq.

• É. Contejean, together with Pierre Courtieu, Julien Forest, Olivier Pons and Xavier Urbain (Cedric Laboratory, CNAM & ENSIIE) presented the last version of the rewriting toolkit CiME3 at RTA 2011 \[25\]. Amongst other original features, this version enjoys two kinds of engines: to handle and discover proofs of various properties of rewriting systems, and to generate Coq scripts from proof traces given in certification problem format in order to certify them with a skeptical proof assistant like Coq. Thus, these features open the way for using CiME3 to add automation to proofs of termination or confluence in a formal development in the Coq proof assistant.
• In their TACAS paper [24], S. Conchon, É. Contejean and M. Iguernelala present a modular extension of ground AC-completion for deciding formulas in the combination of the theory of equality with user-defined AC symbols, uninterpreted symbols and an arbitrary signature disjoint Shostak theory X.

• F. Bobot and A. Paskevich studied translation from a first-order logic with polymorphic types à la ML (which is the base logic of the Why platform and the Alt-Ergo theorem prover) to a many-sorted or one-sorted logic implemented in mainstream automated theorem provers. They devised a three-stage scheme where the last stage eliminates polymorphic types while adding the necessary “annotations” to preserve soundness, and the first two stages serve to protect certain terms so that they can keep their original types and unannotated form. Such protection allows to make use of provers’ built-in theories and operations. This work generalizes the previous study by S. Lescuyer and J.-F. Couchot [65] onto arbitrary monomorphic types, e.g. array types. It was presented at FroCoS 2011 [22] (see also an extended version with full proofs [42]). These results are part of F. Bobot’s PhD thesis [12].

6.4. Floating-Point and Numerical Programs

• T. Nguyen and C. Marché have worked on how to prove floating-point programs while taking into account architecture- and compiler-dependent features such as the use of the x87 stack in Intel microprocessors. This is done by analyzing the assembly code generated by the compiler [40], [28]

• S. Boldo and C. Marché published a survey article on the proofs of numerical C programs using both automatic provers and Coq [15].

• S. Boldo and T. Nguyen have worked on how to prove numerical programs on multiple architectures and compilers [17]. More precisely, it covers all the compiler choices about the use of extended registers, FMA, and reorganization of additions.

• S. Boldo and J.-M. Muller (CNRS, Arénaire, LIP, ÉNS Lyon) have worked on new floating-point algorithms for computing the exact and approximated errors of the FMA (fused multiply-and-add) [16].

• S. Boldo and G. Melquiond have developed in Coq a comprehensive formalization of floating-point arithmetic: core definitions, axiomatic and computational rounding operations, high-level properties [23]. It provides a framework for developers to formally certify numerical applications.

• G. Melquiond, in collaboration with F. de Dinechin (Arénaire, LIP, ÉNS Lyon) and C. Lauter (Intel Hillsboro), has improved the methodology for formally proving floating-point mathematical functions when their correctness depends on relative errors [19].

• S. Boldo, J.-C. Filiatré and G. Melquiond, in collaboration with F. Clément (Estime, INRIA Paris-Rocquencourt) and M. Mayero (University Paris 13) have finished a full formal proof of a program solving a partial differential equation (the wave equation) using a finite difference scheme [36]. This proof includes both the mathematical convergence proof (method error) [57], a tricky floating-point proof [56] and proofs of the absence of runtime errors.

• C. Lelay, under the supervision of S. Boldo and G. Melquiond, has worked on differentiability in Coq. The goal was to prove the existence of a solution to the wave equation thanks to D’Alembert’s formula and to automatize the process as much as possible [44] [77].

• G. Melquiond, in collaboration with W. G. Nowak (Institute of Mathematics, Austria) and P. Zimmermann (Caramel, INRIA Nancy-Lorraine), has designed new methods for computing guaranteed enclosures of the Masser-Gramain constant, a two-dimensional analogue of the Euler-Mascheroni constant [86].

• G. Melquiond, in collaboration with J-M. Muller (CNRS, Arénaire, LIP, ÉNS Lyon) and E. Martin-Dorel (Arénaire, LIP, ÉNS Lyon), has worked on weakening the assumptions floating-point error-free transformations rely on [39].
6. New Results

6.1. Petri Nets and their Synthesis

Participant: Philippe Darondeau.

6.1.1. Petri Net Reachability Graphs: Decidability Status of FO Properties

In [24], we investigate the decidability and complexity status of model-checking problems on unlabelled reachability graphs of Petri nets by considering first-order, modal and pattern-based languages without labels on transitions or atomic propositions on markings. We consider several parameters to separate decidable problems from undecidable ones. Not only are we able to provide precise borders and a systematic analysis, but we also demonstrate the robustness of our proof techniques.

6.1.2. Separability in Persistent Petri Nets

We prove in [14] that the separability of plain, bounded, reversible and persistent Petri nets, a class of nets that extends the well-known live and bounded marked graphs. We establish first a weak form of separability, already known to hold for marked graphs, in which every firing sequence is simulated by a firing sequence of k parallel instances identical firing counts. We establish on top of this a strong form of separability, in which every firing sequence of is simulated by identical firing sequences.

6.1.3. Petri Net Distributability

A Petri net is distributed if, given an allocation of transitions to (geographical) locations, no two transitions at different locations share a common input place. A system is distributable if there is some distributed Petri net implementing it. We address in [23] the question of which systems can be distributed, while respecting a given allocation. The paper states the problem formally and discusses several examples illuminating — to the best of the authors’ knowledge — the current status of this work.

6.2. Heterogeneous Systems

Participants: Eric Badouel, Albert Benveniste, Timothy Bourke, Benoît Caillaud.

6.2.1. Hybrid Modeling

Hybrid modeling tools like Simulink have evolved from simulation platforms into development platforms on which testing, verification and code generation are also performed. It is critical to ensure that the results of simulation, compilation and verification are consistent. Synchronous languages have addressed these issues but only for discrete systems. Reprising earlier work [32], we present in [21] a hybrid modeler built from a synchronous language and an off-the-shelf numerical solver. The main novelty is a language with hierarchical automata that can be arbitrarily mixed with data-flow and ordinary differential equations (ODEs). A type system statically ensures that discrete state changes are aligned with zero-crossing events and that the function passed to the numerical solver has no side-effects during integration. Well-typed programs are compiled by source-to-source translation into synchronous code which is then translated into sequential code using an existing synchronous language compiler.
Starting from a minimal, yet full-featured, Lustre-like synchronous language, we present in [22] a conservative extension where data-flow equations can be mixed with ordinary differential equations (ODEs) with possible reset. A type system is proposed to statically distinguish discrete computations from continuous ones and to ensure that signals are used in their proper domains. We propose a semantics based on non-standard analysis which gives a synchronous interpretation to the whole language, clarifies the discrete/continuous interaction and the treatment of zero-crossings, and also allows the correctness of the type system to be established. The extended data-flow language is realized through a source-to-source transformation into a synchronous subset, which can then be compiled using existing tools into routines that are both efficient and bounded in their use of memory. These routines are orchestrated with a single off-the-shelf numerical solver using a simple but precise algorithm which treats causally-related cascades of zero-crossings. We have validated the viability of the approach through experiments with the Sundials library.

6.2.2. Distributed Structured Documents

Evaluation of attributes w.r.t. an attribute grammar can be obtained by inductively computing a function expressing the dependencies of the synthesized attributes on inherited attributes. This higher-order functional approach to attribute grammars leads to a straightforward implementation using a higher-order lazy functional language like Haskell. The resulting evaluation functions are, however, not easily amenable to optimization rules. We present in [12] an alternative first-order functional interpretation of attribute grammars where the input tree is replaced with an extended cyclic tree each node of which is aware of its context viewed as an additional child tree. By the way, we demonstrate that these cyclic representations of zippers (trees with their context) are natural generalizations of doubly-linked lists to trees over an arbitrary signature.

6.3. Component-Based Design

Participants: Eric Badouel, Albert Benveniste, Benoît Caillaud, Benoît Delahaye, Sophie Pinchinat.

6.3.1. The Modal Interface Theory

In [18], we present the modal interface theory, a unification of interface automata and modal specifications, two radically dissimilar models for interface theories. Interface automata is a game-based model, which allows the designer to express assumptions on the environment and which uses an optimistic view of composition: two components can be composed if there is an environment where they can work together. Modal specifications are a language theoretic account of a fragment of the modal mu-calculus logic with a rich composition algebra which meets certain methodological requirements but which does not allow the environment and the component to be distinguished. The present paper contributes a more thorough unification of the two theories by correcting a first attempt in this direction by Larsen et al., drawing a complete picture of the modal interface algebra, and pushing the comparison between interface automata, modal automata and modal interfaces even further.

6.3.2. A Stochastic Interface Theory

Notions of specification, implementation, satisfaction, and refinement, together with operators supporting stepwise design, constitute a specification theory. In [16], we construct such a theory for Markov Chains (MCs) employing a new abstraction of a Constraint MC. Constraint MCs permit rich constraints on probability distributions and thus generalize prior abstractions such as Interval MCs. Linear (polynomial) constraints suffice for closure under conjunction (respectively parallel composition). This is the first specification theory for MCs with such closure properties. We discuss its relation to simpler operators for known languages such as probabilistic process algebra. Despite the generality, all operators and relations are computable.

6.3.3. Contract-Based Compositional Analysis of Stochastic Systems

A contract allows to distinguish hypotheses made on a system (the guarantees) from those made on its environment (the assumptions). In [17], we focus on models of Assume/Guarantee contracts for (stochastic) systems. We consider contracts capable of capturing reliability and availability properties of such systems. We also show that classical notions of Satisfaction and Refinement can be checked by effective methods thanks to a reduction to classical verification problems. Finally, theorems supporting compositional reasoning and enabling the scalable analysis of complex systems are also studied.
6.3.4. Modal event-clock specifications for timed component-based design

On the one hand, modal specifications are classic, convenient, and expressive mathematical objects to represent interfaces of component-based systems. On the other hand, time is a crucial aspect of systems for practical applications, e.g. in the area of embedded systems. And yet, only few results exist on the design of timed component-based systems. In [13], we propose a timed extension of modal specifications, together with fundamental operations (conjunction, product, and quotient) that enable reasoning in a compositional way about timed system. The specifications are given as modal event-clock automata, where clock resets are easy to handle. We develop an entire theory that promotes efficient incremental design techniques.

6.4. Scheduling and Supervisory Control

Participants: Eric Badouel, Benoît Caillaud, Philippe Darondeau.

6.4.1. Model Identification and Synthesis of Discrete-Event Systems

Book chapter [28] focuses on two important and tightly related problems, namely the identification and synthesis of discrete-event systems. Particular attention is devoted to two main formalisms in this area, i.e., finite state automata and Petri nets. The goal of this chapter is to provide a collection of references in this framework, and discuss the main research areas where such problems have been investigated. Due to the extensive literature, only some of the results are discussed in a certain detail, such as the basic ideas related to the theory of regions and the synthesis of labeled Petri nets, while other results are simply mentioned and the reader is addressed to the specific contributions for more details.

6.4.2. Assembling Sessions

Sessions are a central paradigm in Web services to implement decentralized transactions with multiple participants. Sessions enable the cooperation of workflows while at the same time avoiding the mixing of workflows from distinct transactions. Languages such as BPEL, ORC, AXML that implement Web Services usually realize sessions by attaching unique identifiers to transactions. The expressive power of these languages makes the properties of the implemented services undecidable. In [25], we propose a new formalism for modelling web services. Our model is session-based, but avoids using session identifiers. The model can be translated to a dialect of Petri nets that allows the verification of important properties of web services.

6.4.3. Towards Distributed Control of Discrete Event Systems

To initiate a discussion on the modeling requirements for distributed control of discrete-event systems, a partially-automated region based methodology is presented in [26]. The methodology is illustrated via a well-known example from distributed computing: the dining philosophers.

6.4.4. Communicating Decentralized Control

Frameworks that incorporate communication into decentralized supervisory control theory address the following problem: find locations in the evolution of the plant behavior where some supervisors send information so that a supervisor that was unable to make the correct control decision prior to receiving external information, is now capable of making the correct control decision. We propose in [19] a solutions to this problem and identify an earliest and a latest placement where such communication results in the synthesis of a correct control solution. In addition to a first and last communication opportunity, there may be a selection of intermediate possibilities where communication would produce the correct control solution. We present a computable procedure to identify a broader range of suitable communication locations.

6.4.5. Residuation of tropical series: rationality issues

In [20], the decidability of existence, rationality of delay controllers and robust delay controllers are investigated for systems with time weights in the tropical and interval semifields. Depending on the (max,+)- or (min,+) rationality of the series specifying the controlled system and the control objective, cases are identified where the controller series defined by residuation is rational, and when it is positive (i.e., when delay control is feasible). When the control objective is specified by a tolerance, i.e. by two bounding rational series, a nice case is identified in which the controller series is of the same rational type as the system specification series.
6.5. Games, Logic and System Synthesis
Participants: Bastien Maubert, Sophie Pinchinat.

6.5.1. Opacity Issues in Games with Imperfect Information

In [27], we study the class of games with opacity condition, which are two-player games with imperfect information in which one of the players only has imperfect information, and where the winning condition relies on the information he has along the play. Those games are relevant for security aspects of computing systems: a play is opaque whenever the player who has imperfect information never "knows" for sure that the current position is one of the distinguished "secret" positions. We study the problems of deciding the existence of a winning strategy for each player, and we call them the opacity-violate problem and the opacity-guarantee problem. Focusing on the player with perfect information is new in the field of games with imperfect-information because when considering classical winning conditions it amounts to solving the underlying perfect-information game. We establish the EXPTIME-completeness of both above-mentioned problems, showing that our winning condition brings a gap of complexity for the player with perfect information, and we exhibit the relevant opacity-verify problem, which noticeably generalizes approaches considered in the literature for opacity analysis in discrete-event systems. In the case of blindfold games, this problem relates to the two initial ones, yielding the determinacy of blindfold games with opacity condition and the PSPACE-completeness of the three problems.

6.5.2. Hardness of preorder checking for basic formalisms

We investigate in [15] the complexity of preorder checking when the specification is a flat finite-state system whereas the implementation is either a non-flat finite-state system or a standard timed automaton. In both cases, we show that simulation checking is Exptime-hard, and for the case of a non-flat implementation, the result holds even if there is no synchronization between the parallel components and their alphabets of actions are pairwise disjoint. Moreover, we show that the considered problems become Pspace-complete when the specification is assumed to be deterministic. Additionally, we establish that comparing a synchronous non-flat system with no hiding and a flat system is Pspace-hard for any relation between trace containment and bisimulation equivalence.
6. New Results

6.1. Real Solving Polynomial Systems

In [20], we describe an algorithm (VQE) for a variant of the real quantifier elimination problem (QE). The variant problem requires the input to satisfy a certain extra condition, and allows the output to be almost equivalent to the input. The motivation/rationale for studying such a variant QE problem is that many quantified formulas arising in applications do satisfy the extra conditions. Furthermore, in most applications, it is sufficient that the output formula is almost equivalent to the input formula. The main idea underlying the algorithm is to substitute the repeated projection step of CAD by a single projection without carrying out a parametric existential decision over the reals. We find that the algorithm can tackle important and challenging problems, such as numerical stability analysis of the widely-used MacCormack’s scheme. The problem has been practically out of reach for standard QE algorithms in spite of many attempts to tackle it. However the current implementation of VQE can solve it in about 12 hours. This paper extends the results reported at the conference ISSAC 2009.

We also focused on the interaction of real solving polynomial system with global optimization. Let \( f \in \mathbb{Q}[X_1, \ldots, X_n] \) of degree \( D \). Algorithms for solving the unconstrained global optimization problem \( f_{\mathbb{R}}^* = \inf_{x \in \mathbb{R}^n} f(x) \) are of first importance since this problem appears frequently in numerous applications in engineering sciences. This can be tackled by either designing appropriate quantifier elimination algorithms or by certifying lower bounds on \( f_{\mathbb{R}}^* \) by means of sums of squares decompositions but there is no efficient algorithm for deciding if \( f_{\mathbb{R}}^* \) is a minimum. The paper [41] is dedicated to this important problem. We design an algorithm that decides if \( f_{\mathbb{R}}^* \) is reached over \( \mathbb{R}^n \) and computes a point \( x_{\mathbb{R}}^* \in \mathbb{R}^n \) such that \( f(x_{\mathbb{R}}^*) = f_{\mathbb{R}}^* \) if such a point exists. If \( L \) is the length of a straight-line program evaluating \( f \), a probabilistic version of the algorithm runs in time \( \tilde{O}(n^2(L + n^2)(D(D-1)^{n-1})^2) \). Experiments show its practical efficiency.

Global optimization problems can also be tackled by computing algebraic certificates of positivity through sums of squares decompositions. Let \( f_1, \ldots, f_p \) and \( f \) be polynomials in \( \mathbb{Q}[X_1, \ldots, X_n] \) and let \( V = V(f_1, \ldots, f_p) \subset \mathbb{C}^n \) be the algebraic variety defined by \( f_1 = \cdots = f_p = 0 \) whose dimension is denoted by \( d \). Assume in the sequel that the ideal \( \langle f_1, \ldots, f_p \rangle \) is radical and equidimensional and that \( V \) is smooth. In [18], up to a generic linear change of variables, we construct families of polynomials \( M_0, \ldots, M_d \) in \( \mathbb{Q}[X_1, \ldots, X_n] \) such that \( f(x) \geq 0 \) for all \( x \in V \cap \mathbb{R}^n \) if and only if \( f \) can be written as a sum of squares of polynomials in \( \mathbb{R}[X_1, \ldots, X_n] \) modulo \( \langle M_i \rangle \) for \( 0 \leq i \leq d \). Such an algebraic certificate of positivity is simpler than the more general Positivstellensatz in Real Algebra. It can be used to certify lower bounds on \( f_{\mathbb{R}}^* = \inf_{x \in V \cap \mathbb{R}^n} f(x) \). Also, computing a numerical approximation of such certificates of positivity can be reformulated as a semidefinite program which can be solved efficiently. We provide numerical experiments showing the effectiveness of our approach.

In [25], we consider the problem of constructing roadmaps of real algebraic sets. This problem was introduced by Canny to answer connectivity questions and solve motion planning problems. Given \( s \) polynomial equations with rational coefficients, of degree \( D \) in \( n \) variables, Canny’s algorithm has a Monte Carlo cost of \( s^{n\log (s)D^{O(n^2)}} \) operations in \( \mathbb{Q} \); a deterministic version runs in time \( s^{n \log (s)D^{O(n^2)}} \). A subsequent improvement was due to Basu, Pollack and Roy, with an algorithm of deterministic cost \( s^{D + 1}D^{O(n^2)} \) for the more general problem of computing roadmaps of a semi-algebraic set \( (d \leq n \) is the dimension of an associated object). We give a probabilistic algorithm of complexity \((nD)^{O(n^{1.5})}\) for the problem of computing a roadmap of a closed and bounded hypersurface \( V \) of degree \( D \) in \( n \) variables, with a finite number of singular points. Even under these extra assumptions, no previous algorithm featured a cost better than \( D^{O(n^2)} \).
6.2. Zero dimensional Solve

Let \( I \subset \mathbb{K}[x_1, \ldots, x_n] \) be a 0-dimensional ideal of degree \( D \) where \( \mathbb{K} \) is a field. It is well-known that obtaining efficient algorithms for change of ordering of Gröbner bases of \( I \) is crucial in polynomial system solving. Through the algorithm \( \text{FGLM} \), this task is classically tackled by linear algebra operations in \( \mathbb{K}[x_1, \ldots, x_n]/I \).

With recent progress on Gröbner bases computations, this step turns out to be the bottleneck of the whole solving process.

In [38], we present an efficient algorithm that takes advantage of the sparsity structure of multiplication matrices appearing during the change of ordering. This sparsity structure arises even when the input polynomial system defining \( I \) is dense. As a by-product, we obtain an implementation which is able to manipulate 0-dimensional ideals over a prime field of degree greater than 30000. It outperforms the Magma/Singular/FGb implementations of \( \text{FGLM} \).

In [38], we investigate the particular but important shape position case. The obtained algorithm performs the change of ordering within a complexity \( O(D(N_1 + n \log (D))) \), where \( N_1 \) is the number of nonzero entries of a multiplication matrix (the density of the matrix). This almost matches the complexity of computing the minimal polynomial of one multiplication matrix. Then, we address the general case and give corresponding complexity results. Our algorithm is dynamic in the sense that it selects automatically which strategy to use depending on the input. Its key ingredients are the Wiedemann algorithm to handle 1-dimensional linear recurrence (for the shape position case), and the Berlekamp–Massey–Sakata algorithm from Coding Theory to handle multi-dimensional linearly recurring sequences in the general case.

6.3. Solving structured systems

Solving multihomogeneous systems, as a wide range of structured algebraic systems occurring frequently in practical problems, is of first importance. Experimentally, solving these systems with Gröbner bases algorithms seems to be easier than solving homogeneous systems of the same degree. Nevertheless, the reasons of this behaviour are not clear. In [16], we focus on bilinear systems (i.e. bihomogeneous systems where all equations have bidegree \((1,1)\)). Our goal is to provide a theoretical explanation of the aforementioned experimental behaviour and to propose new techniques to speed up the Gröbner basis computations by using the multihomogeneous structure of those systems. The contributions are theoretical and practical.

First, we adapt the classical \( F_5 \) criterion to avoid reductions to zero which occur when the input is a set of bilinear polynomials. We also prove an explicit form of the Hilbert series of bihomogeneous ideals generated by generic bilinear polynomials and give a new upper bound on the degree of regularity of generic affine bilinear systems. This leads to new complexity bounds for solving bilinear systems. We propose also a variant of the \( F_5 \) Algorithm dedicated to multihomogeneous systems which exploits a structural property of the Macaulay matrix which occurs on such inputs. Experimental results show that this variant requires less time and memory than the classical homogeneous \( F_5 \) Algorithm. Lastly, we investigate the complexity of computing a Gröbner basis for the grevlex ordering of a generic 0-dimensional affine bilinear system over \( k[x_1, \ldots, x_{n_x}, y_1, \ldots, y_{n_y}] \). In particular, we show that this complexity is upper bounded by \( O\left( \left(n_x + n_y + \min (n_x + 1, n_y + 1) \right)^{\omega \frac{N_1}{\min(n_x + 1, n_y + 1)}} \right) \), which is polynomial in \( n_x + n_y \) (i.e. the number of unknowns) when \( \min(n_x, n_y) \) is constant.

6.4. Structured Systems and Applications to Cryptanalysis

The Goppa Code Distinguishing (GCD) problem consists in distinguishing the matrix of a Goppa code from a random matrix. Up to now, it is widely believed that the GCD problem is a hard decisional problem. In [36], we present the first technique allowing to distinguish alternant and Goppa codes over any field. Our technique can solve the GCD problem in polynomial-time provided that the codes have rates sufficiently large. The key ingredient is an algebraic characterization of the key-recovery problem which reduces to the solving of a system of bi-homogeneous polynomial equations. The idea is to consider the dimension of the solution space of a linearized system deduced from the algebraic system describing the key-recovery. It turns
out that experimentally this dimension depends on the type of code. Explicit formulas derived from extensive experimentations for the value of the dimension are provided for “generic” random, alternant, and Goppa code over any alphabet. Finally, we give explanations of these formulas in the case of random codes, alternant codes over any field and binary Goppa codes.

6.5. Algebraic Cryptanalysis

The Isomorphism of Polynomials (IP) is one of the most fundamental problems in multivariate public key cryptography (MPKC). In In [23], we introduce a new framework to study the counting problem associated to IP. Namely, we present tools of finite geometry allowing to investigate the counting problem associated to IP. Precisely, we focus on enumerating or estimating the number of isomorphism equivalence classes of homogeneous quadratic polynomial systems. These problems are equivalent to finding the scale of the key space of a multivariate cryptosystem and the total number of different multivariate cryptographic schemes respectively, which might impact the security and the potential capability of MPKC. We also consider their applications in the analysis of a specific multivariate public key cryptosystem. Our results not only answer how many cryptographic schemes can be derived from monomials and how big the key space is for a fixed scheme, but also show that quite many HFE cryptosystems are equivalent to a Matsumoto-Imai scheme.

In [34], we present a practical cryptanalysis of the Identification Scheme proposed by Patarin at Crypto 1996. This scheme relies on the hardness of the Isomorphism of Polynomial with One Secret (IP1S), and enjoys shorter key than many other schemes based on the hardness of a combinatorial problem (as opposed to number theoretic problems). We present two new deterministic algorithms to attack the IP1S problem, and we rigorously analyze their complexity and success probability. We show that they can solve a constant fraction of all the instances of degree two in polynomial time.

In [33], we investigate the security of a generalization of HFE (multivariate and odd-characteristic variants). We propose an improved version of the basic Kipnis-Shamir key recovery attack against HFE. We then generalize the Kipnis-Shamir attack to Multi-HFE. The attack reduces to solve a MinRank problem directly on the public key. This leads to an improvement of a factor corresponding to the square of the degree of the extension field. We used recent results on MinRank to show that our attack is polynomial in the degree of the extension field. It appears that multi-HFE is less secure than original HFE for equal-sized keys. Finally, adaptations of our attack overcome several variants (i.e. minus modifier and embedding). As a proof of concept, we have practically broken the most conservative parameters given by Chen, Chen, Ding, Werner and Yang in 9 days for 256 bits security. In all, our results give a more precise picture on the (in)security of several variants of HFE proposed these last years.

In [31], we initiate the formal treatment of cryptographic constructions (“Polly Cracker”) based on the hardness of computing remainders modulo an ideal. We start by formalising and studying the relation between the ideal remainder problem and the problem of computing a Gröbner basis. We show both positive and negative results. On the negative side, we define a symmetric Polly Cracker encryption scheme and prove that this scheme only achieves bounded CPA security under the hardness of the IR problem. Furthermore, we show that a large class of algebraic transformations cannot convert this scheme to a fully secure Polly Cracker-style scheme. On the positive side, we formalise noisy variants of the ideal related problems. These problems can be seen as natural generalisations of the LWE problem and the approximate GCD problem over polynomial rings. After formalising and justifying the hardness of the noisy assumptions we show that noisy encoding of messages results in a fully IND-CPA secure somewhat homomorphic encryption scheme. Together with a standard symmetric-to-asymmetric transformation for additively homomorphic schemes, we provide a positive answer to the long standing open problem of constructing a secure Polly Cracker-style cryptosystem reducible to the hardness of solving a random system of equations. Indeed, our results go beyond that by also providing a new family of somewhat homomorphic encryption schemes based on new, but natural, hard problems. Our results also imply that Regev’s LWE-based public-key encryption scheme is (somewhat) multiplicatively homomorphic for appropriate choices of parameters.
6.6. Computer Algebra and Algorithmic Number Theory

The Elliptic Curve Discrete Logarithm Problem (ECDLP) has become the most attractive alternative to factoring for public key cryptography. Whereas subexponential factoring algorithms exist, solving the ECDLP in general can only be done in exponential time. Provided that a certain heuristic assumption holds, we present in [39] an index calculus algorithm solving ECDLP over any binary field $\mathbb{F}_q$ in time $O(2^c n^{2/3} \log n)$, where $c$ is a small constant. Our algorithm follows the index calculus method that was first introduced by Semaev and later developed by Gaudry and Diem. In particular, the main step consists in decomposing points of the curve with respect to an appropriately chosen factor basis. This part can be nicely reformulated as a purely algebraic problem consisting in finding solutions to a multivariate polynomial $f(x_1, \ldots, x_m) = 0$ such that all the variables $x_i$ belong to some vector subspace of $\mathbb{F}_q/\mathbb{F}_p$. We solve this problem by means of Gröbner basis techniques and analyse its complexity using the multihomogeneous structure of the equations. Even, if this paper is essentially theoretical and is not aiming for practical attacks, the new ideas developed here could be used to have practical attacks in the future. This of course represents a challenging open problem.
SECRET Project-Team

5. New Results

5.1. Symmetric cryptosystems

Participants: Céline Blondeau, Christina Boura, Baudoin Collard, Anne Canteaut, Pascale Charpin, Stéphane Jacob, Gohar Kyureghyan.

From outside, it might appear that symmetric techniques become obsolete after the invention of public-key cryptography in the mid 1970’s. However, they are still widely used because they are the only ones that can achieve some major features as high-speed or low-cost encryption, fast authentication, and efficient hashing. Today, we find symmetric algorithms in GSM mobile phones, in credit cards, in WLAN connections. Symmetric cryptology is a very active research area which is stimulated by a pressing industrial demand for low-cost implementations (in terms of power consumption, gate complexity,...). These extremely restricting implementation requirements are crucial when designing secure symmetric primitives and they might be at the origin of some weaknesses. Actually, these constraints seem quite incompatible with the rather complex mathematical tools needed for constructing a provably secure system.

The specificity of our research work is that it considers all aspects of the field, from the practical ones (new attacks, concrete specifications of new systems) to the most theoretical ones (study of the algebraic structure of underlying mathematical objects, definition of optimal objects). But, our purpose is to study these aspects not separately but as several sides of the same domain. Our approach mainly relies on the idea that, in order to guarantee a provable resistance to the known attacks and to achieve extremely good performance, a symmetric cipher must use very particular building blocks, whose algebraic structures may introduce unintended weaknesses. Our research work captures this conflict for all families of symmetric ciphers. It includes new attacks and the search for new building blocks which ensure both a high resistance to the known attacks and a low implementation cost. This work, which combines cryptanalysis and the theoretical study of discrete mathematical objects, is essential to progress in the formal analysis of the security of symmetric systems.

In this context, the very important challenges are the designs of low-cost ciphers and of secure hash functions. Most teams in the research community are actually working on the design and on the analysis (cryptanalysis and optimization of the performance) of such primitives.

5.1.1. Hash functions.

Following the recent attacks against almost all existing hash functions (MD5, SHA-0, SHA-1...), we have initiated a research work in this area, especially within the Saphir-2 ANR project and with several PhD theses. Our work on hash functions is two-fold: we have designed two new hash functions, named FSB and Shabal, which have been submitted to the SHA-3 competition, and we have investigated the security of several hash functions, including the previous standards (SHA-0, SHA-1...) and some other SHA-3 candidates.

Recent results:

- study of the algebraic properties of the recent hash function proposals, including the SHA-3 candidates Keccak and Luffa. This work includes a theoretical study of the algebraic degree of iterated functions composed of parallel applications of a smaller function [24].
- Upper bounds on the degree of an iterated permutation from the degree of the inverse of the inner transformation; this result has been applied both to hash functions and to block ciphers [31], [44].

5.1.2. Stream ciphers.

Our research work on stream ciphers is a long-term work which has been developed within the 4-year ANR RAPIDE project. It includes an important cryptanalytic effort on stream ciphers.
Recent results:

- Evaluation of the bias of parity-check relations in the context of cryptanalysis of combination generators with constituent devices which generate period sequences [13].
- Cryptanalysis of the recent stream cipher proposal Armadillo [21].

5.1.3. Block ciphers.

Even if the security of the current block cipher standard, AES, is not threaten when it is used in a classical context, there is still a need for the design of improved attacks, and for the determination of design criteria which guarantee that the existing attacks do not apply. This notably requires a deep understanding of all previously proposed attacks.

Recent results:

- Differential cryptanalysis with multiple differentials, multiple differential cryptanalysis on the lightweight block cipher Present [23].
- Use of tools from error correcting theory in linear cryptanalysis [36].
- Determination of the data complexity (i.e., of the required number of plaintexts-ciphertexts) and of the success probability of all statistical attacks against block ciphers [12].

5.1.4. Cryptographic properties and construction of appropriate building blocks.

The construction of building blocks which guarantee a high resistance to the known attacks is a major topic within our project-team, for stream ciphers, block ciphers and hash functions. The use of such optimal objects actually leads to some mathematical structures which may be the origin of new attacks. This work involves fundamental aspects related to discrete mathematics, cryptanalysis and implementation aspects. Actually, characterizing the structures of the building blocks which are optimal regarding to some attacks is very important for finding appropriate constructions and also for determining whether the underlying structure induces some weaknesses or not.

For these reasons, we have investigated several families of filtering functions and of S-boxes which are well-suited for their cryptographic properties or for their implementation characteristics. For instance, bent functions, which are the Boolean functions which achieve the highest possible nonlinearity, have been extensively studied in order to provide some elements for a classification, or to adapt these functions to practical cryptographic constructions. We have also been interested in functions with a low differential uniformity (e.g., APN functions), which are the S-boxes ensuring an (almost) optimal resistance to differential cryptanalysis.

Recent results:

- Study of the properties of the family of power functions with exponents $2^t - 1$. This family notably includes the cube function $x^3$ and the inverse function over a finite field with characteristic 2. In this work, the whole Walsh spectrum of $x^7$ is determined [11].
- Construction and study of the properties of new families of permutation polynomials over the field with $2^m$ elements; study of permutations with a linear structure: [14].
- Study of the algebraic properties (e.g. the algebraic degree) of the inverses of APN power permutations [47].

5.2. Code-based cryptography

Participants: Matthieu Finiasz, Grégory Landais, Rafael Misoczki, Ayoub Otmani, Nicolas Sendrier, Jean-Pierre Tillich.
Most popular public-key cryptographic schemes rely either on the factorization problem (RSA, Rabin), or on the discrete logarithm problem (Diffie-Hellman, El Gamal, DSA). These systems have evolved and today instead of the classical groups \((\mathbb{Z}/n\mathbb{Z})\) we may use groups on elliptic curves. They allow a shorter block and key size for the same level of security. An intensive effort of the research community has been and is still being conducted to investigate the main aspects of these systems: implementation, theoretical and practical security. It must be noted that these systems all rely on algorithmic number theory. As they are used in most, if not all, applications of public-key cryptography today (and it will probably remain so in the near future), cryptographic applications are thus vulnerable to a single breakthrough in algorithmics or in hardware (a quantum computer can break all those scheme).

Diversity is a way to dilute that risk, and it is the duty of the cryptographic research community to prepare and propose alternatives to the number theoretic based systems. The most serious tracks today are lattice-based cryptography (NTRU,...), multivariate cryptography (HFE,...) and code-based cryptography (McEliece encryption scheme,...). All these alternatives are referred to as post-quantum cryptosystems, since they rely on difficult algorithmic problems which would not be solved by the coming-up of the quantum computer.

The code-based primitives have been investigated in details within the project-team. The first cryptosystem based on error-correcting codes was a public-key encryption scheme proposed by McEliece in 1978; a dual variant was proposed in 1986 by Niederreiter. We proposed the first (and only) digital signature scheme in 2001. Those systems enjoy very interesting features (fast encryption/decryption, short signature, good security reduction) but also have their drawbacks (large public key, encryption overhead, expensive signature generation). Some of the main issues in this field are

- security analysis, implementation and practicality of existing solutions,
- reducing the key size, \(e.g.,\) by using rank metric instead of Hamming metric, or by using particular families of codes,
- address new functionalities, like hashing or symmetric encryption.

Recent results:

- A distinguishing attack on high rate Goppa codes [25]. This results does not lead to an attack on any code based cryptosystem, but, in some particular cases, it invalidates the security reduction. It was conjectured that there was no such distinguishers.
- A new class of codes for McEliece type cryptosystems offering more versatility [22]
- A generic attack on one-time signature based on codes (KKS type) [27].
- A improvement of generic decoding techniques when addressing multiple targets [28].

5.3. Error-correcting codes and applications

Participants: Mamdouh Abbara, Matthieu Finiasz, Vincent Herbert, Denise Maurice, Nicolas Sendrier, Jean-Pierre Tillich.

Decoding algorithms are extensively used for cryptanalyses. For instance, a classical cryptanalysis of the stream ciphers which rely on linear feedback shift register filtered by a Boolean function models the attacked cipher as the result of the transmission of a linear function through a very highly noisy channel. Then, removing the noise amounts to decoding a certain linear code. This code is highly structured, and one of the most efficient methods to decode it exploits the fact that it has low density parity-check equations, and thus can be decoded as a low-density parity-check code, with iterative algorithms. Furthermore, the problem of finding good approximations of ciphers amounts to a decoding problem of the first order Reed-Muller code. Local decoding is then used in this context, and enables various attacks, such as correlation attacks or linear cryptanalysis.

Besides the cryptographic applications of decoding algorithms, we also investigate two new application domains for decoding algorithms: reverse engineering of communication systems, and quantum error correcting codes for which we have shown that some of them can be decoded successfully with iterative decoding algorithms.
5.3.1. Algebraic error-correcting codes.

Finding lower bounds on the minimum distance of cyclic codes is an old and difficult problem. Cyclic codes with three zeroes correct at most three errors, that is have minimum distance at most 7. It is an interesting question to determine which cyclic codes with three zeroes have minimum distance 7. Vincent Herbert revisit this problem by using an algorithm due to Shaub. Some classification questions are addressed about three error correcting cyclic codes and some new results involving intensive computer search have been obtained [10], [26].

5.3.2. Quantum codes.

The knowledge we have acquired in iterative decoding techniques has also led to study whether or not the very same techniques could also be used to decode quantum codes. Part of the old ACI project “RQ” in which we were involved and the new ANR project “COCQ” are about this topic. It is worth noticing that protecting quantum information from external noise is an issue of paramount importance for building a quantum computer. It also worthwhile to notice that all quantum error-correcting code schemes proposed up to now suffer from the very same problem that the first (classical) error-correcting codes had: there are constructions of good quantum codes, but for the best of them it is not known how to decode them in polynomial time. Our approach for overcoming this problem has been to study whether or not the family of turbo-codes and LDPC codes (and the associated iterative decoding algorithms) have a quantum counterpart.

Recent results:
- a construction of a family of quantum turbo-codes with excellent error reducing performance under iterative decoding and this even for very noisy channels [29];
- a proof that this family has unbounded minimum distance [20].

5.3.3. Reverse engineering of communication systems.

To evaluate the quality of a cryptographic algorithm, it is usually assumed that its specifications are public, as, in accordance with Kerckhoffs principle\textsuperscript{2}, it would be dangerous to rely, even partially, on the fact that the adversary does not know those specifications. However, this fundamental rule does not mean that the specifications are known to the attacker. In practice, before mounting a cryptanalysis, it is necessary to strip off the data. This reverse engineering process is often subtle, even when the data formatting is not concealed on purpose. A typical case is interception; some raw data, not necessarily encrypted, is observed out of a noisy channel. To access the information, the whole communication system has first to be disassembled and every constituent reconstructed. Our activity within this domain, whose first aim is to establish the scientific and technical foundations of a discipline which does not exist yet at an academic level, has been supported by some industrial contracts driven by the DGA.

\textsuperscript{2}Kerckhoffs stated that principle in a paper entitled \textit{La Cryptographie militaire}, published in 1883.
6. New Results

6.1. Indistinguishability Proofs

Participants: Rohit Chadha, Vincent Cheval, Ştefan Ciobăcă, Hubert Comon-Lundh, Stéphanie Delaune, Steve Kremer.

Most existing results in verification of security protocols focus on trace properties such as secrecy or authentication. There are however several security properties that cannot be defined (or cannot be naturally defined) as trace properties and require the notion of indistinguishably. Typical examples are anonymity, privacy related properties or statements closer to security properties used in cryptography.

In the framework of the applied pi-calculus [54], as in similar languages based on equational logics, indistinguishability corresponds to a relation called trace equivalence. Roughly, two processes are trace equivalent when an observer cannot see any difference between the two processes.

Under some conditions, trace equivalence can be reduced to the problem of deciding symbolic equivalence, an equivalence relation introduced by M. Baudet [55]. However, the procedure proposed by Mathieu Baudet for deciding symbolic equivalence is complex and cannot be implemented in its current state. Moreover, this method can only deal with simple processes with trivial else branches and is restricted to the class of subterm-convergent equational theories. Unfortunately, this makes it unsuitable for some case studies of interest to the SECSI team, among which the FOO electronic voting protocol, and the electronic passport protocols.

In order to provide tool support to decide trace equivalence, Rohit Chadha, Ştefan Ciobăcă, and Steve Kremer propose a procedure that can handle a large set of cryptographic primitives. The procedure has been implemented in a prototype tool and has been effectively tested on examples (e.g., the FOO e-voting protocol). This paper is currently under submission.

Vincent Cheval, Hubert Comon-Lundh and Stéphanie Delaune have designed another procedure that allows one to check trace equivalence for a general class of processes [31]. In their class, they can model conditionals (with non-trivial else branches), private channels, and non-deterministic choice. The private authentication protocol and the various versions of the electronic passport protocol fall into their class.

6.2. Anonymous Credentials

Participants: Stéphanie Delaune, Malika Izabachène, Graham Steel.

Anonymous credentials plays an important role in non-interactive anonymous authentication: they allow a user to obtain certificates from organization and subsequently prove their possession in such a way that transactions of a same user remain unlinkable. In collaboration with Benoit Libert and Damien Vergnaud, Malika Izabachène present an anonymous credential scheme [39] in which a user can prove possession of appropriate attributes in an non-interactive fashion, by showing that these attributes satisfy a certain predicate (different type of predicates are handled).

Following this line of research on anonymous protocols, Stéphanie Delaune, Malika Izabachène and Graham Steel formalize unlinkability in the pi-calculus framework. They are exploring several scenarios in order to capture many adversarial strategies, especially in the context of low-cost devices, in which sensitive data are stored and identifier means are exchanged through public channels.

6.3. Security APIs

Participants: Stéphanie Delaune, Steve Kremer, Robert Künnemann, Graham Steel, Yusuke Kawamoto, Joe-Kai Tsay.
Security APIs allow untrusted code to access sensitive resources in a secure way. The idea is to design an interface between a trusted component, such as a smart card or cryptographic security module, and the untrusted outside world such that no matter what sequence of commands in the interface are called, and no matter what the parameters, certain good properties will continue to hold, e.g. the secret long term keys on the smartcard are never revealed. Designing such interfaces is very tricky, and several vulnerabilities in APIs in common use have come to light in recent years.

The members of the SECSI team have been studying the application of formal security analysis techniques to APIs, for the last few years. These APIs include industrial standards such as PKCS#11 and the Trusted Platform Module (TPM).

In [37], Delaune, Kremer and Steel present a Horn-clause-based framework for analyzing security protocols that use platform configuration registers (PCRs), which are registers for maintaining state inside the Trusted Platform Module (TPM). In their model, the PCR state space is unbounded, and experience shows that a naïve analysis using verification tools such as ProVerif or SPASS does not terminate. To address this, the authors extract a set of instances of the Horn clauses of the model, for which ProVerif does terminate on the chosen examples. The authors prove the soundness of this extraction process: no attacks are lost, that is, any query derivable in the more general set of clauses is also derivable from the extracted instances. The effectiveness of this framework is demonstrated in two case studies: a simplified version of Microsoft Bitlocker, and a digital envelope protocol that allows a user to choose whether to perform a decryption, or to verifiably renounce the ability to perform the decryption.

One of the reasons for the existence of security flaws that the members of the SECSI team identified when studying security APIs is the absence of definitions stating the expected security properties.

In [40], Kremer, Steel and Warinschi propose a much-needed formal definition of security for cryptographic key management APIs. The advantages of this definition are that it is general, intuitive, and applicable to security proofs in both symbolic and computational models of cryptography. This definition relies on an idealized API which allows only the most essential functions for generating, exporting and importing keys, and takes into account dynamic corruption of keys. Based on this the authors can define the security of more expressive APIs which support richer functionality. They illustrate their approach by showing the security of APIs both in symbolic and computational models.

More recently, Kremer, Künnemann and Steel go even a step further in that direction and present the first key-management functionality in Canetti’s Universal Composability (UC) framework. It allows one to enforce a wide range of security policy and is highly extensible. The authors illustrate its use by proving an implementation of a Security API secure with respect to arbitrary key-usage operations and explore a proof technique that allows to store cryptographic keys externally, a novelty in the UC framework. This work is currently submitted.

In other recent work, in collaboration with Riccardo Focardi at the University of Venice, Kawamoto, Steel and Tsay have investigated the error behaviour of functions in the PKCS#11 API of various cryptographic devices including security tokens, electronic ID cards and Hardware Security Modules (HSMs). In certain circumstances attackers can take advantage of errors reported to make cryptanalytic attacks on functions in the API. Taking the example of the command used to import and encrypted key onto the device, they have discovered a number of so-called ‘error oracle attacks’ based on variations of well-known padding attacks due to Bleichenbacher and Vaudenay. This work has also recently been submitted. A number of vulnerability reports have been sent to manufacturers and national agencies.

6.4. Mobile Ad-Hoc Networks

Participants: Mathilde Arnaud, Morten Dahl, Stéphanie Delaune, Graham Steel.

Mobile ad hoc networks consist of mobile wireless devices which autonomously organize their communication infrastructure: each node provides the function of a router and relays packets on paths to other nodes. Finding these paths in an a priori unknown and constantly changing network topology is a crucial functionality of any ad hoc network. Specific protocols, called routing protocols, are designed to ensure this functionality known as
Secure routing protocols use cryptographic mechanisms in order to prevent a malicious node from compromising the discovered route and they often perform some recursive tests on received messages. Mathilde Arnaud, Véronique Cortier and Stéphanie Delaune provide \( \text{NPTIME} \) decision procedures for protocols with recursive tests and for a bounded number of sessions [26]. They also revisit constraint system solving, providing a complete symbolic representation of the attacker knowledge.

In the context of vehicular ad-hoc networks, to improve road safety, a vehicle-to-vehicle communication platform is currently being developed by consortia of car manufacturers and legislators. In [35], Morten Dahl, Stéphanie Delaune and Graham Steel propose a framework for formal analysis of privacy in location based services such as anonymous electronic toll collection. They give a formal definition of privacy, and apply it to the VPriv scheme for vehicular services. They analyse the resulting model using the ProVerif tool, concluding that the privacy property holds only if certain conditions are met by the implementation. Their analysis includes some novel features such as the formal modelling of privacy for a protocol that relies on interactive zero-knowledge proofs of knowledge and list permutations.

6.5. Composition Results

Participants: Céline Chevalier, Stéphanie Delaune, Steve Kremer.

Céline Chevalier, Stéphanie Delaune, and Steve Kremer investigate the composition of protocols that share a common weak secret [32]. This situation arises when users employ the same password on different services. More precisely they study whether resistance against guessing attacks composes when a same password is used. More precisely, they present a transformation which maps a password protocol that is secure for a single protocol session (a decidable problem) to a protocol that is secure for an unbounded number of sessions. Their result provides an effective strategy to design secure password protocols: (i) design a protocol intended to be secure for one protocol session; (ii) apply the transformation and obtain a protocol which is secure for an unbounded number of sessions. This technique also applies to compose different password protocols allowing one to obtain both inter-protocol and inter-session composition.

6.6. Protecting Hypervisors from Denial of Service Attacks

Participant: Hedi Benzina.

Hedi Benzina showed that hypervisors can be protected from some denial of service attacks by allowing administrators to write security policies in a simple language [41]. He implemented the RuleGen tool, which translates these policies into Orchids signatures.

6.7. Soundness Results: Some Limitations

Participant: Hubert Comon-Lundh.

Soundness results aim at bridging the gap between computational and symbolic security; they show that some symbolic model, in which messages are terms and the attacker is a formal process, faithfully abstracts the computational model, in which messages are bitstrings and the attacker is any probabilistic polynomial time Turing machine. Such results allow one to derive strong security guarantees, while reasoning at an abstract level. They have been developed for several cryptographic primitives (e.g. symmetric and asymmetric encryption, signatures, hash) and security properties.

These results however suffer from some severe limitations, as Hubert Comon-Lundh and Véronique Cortier demonstrate [34], focusing on symmetric encryption.

6.8. Model-Checking Reactive Probabilistic Systems

Participant: Rohit Chadha.
Rohit Chadha along with A. Prasad Sistla and Mahesh Viswanathan continued their study on reactive probabilistic systems modeled as Probabilistic Büchi Automata (PBA) in [30]. Reactive probabilistic systems are probabilistic non-deterministic systems in which the nondeterminism is resolved by an external environment which is oblivious of the “current” state of the system. This paper investigates the power of PBA when the threshold probability of acceptance is non-extremal, i.e., is a value strictly between 0 and 1. Many practical randomized algorithms are designed to work under non-extremal threshold probabilities and thus it is important to study power of PBAs for such cases. The paper presents a number of surprising expressiveness and decidability results for PBAs when the threshold probability is non-extremal. Some of these results sharply contrast with the results for extremal threshold probabilities. The paper also presents results for Hierarchical PBAs and for an interesting subclass of them called simple PBAs.

Rohit Chadha along with V. Korthikranthi, M. Viswanathan, G. Agha and Y. Kwon also study reactive probabilistic systems in [28]. In [28], reactive probabilistic systems are viewed as transformers of probability distributions, giving rise to a labeled transition system over the probability distributions over the states of the system. Thus, a reactive probabilistic system can be seen as defining a set of executions where each execution is a sequence of probability distributions. Reasoning about sequences of distributions allows one to express properties not expressible in standard probabilistic logics like PCTL; examples include expressing bounds on transient rewards and expected values of random variables, as well as comparing the probability of being in one set of states at a given time with another set of states. With respect to such a semantics, the model-checking problem is undecidable. In this paper, the authors identify a special class of systems called semi-regular Markov Decision Processes that have a unique non-empty, compact, invariant set of distributions, for which they show that checking any $\omega$-regular property is decidable. Their decidability result also implies that for semi-regular probabilistic finite automata with isolated cut-points, the emptiness problem is decidable.

6.9. Continuous Random Variables

Participant: Jean Goubault-Larrecq.

Continuing work on probabilistic and non-deterministic choice in a domain-theoretic setting, Jean Goubault-Larrecq and Daniele Varacca (PPS, University Paris 7) proposed a new monad for probabilistic choice, that of continuous random variables [38]. The usual Jones-Plotkin monad of continuous valuations, although simple enough, suffers from the defect that no category of continuous domains is known that would be both Cartesian-closed (i.e., would allow one to interpret functions) and stable under the Jones-Plotkin monad.

Jean Goubault-Larrecq and Daniele Varacca managed to show that a related monad, that of continuous random variables, inspired from the notion of a random variable in probability theory, did not suffer from this defect: the category of bc-domains is indeed both Cartesian-closed and stable under this monad. Moreover, the authors showed that using one or the other monad gave semantics to higher-order probabilistic programs that were indistinguishable at ground types. Finally, they used this to solve an open problem by Escardò, namely that observational equivalence of probabilistic higher-order programs is recursively enumerable.

6.10. Choquet-Kendall-Matheron Theorems

Participant: Jean Goubault-Larrecq.

One of the results obtained by Jean-Goubault-Larrecq in his theory of semantics for mixed non-deterministic and probabilistic choice [60] is that there is a one-to-one correspondence between continuous credibilities over some (state) space $X$ and certain compact subsets of the space of all continuous valuations over $X$, under mild assumptions on $X$. Similar theorems were produced by Choquet in the 1950s, refined by Kendall, then by Matheron in the 1970s, with applications in random set theory, among others.

Klaus Keimel and Jean Goubault-Larrecq produced an extremely simple proof of this fact [22], based on a simple special case of Groemer’s integral theorem. This proof also produces a much more general result than what was known earlier, as it does not assume that $X$ is second-countable or Hausdorff, and only local compactness.
A domain-theoretic view is that this is a representation theorem for mixed demonic choice and probabilistic choice; the angelic and erratic cases are also covered by Goubault-Larrecq and Keimel. These results had been presented at Dagstuhl Seminar 10232, June 2010.

6.11. Full Abstraction for Call-by-Value Programs with Choice

Participant: Jean Goubault-Larrecq.

Consider a programming language, with both an operational semantics, stating how one can implement a machine for this language, and a denotational semantics, which states what programs compute (not how). A classical question in programming language semantics is whether equality of denotations (from denotational semantics) coincides with contextual equivalence (from operational semantics). This is called full abstraction.

This question was first formulated for PCF by G. Plotkin in 1977, who showed that PCF was not fully abstract, although PCF plus a form of parallel or was. PCF is a simply-typed higher-order language, which one could see as a simple variant of the ML language without mutable state.

Jean Goubault-Larrecq examined the question for variants of PCF with various forms of non-deterministic and probabilistic choice. The latter are modeled denotationally by using his theory of previsions [61]. The most startling result is that the call-by-value variant of PCF with only angelic non-determinism is fully abstract, without the need for parallel or. Jean Goubault-Larrecq also showed that call-by-value PCF with angelic non-determinism and probabilistic choice is not fully abstract, but that this language plus so-called statistical test primitives is fully abstract. These results were presented at the Domains X Workshop, Swansea, Wales, UK, September 2011.
6. New Results

6.1. Point counting

In joint work with Pierrick Gaudry (CARAMEL) and David Kohel (Marseille), B. Smith developed an accelerated Schoof-type point counting algorithm for genus 2 curves with efficiently computable real multiplication endomorphisms. This project has made the computation of cryptographic-sized group orders practical for curves of genus 2 over prime finite fields. Going way beyond the current cryptographic range, the algorithm has been used to compute the group order of a 1024-bit Jacobian (smashing the previous 256-bit record of Gaudry and Schost). The article describing this algorithm has been awarded the Best Paper prize at ASIACRYPT 2011 [26], and an extended version has been invited for submission to Journal of Cryptology (the leading journal in the field).

6.2. Complex multiplication

F. Morain has been investigating new invariants for building class polynomials with small coefficients. This is still work in progress, though advertised in some talks of his.

6.3. Steganography

D. Augot, M. Barbier and Caroline Fontaine randomized the bounded syndrome coding problem on wet paper—an important embedding problem in steganography—such that this problem always has a solution [24]. This randomization is inspired the Courtois–Finiasz–Sendrier signature scheme, and shows nice results for linear perfect codes. In the special case of binary Hamming codes, this new method reaches exactly the necessary and sufficient bounds to ensure the embedding. The previous bounds were introduced by Carlos Munuera and M. Barbier [19]. These bounds depend on the dual distance of the code used. Thanks to the generalized Hamming weight, they proved that codes with low MDS rank are better in this context. Since the nature of their results are combinatorial, the authors generalized a bound for systematic non linear codes and showed that the non-linear systematic codes could be good candidates, as shown by the example of the Nadler code.

6.4. Homomorphic encryption

D. Augot, in collaboration with L. Perret from Salsa team, and Bochum Universität [22], designed a “secret-key” homomorphic encryption scheme, which is much more efficient than the public-key ones. It is based on \(q\)-ary Reed-Muller codes (or multi-variate evaluation-interpolation schemes). The main drawback is a severe restriction on the number of uses of a given secret key, but the ease of decrypting leads to think that the scheme can reencrypt its keys, enabling its reuse.

6.5. List decoding

D. Augot, M. Barbier and A. Couvreur wrote on how to decode binary Goppa codes. Augot, Barbier, and Couvreur presented a simple way, with a clean study of the complexity [23]. Using this list decoding algorithm, Barbier and Paulo Barreto proposed a key reduction for the McEliece cryptosystem [25]. The list decoding algorithm above allowed them to add more errors during the McEliece encryption step, making decoding attacks more difficult. At the same complexity of these attacks, using the list decoding algorithm decreases the public key size, which is the main drawback of this cryptosystem.
6.6. Explicit isogeny constructions

B. Smith constructed six infinite series of families of pairs of algebraic curves of arbitrarily high genus [20], defined over number fields, together with an explicit isogeny between the Jacobians of the curves splitting multiplication by 2, 3, or 4.

6.7. Quasi-cyclic codes

M. Barbier, Christophe Chabot and G. Quintin exhibited a bijective correspondence between the $\ell$-quasi-cyclic codes over $\mathbb{F}_q$ of length $m\ell$ and the set of ideals of $M_{\ell}(\mathbb{F}_q)[X]/(X^m - 1)$ [29]. They proposed also two new classes called the quasi-BCH and quasi-evaluation codes. For the first one, they introduced a unambiguous decoding algorithm, and thanks to the second one they designed 49 new codes over $\mathbb{F}_4$ which have a bigger minimum distance than previously known codes.

6.8. Root-finding over Galois rings

Jérémy Berthomieu, Grégoire Lecerf and G. Quintin presented a new algorithm to find all the roots of a given polynomial with coefficients in a Galois ring [30]. It has been used to study the behavior of the Sudan algorithm for Reed-Solomon codes over Galois rings. The algorithm has been adapted to work over rings of power series in several variables. It was implemented in the Quintix package of Mathemagix.
6. New Results

6.1. Octagonal Domains for Continuous Constraints (continuous)

Participants: Marie Pelleau, Charlotte Truchet, Frédéric Benhamou.

Domains in Continuous Constraint Programming (CP) are generally represented with intervals whose n-ary Cartesian product (box) approximates the solution space. We propose a new representation for continuous variable domains based on octagons. We generalize local consistency and split to this octagon representation, and we propose an octagonal-based branch and prune algorithm. Experimental results in IBEX on the COCONUT benchmarks suite show promising performance improvements on several classical benchmarks.

The corresponding paper Octagonal Domains for Continuous Constraints got the Best Student Paper Award at the 17th International Conference on Principles and Practice of Constraint Programming (CP 2011).

6.2. A Constraint Seeker (classification, constraints and application)

Participants: Nicolas Beldiceanu, Helmut Simonis.

We design a Constraint Seeker application which provides a web interface to search for global constraints in the global constraint catalog, given positive and negative ground examples. Based on the given instances the tool returns a ranked list of matching constraints, the rank indicating whether the constraint is likely to be the intended constraint of the user. A systematic evaluation is provided over the complete global constraint catalog.

The corresponding paper A Constraint Seeker: Finding and Ranking Global Constraints from Examples was published at the 17th International Conference on Principles and Practice of Constraint Programming (CP 2011) and the corresponding tool is available as a web service at http://www.emn.fr/z-info/sdemasse/gccat/.

6.3. Soft Problems (modelling and filtering)

Participants: Thierry Petit, Alexis De Clerq, Nicolas Beldiceanu, Narendra Jussien.

1. Side-constrained problems We experimentally shown that solving some classes of over-constrained problems requires an efficient (global) propagation of side constraints on variables representing violations. This work completes our previous studies, which highlighted the interest of a variable-based representation of violations for sake of modelling.


2. Soft cumulative scheduling We proposed a new constraint for handling cumulative problems with exceeds of capacities, in the case where the time horizon is fixed and the capacity can vary over time. Sweep and Edge-finding algorithms for classical cumulative problems have been modified so as to provide a filtering algorithm for our constraint. Experiments shown that instances involving several hundreds of activities can be solved with our approach.

The corresponding paper Filtering Algorithms for Discrete Cumulative Problems with Over-loads of Resource was published at the 17th International Conference on Principles and Practice of Constraint Programming (CP 2011).
6.4. Constraints with costs (modelling and filtering)

Participants: Thierry Petit, Nicolas Beldiceanu.

1. **Distribution of costs** We presented a new cardinality constraint dedicated to sequences of totally ordered cost. This constraint is useful to impose a precise (fair) distribution of the values taken by the cost variables in a given sequence, for instance in a bin packing with safety rules or in cumulative scheduling with overloads of resource. We came up with a generalized arc-consistency filtering algorithm, whose time complexity is linear in the sum of the number of variables and the number of values in the union of their domains.

The corresponding paper *the Ordered Distribute Constraint* was published in the *International Journal on Artificial Intelligence Tools* [15].

2. **The objective sum constraint.** Constraint toolkits generally provide a sum constraint whose propagation is poor to solve optimization problems. Therefore, solving real-world problems requires to develop ad-hoc techniques for handling sums, based on the particular properties of each problem. We proposed a generic technique which improves the standard sum constraint by exploiting the propagation of a set of constraints defined on the variables involved in a sum.

The corresponding paper *The Objective Sum Constraint* was published in the 8th International Conference on Integration of Artificial Intelligence and Operations Research Techniques in Constraint Programming for Combinatorial Optimization Problems (CPAIOR 2011) [25].

3. **the increasing sum constraint** Given a sequence of variables \(X\) of length \(n\), we consider the *increasing sum* constraint, which imposes the variables of \(X\) to be sorted in non strictly order, and that the sum of the variables of \(X\) is equal to \(s\). We propose an linear time bound-consistency algorithm for *increasing sum*. This work is related to problems with variable symmetries, when some of the symmetric variables are involved in sum constraints.

The paper *A Theta(n) Bound-Consistency Algorithm for the Increasing Sum Constraint* was published at the 17th International Conference on Principles and Practice of Constraint Programming (CP 2011), [24].

These works were all done in collaboration with J.-C. Régin (Univ. Nice).

6.5. Efficient Filtering Algorithms for Generic Constraints (filtering)

Participants: Nicolas Beldiceanu, Xavier Lorca, Thierry Petit.

**Counting constraints** We identified a family of counting constraints for which performing a complete filtering is a tractable problem. We provided a generalized arc-consistency algorithm and its specialization to some well-known global constraints. For some of them the obtained time complexity is linear in the sum of domain sizes, which improve or equals the best known results in the literature.

The corresponding paper *A Generalized Arc-Consistency Algorithm for a Class of Counting Constraints* was published at the 22th International Joint Conference on Artificial Intelligence (IJCAI’11) [23].

6.6. Efficient Filtering Algorithms and Heuristics for Dedicated Constraints (filtering)

Participants: Jean-Guillaume Fages, Xavier Lorca, Arnaud Malapert, Narendra Jussien.

1. **Revisiting the tree constraint** We revisit the *tree* constraint introduced at CPAIOR 2005 in [35] which partitions the nodes of a \(n\)-nodes, \(m\)-arcs directed graph into a set of node-disjoint anti-arborescences for which only certain nodes can be tree roots. We introduce in a new filtering algorithm that enforces generalized arc-consistency in \(O(n+m)\) time while the original filtering algorithm reaches \(O(nm)\) time. This result allows to tackle larger scale problems involving graph partitioning in CHOCO.
The corresponding paper *Revisiting the tree Constraint* was published at the 17th International Conference on Principles and Practice of Constraint Programming (CP 2011), [18].

2. **An Optimal Constraint Programming Approach to the Open-Shop Problem** We present an optimal constraint programming approach for the Open-Shop problem, which integrates recent constraint propagation and branching techniques with new upper bound heuristics for the Open-Shop problem. Randomized restart policies combined with nogood recording allow to search diversification and learning from restarts. This approach closed all remaining problems of the Brucker et al. and Guéret and Prins benchmarks with cpu times that are orders of magnitude lower than the best known metaheuristics.

The corresponding paper *An Optimal Constraint Programming Approach to the Open-Shop Problem* was published in the INFORMS Journal on Computing [13]. This work was done in collaboration with H. Cambazard (4C, INP Grenoble), C. Guéret (IRCCyN), A. Langevin (Ecole Polytechnique Montréal) and L.-M. Rousseau (Ecole Polytechnique Montréal).

### 6.7. Explinations for Contrain Programming (solver)

**Participants:** Narendra Jussien, Charles Prud’homme.

Constraint programming, despite its numerous successes in practice, suffers from not being really user-friendly when used by software engineers. Indeed, when faced with a no solution message from a constraint solver, it is hard yet impossible to identify the cause of this message: is it from a bad modelling, an ill-written constraint, a bug in the solver, .... Explinations for constraint programming have addressed this issue but are not yet widely used in the CP community. Recent work in the field tend to demonstrate that providing explanation-based user-oriented features can be done quite easily in modern constraint solvers. The objective of this line of work is to specify an user-oriented explanation-module for flexible solver architectures. This line of work is financed through a Google focused research grant. First results provide a complete solver independent specification of explanation algorithms, data structure for encoding nogoods and treatment algorithms. A reference implementation is being made within the new version of our solver CHOCO.

### 6.8. Bin repacking (constraints and application)

**Participants:** Sophie Demassey, Xavier Lorca, Fabien Hermenier.

A datacenter can be viewed as a dynamic bin packing system where servers host applications with varying resource requirements and varying relative placement constraints. When those needs are no longer satisfied, the system has to be reconfigured. Virtualization allows to distribute applications into Virtual Machines (VMs) to ease their manipulation. In particular, a VM can be freely migrated without disrupting its service, temporarily consuming resources both on its origin and destination.

We introduce the Bin Repacking Scheduling Problem in this context. This problem is to find a final packing and to schedule the transitions from a given initial packing, accordingly to new resource and placement requirements, while minimizing the average transition completion time. Our CP-based approach uses CHOCO and is implemented into Entropy, an autonomous VM manager which detects reconfiguration needs, generates and solves the CP model, then applies the computed decision. CP provides the awaited flexibility to handle heterogeneous placement constraints and the ability to manage large datacenters with up to 2000 servers and 10000 VMs.

The corresponding paper *Bin-Repacking Scheduling in Virtualized Datacenters* was published at the 17th International Conference on Principles and Practice of Constraint Programming (CP 2011), [19].

### 6.9. A Global Constraint for Multi-Agent Localization (constraints and application)

**Participants:** Gilles Chabert, Sophie Demassey.
This work has been initiated in the context of the Angels research project, in which G. Chabert has been involved during two years. The idea was to provide a new method for inter-localizing a group of autonomous underwater robots, traditional Kalman-based methods being inadequate in this context due to the highly nonlinear models derived from the sensing technology (electric fish robots).

We proposed, through a rough discretization of the signals, to consider the problem as a whole and under a combinatorial form. The level of the signal is basically associated to a cardinality of surrounding objects. This led to a global constraint, namely a conjunction of among constraint with interval value domains and in multiple dimension (objects are variables with several components).

Conjunction of among constraints had been already studied but not in the case of interval value domains. We therefore conducted a theoretical study and proved that the problem was tractable in the one-dimensional case, but not in higher dimension. We have also investigated different decompositions and filtering algorithms. This work is submitted to CPAIOR 2012. An INRIA research report has also been issued in June 2011, where the robotics aspects are described.
6. New Results

6.1. Real-time services and protocols

In this area, we developed, on the one hand, policies for managing the quality of service of operating support (mainly, networks and protocols) in order to meet the properties required by real time applications (hard real time, weakly hard real time) and, on the other hand, strategies for scheduling activities and memory management.

6.1.1. Network-MAC cross-layer framework for differentiated QoS in wireless sensor networks

Participants: Bilel Nefzi, Ye-Qiong Song.

Self-adaptive QoS mechanism is preferable in large-scale wireless sensor networks because of frequent network condition changes and the difficulty to statically configure the network parameters. A network-MAC cross-layer framework has been developed for facilitating packet scheduling, congestion control and energy consumption minimization. The work is based on a very simple idea of ‘collecting-and-transmitting burst’ scheme, called CoSenS (Collecting and Sending burst Scheme). The underlying MAC protocol is the widely adopted and deployed unslotted CSMA/CA of IEEE802.15.4. An algorithm is designed making the network self-adapts to the dynamic traffic changes. CoSenS provides a simple but efficient improvement of the MAC layer of IEEE 802.15.4 in terms of reliability, delay and throughput. Two extensions have been made: P-CoSenS for integrating priority management and S-CoSenS which adds dynamic sleeping period management [15], [8].

6.1.2. QoS in UWB-based sensor networks

Participants: Jamila Ben Sliman, Mounir Frikha [INIT, SupCom, Tunisia], Anis Koubaa [ISEP-IPP-Politechnic Institute of Porto, Portugal], Ye-Qiong Song.

IEEE802.15.4a provides higher data rates with smaller energy consumption thanks to the UWB (Ultra Wide Band) technology. However there exist few solutions on how to optimally exploit the great potential of this new standard. Similar to the industrial wireless network initiatives (e.g. WirelessHART, ISA SP100, IEEE802.15.4e), we developed PMCMTP, a multi-channel multi-time slot MAC protocol, which dynamically assigns channels and time slots for dense and large-scale WSNs with QoS support. The most challenging issue is providing a tradeoff between the resource efficiency and the multi-constrained QoS support. For this purpose, we propose a cross-layer algorithm JSAR (Joint duty-cycle Scheduling, resource Allocation and multi-constrained QoS Routing algorithm), based on multi-channel multi-time slot PMCMTP MAC. JSAR simultaneously combines a duty-cycle scheduling scheme for energy saving, a resource allocation scheme for efficient use of frequency channels and time slots, and an heuristic for multi-constrained routing protocol. The performance of JSAR has been evaluated, showing that it is suitable for on-line implementation [20].

6.1.3. Wireless Networked control systems (WNCS)

Participants: Najet Boughanmi, Eric Rondeau [CRAN UMR 7039, Nancy], Ye-Qiong Song.

With recent technology progress, it is becoming attractive to use wireless solutions for industrial process monitoring and control. Our approach for developing wireless networked control systems (WNCS) is based on the application and network co-design principle. The idea is to adjust on-line the network parameters (QoS) according to the needs of the control loops or Quality of Control (QoC). For achieving this on a WSN (Wireless Sensor Network) which is based on CSMA/CA MAC protocol, several enhancements have been done. In the PhD work of N. Boughanmi [7], we proposed several QoS management mechanisms with priority (based on blackburst scheme) for both the beacon enabled mode and the non-beacon enabled mode of the IEEE 802.15.4 protocol. QoS online adaptation protocols have also been designed which take as parameter the QoC of the system. These proposals are validated through simulations (using TrueTime) and partially with fixed priority scheduling analysis approach.
6.1.4. Wireless networks and middleware for ambient assisted living systems  
**Participants:** Claude Deroussent [MEDeTIC], Shahram Nourizadeh, Ye-Qiong Song, Jean-Pierre Thomesse.  
Wireless sensor networks have a great potential for contributing to build the ambient assisted living environment for elderly people at home (PhD work of S. Nourizadeh under LORIA-MEDETIC contract). However, several problems have to be addressed for the integration of WSN into the existing home automation networks. The PhD thesis of Shahram Nourizadeh addresses the problem of QoS in context-aware heterogeneous healthcare systems [9]. For providing real-time data collecting in a telehealthcare system composed of wireless sensor network and home automation network, a middleware called CodaQ is designed. It provides context data and takes into account QoS requirements of the applications. In [30], we showed how the data are modeled for including context information and how the QoS requirements are handled within a middleware. First measurements on a test bed have been carried out, showing the good performance of our design.

6.2. Evaluation and optimal dimensioning of real-time systems  
6.2.1. Code analyses and advanced visualization of software in real-time  
**Participants:** Damien Bodenes, Pierre Caserta, Olivier Zendra.

Last years, strong developments for our instrumentation, tracer and analyzer, had been performed, allowing us to really enter the experimental phase and getting first interesting results. A thorough state of the art had also been written.  
This year, in 2011, this state of the art paper was finally published in TVCG, a leading journal in computer visualization [10]. Thanks to the experimental setup efforts of previous years, we were in 2011 able to conduct good experiments. We designed and implemented a new way to visualize relations between software elements. These relations include static relations (is-a, direct heir, caller, callee, etc.) and dynamic ones (runtime caller, runtime callee). Our new relation visualization comprises a new way of placing way points so as to significantly decrease spatial and visual clutter when visualizing software systems with large numbers (thousands up to millions) of relations. This lead to a publication in VISSOFT, one of the most recognized conferences in the software visualization domain [23], as well as a Best Poster in ECOOP, one of the most recognized conferences in the object-oriented domain [46]. The important design and implementation work realized on the tracing and analysis software also lead to the publication of our method in ICOOOLPS 2011 [24].  
Work is going onto analyze polymorphism in Java programs, answering an apparently simple yet so far unanswered question: how much polymorphism is there actually in Java programs. This is of paramount importance, since a lot of work occur around polymorphism, which is an important concept, but no one is currently able to tell how much it impact programs in real life. We have begun writing this paper in cooperation with the LIRMM lab in Montpellier. In addition, we are in the process of finishing work pertaining to analyzing program evolutions, looking at differences between versions, and analyzing how dynamic metrics and static metrics correlate to evolution rate.

6.2.2. Open Power and Energy Optimization PLatform and Estimator  
**Participants:** Sophie Alexandre, Jonathan Ponroy, Kévin Roussel, Olivier Zendra.

Work in this domain was performed in the context of the ANR Open-PEOPLE (Open Power and Energy Optimization PLatform and Estimator) project, financed since the very end of 2008. INRIA Nancy Grand Est is responsible for the software part of the platform and is involved in memory management for low-power issues. Work in this project begun in April 2009 (kick-off meeting). We have finished setting up the very important infrastructure for the software part of the Open-PEOPLE platform. We have finished expressing the requirements for the platform, in order to start the actual developments and the actual integration of tools provided by the different partners. In 2011, we have finished expressing the platform architecture and user interface (GUI). We have also finished implementing the part of the software platform that is the remote control to the hardware platform. We finally have finished implementing the core of the software platform and canonical models handling. Several technical reports were written in relation to this work [38], [39], [40], [42], [43], [44].
We are now in the process of finishing the design and implementation of the PCMD (Power Consumption Model Development) and the PCAO (Power Consumption Analysis and Optimization) parts of the software platform, as well as the external tools integration work. The very first release of the whole Open-PEOPLE platform is expected early 2012. This lead to the several presentations and posters in conferences [51], [47], [52].

6.2.3. **Robustness evaluation for a critical distributed system**

**Participants:** Adrien Guénard, Lionel Havet, Françoise Simonot-Lion.

Wireless Sensor and Actuator Networks (WSANs) combine sensors and actuators interconnected by wireless networks in order to perform distributed sensing and acting tasks. Closed-loop controllers can therefore be deployed on WSANs; such systems have to meet specific requirements in terms of performance, dependability, energy and cost which raises great challenges due to the unreliability of wireless communications. A way to ensure that a system meets the required properties is to model it and go through its analysis. Building a model requires both deep knowledge on the system as well as on the used framework. Therefore there is a need for frameworks well-suited to the targeted systems and to the properties to verify. We proposed an approach meeting these conditions and a simulation framework, Samovar, based on Matlab / Simulink, allowing the modeling of the network protocols (Mac and routing services) and the resources sharing policy thanks to the TrueTime toolbox. Several classes of components (application, nodes, networks and middleware) and a clear semantics for their composition are identified. Furthermore, the design of Samovar was also driven by the need to transfer easily software components model between the concrete systems and its simulated model. The modeling and simulation method as well as the Samovar framework were assessed on several case studies. This work is supported by INRIA through the ADT SAMOVAR.

6.3. **Real-time scheduling**

6.3.1. **Scheduling of tasks in automotive multicore ECUs**

**Participants:** Aurélien Monot, Nicolas Navet, Françoise Simonot-Lion.

As the demand for computing power is quickly increasing in the automotive domain, car manufacturers and tier-one suppliers are gradually introducing multicore ECUs in their electronic architectures. Additionally, these multicore ECUs offer new features such as higher levels of parallelism which ease the respect of safety requirements such as the ISO 26262 and the implementation of other automotive use-cases. These new features involve also more complexity in the design, development and verification of the software applications. Hence, car manufacturers and suppliers will require new tools and methodologies for deployment and validation. We address the problem of sequencing numerous elementary software components, called runnables, on a limited set of identical cores. We show how this problem can be addressed as two sub-problems, partitioning the set of runnables and building the sequencing of the runnables on each core, which problems cannot be solved optimally due to their algorithmic complexity. We then present low complexity heuristics to partition and build sequencer tasks that execute the runnable set on each core, and derive lower bounds on their efficiency (i.e., competitive ratio). Finally, we address the scheduling problem globally, at the ECU level, by discussing how to extend this approach in the case where other OS tasks are scheduled on the same cores as the sequencer tasks. An article providing a summary of this line of work will appear in IEEE TII [14].

6.3.2. **Fine-grained hardware modeling in response time analyses**

**Participants:** Dawood Khan, Nicolas Navet.

Early in the design cycle, the two main approaches for verifying timing constraints and dimensioning the networks are worst-case schedulability analysis and simulation. In [29], we advocate that both provide complementary results and that, most often, none of them alone is sufficient. In particular, it is shown on automotive case-studies that response time distributions that can be derived from simulations cannot replace worst-case analysis. On the other hand, it is shown on examples that the analytical models, as used in worst-case analyses, are error-prone and often much simplified abstractions of the real system, which might lead to optimistic (i.e., unsafe) results.
As an illustration of the latter point, the classical WCRT analysis of Controller Area Network (CAN) implicitly assumes an infinite number of transmission buffers which is not the case in practice. This might lead high priority messages to suffer from priority inversion if the buffers are already occupied by low priority messages. This gives rise to an additional delay for high priority messages, which, if not considered, may result in a deadline violation. In an earlier work, we explained the cause of this additional delay and have extended the existing CAN schedulability analysis to integrate it. We have then studied the case where low-priority transmissions cannot be aborted because the communication controller or the driver does not allow it. We show on two case studies that the impact on response times is important and cannot be neglected in most real-time systems. This work was published in \cite{26}.

6.3.3. **Probabilistically analysable real-time system**

**Participants:** Liliana Cucu-Grosjean, Codé Lo, Luca Santinelli, Dorin Maxim.

The adoption of more complex hardware to respond to the increasing demand for computing power in next-generation systems exacerbates some of the limitations of static timing analysis for the estimation of the worst-case execution time (WCET) estimation. In particular, the effort of acquiring (1) detail information on the hardware to develop an accurate model of its execution latency as well as (2) knowledge of the timing behaviour of the program in the presence of varying hardware conditions, such as those dependent on the history of previously executed instructions. These problems are also known as the timing analysis walls. The probabilistic timing analysis, a novel approach to the analysis of the timing behaviour of next-generation real-time embedded systems, provides answers to timing analysis walls. In \cite{11} we have showed how the probabilistic timing analysis attacks the timing analysis walls. We have also presented experimental evidence that shows how probabilistic timing analysis reduces the extent of knowledge about the execution platform required to produce probabilistically-safe and tight WCET estimations.

6.3.4. **Optimal scheduling policies for real-time systems with probabilistic execution times**

**Participants:** Liliana Cucu-Grosjean, Luca Santinelli, Dorin Maxim, Olivier Buffet, Rob Davis [University of York].

We have investigated the problem of optimal priority assignment in fixed priority preemptive single processor systems where tasks have probabilistic execution times. We have identified three sub-problems which optimise different metrics related to the probability of deadline failures. For each sub-problem we have proposed an algorithm that is proved optimal. The first two algorithms are inspired by Audsley’s algorithm which is a greedy (lowest priority first) approach that is optimal in the case of tasks with deterministic execution times. Since we prove that such a greedy approach is not optimal for the third sub-problem, we have proposed a tree search algorithm in this case. These results were published in \cite{27}.

6.3.5. **Statistical analysis of real-time systems**

**Participants:** Liliana Cucu-Grosjean, Lu Yue, Thomas Nolte [Malardelan University], Ian Bate [University of York].

The response time analysis of real-time systems usually needs the knowledge of WCET estimation and this knowledge is not always available, e.g., because of intellectual property issues. This problem may be avoided by estimating statistically either the WCET of a task \cite{18} or the response time of each task \cite{37}.

6.3.6. **Multiprocessor scheduling of real-time systems with probabilistic execution times**

**Participants:** Liliana Cucu-Grosjean, Joel Goossens [Université Libre de Bruxelles].

After providing exact feasibility tests for the case of arbitrary tasks on unrelated processor in \cite{12}, we have proposed feasibility tests for tasks with probabilistic execution times \cite{34}. These tests are based on intervals that are proved to contain the highest probability of having tasks with deadline missed.

6.3.7. **Probabilistic Component-based Approaches**

**Participants:** Luca Santinelli, Patrick Meumeu Yomsi, Dorin Maxim, Liliana Cucu-Grosjean.
We have proposed a probabilistic component-based model which abstracts in the interfaces both the functional and non-functional requirements of such systems. This approach allows designers to unify in the same framework probabilistic scheduling techniques and compositional guarantees that go from soft to hard real-time. We have provided sufficient schedulability tests for task systems using such framework when the scheduler is either preemptive fixed-priority or earliest deadline first. These results were published in [35].

### 6.3.8. Mixed-criticality problems for probabilistic real-time systems

**Participants:** Bader Alahmad, Luca Santinelli, Liliana Cucu-Grosjean, Sathish Gopalakrishnan [University of British Columbia].

Critical embedded systems (CESs) face the need of new functionalities imposed by the end users. These new functionalities of CESs impose the utilization of complex architectures. The complex architectures increase the time variability of programs and this coupled with worst-case reasoning implies over-provisioned systems. Avoiding such over-provision became an important problem within CESs. One model answering such problem in the mixed-criticality problem. It is natural then to combine mixed-criticality with probabilistic approaches known to decrease the over-provision by taking into account the information that worst-case situations have low probability of occurrence. We have proposed and contrasted in [19] two probabilistic execution-behavior models for mixed-criticality independent job systems as they execute on a single machine. The models differ in both the system assumptions and the amount of job information they offer and exploit. While one model is compliant with the current standard practice of fixing jobs’ criticalities, the other is a proposal to treat job criticalities as random entities with predetermined probabilities of jobs being of certain criticalities throughout the lifetime of the system.

### 6.3.9. Energy optimization for real-time systems

**Participants:** Cristian Maxim, Liliana Cucu-Grosjean, Olivier Zendra.

Many embedded real-time systems integrate battery operated microprocessor systems with limited battery autonomy. Minimizing energy consumption is thus crucial. We have proposed in [28] an algorithm that improves energy consumption in real-time systems by combining Dynamic Voltage Scaling and a decrease in the number of preemptions. Our overall purpose is to focus on a specific part of the problem, namely selectively increasing frequency to lower the number of preemptions of a task to try and decrease the total energy consumption.

### 6.4. National Initiatives

#### 6.4.1. ANR Project “QUAlity of ServIce for wireless sensor networks and Mobile Objects – parameter aDaptatiOn (QUASIMODO)

**Participants:** Bilel Nefzi, François Despaux, Abdelkader Lahmadi, Adrien Guenard, Françoise Simonot-Lion, Ye-Qiong Song.

Quasimodo project (http://quasimodo.loria.fr/) is a joint "ANR Programme blanc international" project (March 2011 - February 2014) between LORIA laboratory- Nancy University and SKLICT - Zhejiang University, funded by ANR (n°ANR 2010 INTB 0206 01) and NSFC (n°NSFC 61061130563). The objective of the project is to provide an adaptive real-time quality of service (QoS) in wireless sensor and actuator networks (WSAN). The main QoS parameters are bounded delay and packet transmission success rate under stringent energy constraint and node mobility. The typical application scenario consists of a multi-robots tracking, coordination and cooperation through the WSAN. This first project year has been focused on the application scenario development, MAC layer design. A first scenario is the single mobile target tracking using mobile sensor nodes. Theoretic tracking algorithm has been developed based on Kalman filter estimation (an ellipse) and optimal sensor coverage. Its extension to including both communication delay and mobile node speed are under simulation using SAMOVAR simulator (http://samovar.loria.fr/). The MAC protocol will be based on the CoSenS framework with one enhancement for avoiding collisions during the waiting period of CoSenS.
6.4.2. PRST MISN / SSS Theme: Eco-Sûr2

Participants: Hugo Cruz Sanchez, Jamila Ben Sliman, Najet Boughanmi, Bilel Nefzi, Françoise Simonot-Lion, Ye-Qiong Song.

EcoSur2 aims at controlling and managing the energy production and consumption within a smart space. An important part of the system is the wireless sensor and actuator network (WSAN) which is used to sense devices and to activate actuators. The activities of TRIO team are focused on the design of wireless sensor network architectures that guarantee communication by optimizing the available resources of the WSAN, and the development of the interoperability solution aimed at linking the heterogeneous technologies used in the system. This activities include: 1) Implementation of a modified version of the Collection Tree Protocol (CTP) by using energy resources for routing decisions; 2) Implementation of asynchronous and periodical sensing applications on nodes; 3) The analysis of different platforms allowing to communicate with the available WSN equipment of previous projects and to facilitate the implementation of optimal communication mechanisms over different routing protocols (eg. Zigbee, RPL); 4) The implementation of the WSN system in the MPIGate to allow interoperability with other technologies (eg. building automation networks, WiFi, Ethernet); 5) The design of an energy oriented messaging system in a WSN; 6) The adaptation and the development of the QoS co-design approaches based on our previous results in networked control system co-design. This year we have focused on the network architecture design and the technical implementation. Contiki based sensor nodes have been chosen. Part of code of S-CoSenS protocol has been developed in Cooja simulator before the actual deployment on the sensor nodes.

6.4.3. PRST MISN / Thème IS: Smartroom for personal assisted living

Participants: Hugo Cruz Sanchez, Adrien Guenard, Lionel Havet, Bilel Nefzi, Shahram Nourizadeh, Ye-Qiong Song.

The aim of the smartroom project is to provide an open platform for developing and testing innovative solutions for personal assisted living. The main task of TRIO team is the definition of the communication architecture with interoperability and QoS support. MPIGate is the starting point for this project. The first phase is focused on the platform implementation. MPIGate has been extended to run on Web service platform [25], [45]. Further development will be around the auto-adaptive application-network middleware and the design of extra low-power and low duty-cycle protocols.

6.4.4. INRIA AEN PAL (Personal Assisted Living)

Participants: Hugo Cruz Sanchez, Shahram Nourizadeh, Ye-Qiong Song.

TRIO team has participated to the Large-scale initiative action AEN PAL project (http://pal.inria.fr/) which aims to provide technologies and services for improving the autonomy and quality of life for elderly and fragile persons. Communication is one of the key components for ensuring real-time data gathering and exchange between heterogeneous sensors and actuators (robots). TRIO team’s participation aims to design the most suitable communication architectures with guaranteed QoS. For this purpose the interface part of MPIGate has been revised in order to shift from a web server based gateway to a web service oriented architecture [36]. This part of work will mainly be supported via the upcoming ADT APL-PERCEE project which will start at the end of 2011.

6.5. European Initiatives

6.5.1. NOE High Performance Embedded Architectures and Compilation (HiPEAC)

Participant: Olivier Zendra.

The TRIO team is involved in the HiPEAC (High Performance Embedded Architecture and Compilation) European Network of Excellence (NoE). Olivier Zendra was initiator and leader in this context of a cluster of European Researchers “Architecture-aware compiler solutions for energy issues in embedded systems” from mid-2007 to mid-2009. A STREP proposal tentatively titled “Integrated and generic energy-aware adaptation for extreme computing systems” is currently being written, mostly in the context of this network of excellence, for submission in Call ICT 2011.9.8 FET Proactive: Minimising Energy Consumption of Computing to the Limit (MINECC).
6.5.2. PROARTIS - Probabilistically Analysable Real-Time Systems

Participants: Liliana Cucu-Grosjean, Luca Santinelli, Codé Lo, Dorin Maxim.

PROARTIS (http://www.proartis-project.eu/) is a STREP project within the FP7 call and it started on February 2010. It has six partners: Barcelona Supercomputing, University of York, University of Padova, INRIA and Airbus. The overarching objective of the PROARTIS project is to facilitate a probabilistic approach to timing analysis. The proposed approach will concentrate on proving that pathological timing cases can only arise with negligible probability, instead of struggling to eradicate them, which is arguably not possible and could severely degrade performance. This will be a major turn from previous approaches that seek analyzability by trying to predict with cycle accuracy the state of hardware and software through analysis.

The PROARTIS project will facilitate the production of analysable CRTE systems on advanced hardware platforms with features such as memory hierarchies and multi core processors. PROARTIS has the following overall strategic industrial goals:

- Increased performance, reliability and reduced costs by enabling critical real-time systems to take full advantage of advanced hardware like deep memory hierarchies and multi core processors. The use of these features will allow designers to schedule more tasks while reducing the weight, power consumption and the size of the whole system and maintaining the desired predictability. It will also reduce the risk of temporal budget overruns. Application-level tasks will have an execution behaviour free (with sufficient low probability) from pathological temporal overruns.
- Increased productivity by enabling software engineers to develop more complex real-time software systems through timing-aware systems that reveal crucial timing details while dramatically simplifying analysis. For example, memory latencies will be predicted with less effort, requiring knowledge only of the total number of memory accesses, rather than the exact memory addresses and memory access patterns.
- Reduced time-to-market by enabling trustworthy WCET and other analyses for large-scale real-time systems that will dramatically reduce testing time.

The work within this project during 2011 lead to the following two publications: [11] and [35].

6.5.3. TIMMO-2-USE - Timing Model - TOols, algorithms, languages, methodology, USE cases

Participants: Nicolas Navet, Françoise Simonot-Lion, Liliana Cucu-Grosjean, Ammar Oulamara, Luca Santinelli.

TIMMO-2-USE (http://timmo-2-use.org/) is an ITEA 2 European project and it started in November 2010. TIMMO-2-USE will address the specification, transition and exchange of different types of timing information throughout different steps of the development process. The general goal is to evaluate and enhance standards for different applications in the development by different technical use cases covering multiple abstraction levels and tools. For this, TIMMO-2-USE will bring the AUTOSAR standard, TADL and EAST-ADL2 into different applications like WCET analysis and in-the-loop scenarios. This will bring new algorithms and tools for the transition and conversion of timing information between different tools and abstraction level based on a new advanced methodology which, in turn, will be based on a combination of the TIMMO and the ATESST2 methodologies. The main impact of TIMMO-2-USE will be:

- Improved, predictable development cycle: An extended and further developed infrastructure for handling timing constraints, containing additional features, will increase the predictability and effectiveness of the development cycle even more. As a result, both development cost and development time are expected to go down due to fewer costly design iterations, while at the same time the resulting design will moreover be more reliable.
- Reduced time-to-market by massive reuse: Reusing components annotated with timing information for the construction of a new system will enable the derivation of more accurate system timing
behaviour at early development stages. Therefore the system can be developed with a reduced number of design iterations.

- More efficient communication and collaboration between different parties involved in development: This will support cooperative development scenarios and reduce the risk of mutual misunderstanding between different parties contributing to the design of the same system, for example OEMs and Tier-1 suppliers, and lead to safer and more accurate systems.

- Reduced development risk: A formal and unambiguous foundation for reasoning about time provides a steady basis and a common ground for better cooperation between tools with respect to timing information based on commonly agreed, industry-wide standards like AUTOSAR. The project will further develop methodologies and languages developed in ATESST2 and TIMMO. TADL (Timing Augmented Description Language) and EAST-ADL2 were introduced as a major leap forward and will be further adapted and extended in TIMMO-2-USE.
6. New Results

6.1. Semantics of the Calculus of Inductive Constructions

**Participant:** Bruno Barras [Contact].

Bruno Barras has formalized the meta-theoretical study of strictly positive inductive types. This was built upon the previous work on specific instances: natural numbers and Brouwer ordinals. The main idea of the model construction is to use the property that every strictly positive inductive definition can be encoded in the parameterized type of trees (the so-called W-types). Such tree-types can themselves be encoded as partial functions from paths to labels. The soundness of this translation gives a way to build the closure ordinal of any strictly positive inductive definition.

Bruno Barras has then modelled the inductive families (also called inductive types with indices). He has been able to prove formally the previously known result that inductive family can be constructed in two steps: first build a carrier type (inductively) which is oblivious of the indices, and then define each member of the family as a subset of the carrier type by enforcing the constraints generated by the indices.

He also started to investigate advanced features of inductive definitions, like the possibility to have non-uniform parameters. When this feature was introduced in Coq, it was thought as a conservative one, but the formal analysis showed that this was not obvious. The consistency model could be extended (with one auxiliary result not yet encoded formally). This shows that non-uniform parameters do not extend much the expressivity of Coq, but the strict equivalence remained as an open problem.

6.2. Relative Strengths of set theory and type theory

**Participants:** Bruno Barras [Contact], Benjamin Werner.

Bruno Barras also formalized common translations in proof theory: negated translations and Friedman’s A-translation. This was used to build a model of (classical) ZF set theory in Coq extended with one axiom called TTDA (Type-Theoretical Description Axiom). This was done in two steps: first build a model of IZF (ZF with the collection axiom but not the excluded-middle) in Coq extended with TTDA, and then encode ZF in IZF, as shown by Friedman.

The converse result: an interpretation of Coq +TTDA in ZF (with one inaccessible cardinal!) seems not possible, as TTDA in a classical setting gives a (weaker) form of the axiom of choice. Bruno Barras as devised a new axiom (called the Type-Theoretical Collection Axiom) that still allow the ZF interpretation above, but he hopes that its consistency can be proved in ZF extended with one inaccessible cardinal.

Benjamin Werner has worked with Gyesik Lee on set-theoretical models of Coq’s type theory. This work is described in a paper published in the LMCS journal [18].

6.3. A Consistency model of Coq extended with decision procedures

**Participants:** Bruno Barras [Contact], Qian Wang.

Bruno Barras and Qian Wang are working on the construction of a model for the Calculus of Constructions extended with the type of natural numbers. The definitional equality has been extended to include all equations derivable in Presburger arithmetic. Compared to previous work, this model can support strong eliminations. Since strong eliminations and extensions of the definitional equality with non-satisfiable equations (for instance \(0 = 1\)) leads to non-normalizing terms, it was necessary to give a precise account of Presburger arithmetic, seen as a specific instance of first-order logic. This work is described in a paper published in the proceedings of the LICS conference [21].
6.4. Towards a concurrent architecture for the Coq kernel

Participants: Bruno Barras [Contact], Enrico Tassi.

In the context of the Paral-ITP ANR project, Bruno Barras and Enrico Tassi have started to implement a kernel of Coq where the process of constructing and checking the proof of a lemma can be executed in a parallel thread.

6.5. Physics of computation

Participant: Gilles Dowek.

Together with Pablo Arrighi, Gilles Dowek has extended Gandy’s theorem to quantum physics, by giving a new definition of the notion of finite density of information in this setting. This work has been presented at the congress QIPC [13].

6.6. Binders

Participant: Gilles Dowek.

Together with Jamie Gabbay, Gilles Dowek has given a translation of permissive nominal logic to Higher-order logic and proved its soundness and completeness. This work is described in a paper published in the Transactions on Computational Logic [15].

6.7. Interfacing Coq with SMT solvers

Participants: Germain Faure, Chantal Keller [Contact], Assia Mahboubi, Benjamin Werner.

This is work in close collaboration with the Marelle team (INRIA Sophia Antipolis). The starting point of this work is to note that SMT solvers, deciding the Satisfiability Modulo Theories, are in constant evolution to take into account new decision procedures as well as theories. These systems are rather complex and it is now clearly established that they all contain bugs. The standard approach is to ask the SMT solver to append to the decision result a certificate that can be checked by another tool.

In this context, we are using Coq to check the certificate. The approach is based on computational reflection. The checker is written in Coq, and its architecture is modular and extensible.

We are now able to check certificates coming from the ZChaff SAT solver and from the veriT SMT solver developed at INRIA Nancy – Grand Est. Proofs established by the SMT tool for the theories of congruence closure and linear arithmetic are checked in short time, overtaking the state of the art in terms of time performance. We also use certificates to build a new Coq tactic that can safely call an external SMT solver, thus increasing Coq’s automation. This tactic is new since it is a decision procedure that combines both linear integer arithmetic and equality of uninterpreted functions. This work is described in a paper published in the proceedings of the CPP2011 conference [25].

6.8. SMT techniques for optimization problems

Participant: Germain Faure [Contact].

The TypiCal team has collaborated with the sysmo team at the Laboratoire d’Informatique de l’École Polytechnique in order to integrate the use of automated tools like SMT solvers in the resolution of optimization problems. The case study was a problem of large scale energy management with various constraints, proposed at the ROADEF 2010 challenge won by the sysmo team. We investigated how to delegate to the SMT tool part of the resolution of constraints. A first conclusion of this experiment is that solving optimization problems represents a more important part of the computation time than first expected. As SMT solvers are not geared toward this class of problems, their performance were not satisfactory. This nonetheless opens new perspectives for the development of SMT tools in order to adapt their internal decision procedures to this new kind of benchmarks. We consider that significant progress in that direction could be easily obtained.
6.9. **Formal correctness of embedded programs**  
**Participants:** Gilles Dowek [Contact], Pierre Néron.

Pierre Néron is working on program transformations that remove the operations which create most of the approximations during the computation on floating point numbers, namely square roots and divisions. This kind of formal tool aims at increasing the confidence in embedded programs. The idea of this transformation comes from the elimination of the quantifier on real closed fields, hence the first task is to define a minimal but useful language on which the transformation will apply and then to extend this transformation on formulas to this whole language. Keeping the size of the code produced by this transformation in an acceptable range was a challenging issue in this work. The next objective is to write a formal proof ensuring that the transformation is correct. This work will be done in collaboration with the NASA Langley research center in the Formal Method team: Pierre Néron will visit this center for one month in January 2012.

This work is described in a submitted paper [30].

6.10. **Formal proofs for convex optimization problems**  
**Participants:** Benjamin Werner [Contact], Victor Magron.

Victor Magron is working on the integration of tools that can deal with inequalities on semi-symbolic expressions with real numbers inside proof assistants like Coq. In particular, he is working on new means to provide formally established bounds for multivariate inequalities, using methods inspired from the convex optimization literature like sums of squares (SOS) and the related semi-definite programming (SDP) relaxation.

He has implemented in OCaml a new algorithm which detects and computes automatically the possible bounds of a given expression. He has tested the approach using benchmarks largely built from inequalities issued from the formal proof of Kepler conjecture (by Thomas Hales). The algorithm computes approximation of transcendental functions by solving sum of squares problems, delegated to an external, dedicated tool. The next step of this project is to certify the correctness of these computations using the Coq system.

He has also improved a Coq tactic based on the external computation of decompositions into sums of squares originally developed by Frédéric Besson (INRIA Rennes - Bretagne Atlantique). The improvement consists in linking this tactic with a tool developed by David Monniaux (Verimag).

6.11. **Formal proof in real algebraic geometry**  
**Participants:** Assia Mahboubi [Contact], Cyril Cohen.

Cyril Cohen and Assia Mahboubi have completed the first formal proof of quantifier elimination for the theory of real closed fields. This work includes a significant part of infrastructure code for ordered algebraic theories, intervals, and polynomials. This work is described in a submitted paper [29].

Cyril Cohen has implemented in Coq a construction of real algebraic numbers and proved it had a structure of discrete Archimedian real closed field, in the sense of the previous proof of quantifier elimination. Beside the computational interest of real algebraic numbers, this construction both legitimates the abstraction chosen for the proof of quantifier elimination and provides a basis for complex algebraic numbers needed for the completion of the formal proof of the Feit-Thompson theorem. This work is described in a paper to appear in the proceedings of the JFLA2011 conference.

6.12. **Constructive mathematics**  
**Participant:** Cyril Cohen [Contact].

In collaboration with Thierry Coquand, Cyril Cohen has come up with a constructive proof of a generalization of the fundamental theorem of Algebra. This work show how to formalize the algebraic closure of an arbitrary real closed field. In particular, it can serve as a basis for the construction of complex algebraic numbers from the real algebraic numbers. This work is described in a submitted paper [28].
6.13. Intersection types

Participants: Alexis Bernadet [Contact], Stéphane Lengrand [(CNRS, Lix)].

Alexis Bernadet and Stéphane Lengrand have studied a typing system for the \( \lambda \)-calculus with non-idempotent intersection types. As it is the case in (some) systems with idempotent intersections, a \( \lambda \)-term is typable if and only if it is strongly normalizing. Non-idempotency brings some further information into typing trees, such as a bound on the longest \( \beta \)-reduction sequence reducing a term to its normal form. These results are presented in Klop’s extension of \( \lambda \)-calculus, where the bound that is read in the typing tree of a term is refined into an exact measure of the longest reduction sequence. This complexity result is, for longest reduction sequences, the counterpart of de Carvalho’s result for linear head-reduction sequences. This work is described in a paper published in the proceedings of the FOSSACS 2011 conference [22].

Alexis Bernadet and Stéphane Lengrand have also revisited models of typed \( \lambda \)-calculus based on filters of intersection types. By using non-idempotent intersections, they simplify a methodology that produces modular proofs of strong normalization based on filter models. Non-idempotent intersections provide a decreasing measure proving a key termination property, simpler than the reducibility techniques used with idempotent intersections. Such filter models are shown to be captured by orthogonality techniques: we formalize an abstract notion of orthogonality model inspired by classical realizability, and express a filter model as one of its instances, along with two term-models (one of which captures a now common technique for strong normalization). Applying the above range of model constructions to Curry-style System F describes at different levels of detail how the infinite polymorphism of System F can systematically be reduced to the finite polymorphism of intersection types. This work is described in a paper published in the proceedings of the CSL 2011 conference [23].


Participants: Alexis Bernadet [Contact], Stéphane Lengrand [(CNRS, Lix)].

We introduce here an alternative definition of Hyland’s effective topos, based on a realizability framework with two levels of abstraction: a low level and a high level. With this definition, the proof that this framework forms a topos is almost as simple as proving that the category of sets is a topos. Moreover, the high level of the framework can be directly used as a model of higher-order intuitionistic systems. We can then craft a programming language based on topos theory, which can be given a constructive semantics. In such a programming language, we can only write functions that terminate, as in proof assistants like Coq, so the language cannot be Turing-complete. The main advantage of having a programming language based on topos theory over more usual intuitionistic systems such as Martin-Löf type theory is the notion of equality: it is extensional, has proof-irrelevance, and allows the axiom of unique choice.

This work has been presented at the Chocola-Ens Lyon seminar in December 2011.

6.15. A formal proof of the Feit-Thompson theorem

Participant: Assia Mahboubi [Contact].

Assia Mahboubi has pursued her work in the Mathematical Component team lead by Georges Gonthier at the Microsoft Inria Joint Centre. She has finished the formalization of the Wielandt fixpoint theorem, which is one of the key results at the interface between the two components (local analysis and character theory) of the published revised proof of the Feit-Thompson theorem. The proof of the Wielandt theorem was difficult to formalize because it requires a challenging combination of advanced theories with sophisticated constructive formalization: group representation, module theory, linear algebra and characters.

The documentation of this formalization, as well as the current state of the whole formal proof can be found on the webpage of the Mathematical Components project.

6.16. A formal library for polynomial arithmetics

Participant: Assia Mahboubi [Contact].
Assia Mahboubi has worked on a modular formal library devoted to the divisibility theory of polynomials. The aim of this library is to provide a solid basis for further formal developments involving algorithms on polynomials, an in particular to cover the cases when the coefficients of the polynomials involved are equipped with a structure weaker than the structure of field required by the standard Euclidean algorithm.

The documentation of this formalization can be found on the webpage of the Mathematical Components project.

6.17. Weak Memory Models

Participant: Assia Mahboubi [Contact].

Assia Mahboubi has collaborated with Jade Alglave (Oxford University) and has programmed a complete formalization in Coq of the semantic proposed by Jade Alglave in a PhD for weak memory models. This work is described in [27].
6. New Results

6.1. Models and Verification Techniques

6.1.1. The BCG Format and Libraries

Participants: Hubert Garavel, Frédéric Lang, Wendelin Serwe.

BCG (Binary-Coded Graphs) is both a file format for the representation of explicit graphs and a collection of libraries and programs dealing with this format. Version 1.0 of the BCG format was recently replaced by version 1.1, which can exploit the capabilities of 64-bit addressing.

In 2011, we continued to enhance the BCG libraries as follows:

- We extended the BCG_READ application programming interface with three new primitives so as to increase symmetry with the OPEN/CÆSAR application programming interface (see § 6.1.2 ).
- We fixed a memory corruption problem occurring with very long label strings; this problem would cause random crashes of the DISTRIBUTOR tool (see § 6.1.5 ).

6.1.2. The OPEN/CÆSAR and CÆSAR_SOLVE Libraries

Participants: Iker Bellicot, Hubert Garavel, Yann Genevois, Frédéric Lang, Radu Mateescu, Wendelin Serwe.

OPEN/CÆSAR is an extensible, modular, language-independent software framework for exploring implicit graphs. This key component of CADP is used to build simulation, execution, verification, and test generation tools.

In 2011, a bug in the CÆSAR_TABLE library has been corrected, which would cause segmentation faults when certain primitives of the Application Programming Interface were invoked on a bounded table.

CÆSAR_SOLVE is a generic software library based on OPEN/CÆSAR for solving boolean equation systems of alternation depth 1 (i.e., without mutual recursion between minimal and maximal fixed point equations) on the fly. This library is at the core of several CADP verification tools, namely the equivalence checker BISIMULATOR, the minimization tool REDUCTOR, and the model checkers EVALUATOR 3.5 and 4.0. The resolution method is based on boolean graphs, which provide an intuitive representation of dependencies between boolean variables, and which are handled implicitly, in a way similar to the OPEN/CÆSAR interface [4].

In 2011, we improved the parallel resolution algorithm of CÆSAR_SOLVE (see § 6.1.5 ).

6.1.3. The EVALUATOR Tool

Participants: Iker Bellicot, Hubert Garavel, Yann Genevois, Radu Mateescu.

EVALUATOR is a model checker that evaluates a temporal logic property on a graph represented implicitly using the OPEN/CÆSAR environment. EVALUATOR works on the fly, meaning that only those parts of the implicit graph relevant to verification are explored. The model checking problem is reformulated in terms of solving a boolean equation system. A useful feature of EVALUATOR is the generation of diagnostics (examples and counterexamples) explaining why a formula is true or false.

In version 3.5 of EVALUATOR, properties are described in regular alternation-free µ-calculus, a logic built from boolean operators, possibility and necessity modalities containing regular expressions denoting transition sequences, and fixed point operators without mutual recursion between least and greatest fixed points. The input language of the tool also enables the user to define parameterized temporal operators and to group them into separate libraries.
In version 4.0 of EVALUATOR (5,000 lines of SYNTAX code, 40,500 lines of LOTOS NT code, and 13,100 lines of C code), properties are written in MCL (Model Checking Language) \[18\], an extension of the regular alternation-free \(\mu\)-calculus of EVALUATOR 3.5 with data-handling and fairness operators. In particular, EVALUATOR 4.0 can handle modalities and fixed point operators with data parameters, regular expressions extended with counters, operators inspired from programming languages (“if-then-else”, “for”, etc.), and operators (of alternation depth two) allowing to characterize complex infinite sequences.

In 2011, we continued the extensive testing of EVALUATOR 3.5 and 4.0 using our test base of 10,000 BCG graphs and 3,800 MCL formulas. This revealed three errors in EVALUATOR 4.0, which have been corrected. We also brought the following enhancements to MCL and EVALUATOR 4.0:

- We extended the set of MCL operators of alternation depth two with parameterized versions of the infinite looping and saturation operators, which allow to succinctly encode the presence (respectively, the absence) of accepting cycles in generalized Büchi automata. The evaluation of these parameterized operators is translated into parameterized boolean equation systems, which are instantiated into plain boolean equation systems and solved on the fly using the algorithms A3 and A4 (extended with marked cycle detection) of the CÆSAR SOLVE library. This evaluation procedure has a complexity linear in the size of the degeneralized Büchi automaton, which is represented by the boolean equation system obtained after instantiation.

- We enhanced MCL by adding a data type for manipulating sets of natural numbers. This data type, equipped with the classical set operations (union, intersection, difference, insertion, deletion, membership, etc.), enables a succinct specification of temporal properties referring to the past, such as the fact that a certain set of events (represented by natural numbers) occurred on the transition sequences leading from the initial state to the current state.

- We added a new option to EVALUATOR 4.0 for displaying a set of regular expressions that over-approximate the set of visible actions (transition labels in the LTS) satisfying the action predicates occurring in the MCL formula. This feature enables to improve the efficiency of verification by hiding the set of actions other than those produced from the MCL formula, minimizing the LTS modulo a weak equivalence relation compatible with the formula, and then verifying the MCL formula on the minimized LTS. This may increase the efficiency of verification by one order of magnitude, as reported in \[40\].

EVALUATOR 4.0 was officially integrated in CADP in March 2011. MCL and EVALUATOR 4.0 were used successfully for analyzing mutual exclusion protocols (see § 6.3.1) and hardware architectures (see § 6.3.2).

### 6.1.4. Compositional Verification Tools

**Participants:** Frédéric Lang, Radu Mateescu.

The CADP toolbox contains various tools dedicated to compositional verification, among which EXP.OPEN 2.1, PROJECTOR 3.0, BCG_MIN 2.0, and SVL 2.2 play a central role. EXP.OPEN explores on the fly the graph corresponding to a network of communicating automata (represented as a set of BCG files). PROJECTOR implements behavior abstraction \[58\], \[64\] by taking into account interface constraints. BCG_MIN minimizes behavior graphs modulo strong or branching bisimulation and their stochastic extensions. SVL (Script Verification Language) is both a high level language for expressing complex verification scenarios and a compiler dedicated to this language.

In 2011, we corrected one bug in PROJECTOR, two bugs in EXP.OPEN, and four bugs in SVL. We also enhanced these tools as follows:

- The generalized parallel composition operator proposed in \[13\], including the support for “\(m\) among \(n\)” synchronization, has been added to SVL, leading to version 2.3 of SVL.

- Together with Pepijn Crouzen (Saarland University), we pursued our work on the so-called “smart reduction” techniques for compositional verification. An article about smart reduction was published in an international conference \[32\].
We improved smart reduction for stochastic branching bisimulation so as to cut stochastic transitions (according to the “maximal progress” assumption) as early as possible in intermediate parallel compositions, thus yielding state space reductions. With this new optimization, EXP.OPEN detects when some action (usually, an output) offered by some process can synchronize with corresponding actions (usually, inputs) offered by the other processes in all their states; if so, all stochastic transitions in choice with this action can safely be cut in every intermediate composition. This situation occurs frequently with Input/Output Interactive Markov Chains.

Additionally, we studied an alternative compositional verification approach named partial model checking [47]. Given a temporal logic formula $\varphi$ to be evaluated on a set $S$ of concurrent processes, partial model checking consists in transforming $\varphi$ into another equivalent formula $\varphi'$ to be evaluated on a subset of $S$. Formula $\varphi'$ is constructed incrementally by choosing one process $P$ in $S$ and incorporating into $\varphi$ the behavioral information corresponding to $P$ — an operation called quotenting. Simplifications must be applied at each step, so as to maintain formulas at a tractable size.

We developed a prototype implementation of this approach using the generic software components of CADP:

- We extended the definition of quotenting given by [47] to support all features of the input language of EXP.OPEN 2.1, which enables networks of labeled transition systems to be described using parallel composition operators borrowed from various process algebras such as CCS, CSP, and LOTOS, including also LOTOS NT parallel composition and “$m$ among $n$” synchronisation operators [13].
- We gave an executable definition of quotenting in terms of a synchronous product between a graph representation (called formula graph) of the formula $\varphi$ and the process $P$, thus enabling quotenting to be implemented efficiently in EXP.OPEN. We extended EVALUATOR 3.5 to automatically generate the formula graph corresponding to a temporal logic formula.
- We proposed and implemented efficient formula simplifications by combining reductions modulo bisimulations and partial formula evaluation computed using boolean equation systems.

We used this prototype implementation to verify 28 temporal logic properties on the TFTP avionics protocol [14]. For several of these properties, partial model checking uses hundreds of times less memory than on-the-fly model checking using EVALUATOR. This work led to a publication in an international conference [38].

### 6.1.5. Parallel and Distributed Verification Tools

**Participants:** Iker Bellicot, Hubert Garavel, Rémi Hérilier, Radu Mateescu.

The CADP toolbox contains several components designed to take advantage of distributed computing facilities (such as clusters of machines) to perform large-scale verifications, namely: CÆSAR_NETWORK, a network communication library used by the other tools, DISTRIBUTOR and BCG_MERGE, two companion tools [10] that perform reachability analysis using a distributed state space exploration algorithm, and BES_SOLVE [35], a tool that solves boolean equations systems using the various resolution algorithms provided by the CÆSAR_SOLVE library (see § 6.1.2), including a distributed on-the-fly resolution algorithm.

In 2011, we continued enhancing these tools, taking advantage of the valuable feedback provided by Eric Madeleine (INRIA Sophia Antipolis) who used CADP on the PACAGRID cluster. We brought the following improvements:

- We performed careful code reviews of the CÆSAR_NETWORK library, and corrected nine bugs. We equipped this library with logging primitives that proved to be helpful for debugging distributed algorithms programmed above the CÆSAR_NETWORK library.
- We fixed three bugs in DISTRIBUTOR and two bugs in BCG_MERGE.
- We pursued the intensive testing campaign undertaken in 2010 for BES_SOLVE, using up to 100 concurrent processes running on the PIPOL and GRID5000 platforms. We focused our efforts on the distributed resolution algorithm for boolean equation systems of CÆSAR_SOLVE/BES_SOLVE,
which was tested extensively on examples of boolean equation systems represented explicitly as text files or generated randomly according to various parameters, the resolution results being crosschecked against the sequential resolution algorithms provided by CAESAR\_SOLVE/BES\_SOLVE.

Several bugs (affecting memory leaks, handling of early termination, diagnostic generation, collecting of statistical information about the resolution) were fixed. Changes were also carried out on the code to simplify its structure, increase modularity, and improve readability. The convergence of the distributed resolution algorithm was accelerated by backward propagation of constants as soon as they have been discovered.

- The EVALUATOR 4.0 tool was extended with a new prototype algorithm allowing the distributed verification of an MCL formula on a graph using several machines connected by a network. This functionality was implemented by connecting the tool to the distributed resolution algorithm for boolean equation systems and experimented out on clusters of machines.

- Finally, we added to CADP four new tools named PBG\_CP, PBG\_INFO, PBG\_MV, and PBG\_RM. These tools respectively enable to copy, query, move, and delete the PBG (Partitioned BCG Graph) collection of files generated and used by DISTRIBUTOR and BCG\_MERGE.

### 6.1.6. Other Tool Developments

**Participants:** Hubert Garavel, Yann Genevois, Rémi Hérilier, Frédéric Lang, Radu Mateescu, Wendelin Serwe.

To support the usage of CADP in industry and academia, we pursued our efforts to master the software quality of CADP:

- We added support for Mac OS X 10.7 (“Lion”) and enhanced the documentation for Mac OS X.

- We corrected one bug in the INSTALLATOR installation assistant, two bugs in the TST platform-checking command, and brought two bug fixes and one usability enhancement in the EUCALYPTUS graphical user-interface. We also provided a workaround for supporting recent versions of UBUNTU.

- We continued building a comprehensive validation framework, based on non-regression testing and semantical checking for the CADP tools. This framework allows functional testing of individual tools as well as integration testing for several CADP tools used together to perform complex verification scenarios on various computing platforms and using various compilers.

Other research teams took advantage of the software components provided by CADP (e.g., the BCG and OPEN/CAESAR environments) to build their own research software. We can mention the following developments:

- the KMELIA tools for component-based systems [50], developed at the University of Nantes (France);

- the VERCORS tool for unifying architectural and behavioral specifications of distributed components [52], developed at INRIA Sophia-Antipolis;

- the DAMASCO (Discovery, Adaptation and Monitoring of Context-Aware Services and Components) framework for composition and adaptation based on model transformation [54], developed at the University of Málaga (Spain);

- the SCOOP tool for symbolic optimizations of probabilistic processes [62], [74], developed at RWTH Aachen (Germany);

- the SLCO (Simple Language of Communicating Objects) environment [49], developed at Eindhoven University of Technology (The Netherlands);

- the MOTOR tool for probabilistic analysis of embedded systems [68], developed at the Embedded Systems Institute (Eindhoven, The Netherlands);

- the ALVIS modeling language for design and verification of embedded systems [72], developed at the AGH University of Science and Technology (Krakow, Poland).
6.2. Languages and Compilation Techniques

6.2.1. Compilation of LOTOS

Participants: Hubert Garavel, Wendelin Serwe.

The CADP toolbox contains several tools dedicated to the LOTOS language, namely the CÆSAR.ADT compiler [ 3 ] for the data type part of LOTOS, the CÆSAR compiler [ 11 ] for the process part of LOTOS, and the CÆSAR.INDENT pretty-printer.

In 2011, in addition to fixing four bugs in the CÆSAR and CÆSAR.ADT compilers, we improved the LOTOS-dedicated tools of CADP as follows:

- We revised the predefined type libraries and C code generated by CÆSAR and CÆSAR.ADT to suppress warnings emitted by recent versions of GCC.
- We enhanced the format in which values of singleton and tuple types are displayed to end users.
- We modified the predefined libraries for natural (unsigned) and integer (signed) types so that users can now indicate the precise number of bits (between 1 and 64) to be used for the machine representation of these types, and can also determine the precise range (lower and upper bound) to be used for values of these types.
- We further modified these two libraries to enable (optional) overflow and underflow checking during arithmetic operations on the natural and integer types.

6.2.2. Compilation of LOTOS NT

Participants: Hubert Garavel, Frédéric Lang, Christine McKinty, Vincent Powazny, Wendelin Serwe.

Regarding the LOTOS NT language — a variant of E-LOTOS created by the VASY project team — we worked along two directions:

- We continued enhancing the TRAIAN compiler (see § 5.2 ), which generates C code from LOTOS NT data type and function definitions. TRAIAN is distributed on the Internet and used intensively within the VASY project team as a development tool for compiler construction [ 8 ].

  In 2011, TRAIAN was essentially in maintenance mode. We updated its documentation and added to its predefined library a conversion function that was missing.

- The LNT2LOTOS, LNT.OPEN, and LPP tools convert LOTOS NT code to LOTOS, thus allowing the use of CADP to verify LOTOS NT descriptions. These tools are officially part of CADP since 2010 and have been used successfully for many different systems (see § 6.3.1 , § 6.3.2 , § 6.3.5 , § 6.3.5 , and § 6.3.3 ).

  In 2011, we continued enhancing these tools, of which we delivered four new releases. In addition to 13 bug fixes, the following enhancements have been brought:

    - The LOTOS NT language was extended with range types (which are interval subtypes of character, integer, or natural types) and predicate types (which are subtypes of existing types, a boolean predicate being used to determine the domain of each subtype).
    - The LOTOS NT language was enriched with the concept of “module pragmas”, which specify implementation constraints for predefined types such as naturals, integers, and strings. Also, the predefined operations “first” and “last” have been added for enumerated types.
    - The LNT2LOTOS translator was made semantically stricter by adding checks for overflow and underflow when doing natural and integer arithmetics, checking that range type bounds and array type bounds belong to the domain of admissible values for their parent types, and adding additional checks for type pragmas.
– The LOTOS code generated by the LNT2LOTOS translator was optimized by handling equation premises (in the data part) and boolean guards (in the behavior part) that are always false or always true. Other optimizations have been added for process definitions whose bodies are empty or only contain a call to another process, for “case” statements that are followed by no instruction or only a “null” instruction, and for “while” loops with an empty body.

– The speed of processing LOTOS NT specifications containing several modules has been made between two and three times faster.

– The error and warning messages issued by the LOTOS NT tools have been enhanced.

– The reference manual has been corrected, reorganized and comprehensively edited. Two new appendices have been added, one that lists all the predefined functions, and another one (20 pages) giving the formal semantics of LOTOS NT.

6.2.3. Source-Level Translations between Concurrent Languages

Participants: Hubert Garavel, Rémi Hérilier, Frédéric Lang, Radu Mateescu, Gwen Salaün, Wendelin Serwe, Damien Thivolle.

Although process algebras are, from a technical point of view, the best formalism to describe concurrent systems, they are not used as widely as they could be [2]. Besides the steep learning curve of process algebras, which is traditionally mentioned as the main reason for this situation, it seems also that the process algebra community scattered its efforts by developing too many languages, similar in concept but incompatible in practice. Even the advent of two international standards, such as LOTOS (in 1989) and E-LOTOS (in 2001), did not remedy this fragmentation. To address this problem, we started investigating source-level translators from various process algebras into LOTOS or LOTOS NT, so as to widen the applicability of the CADP tools.

In 2011, in addition to the LNT.OPEN tool suite (see § 6.2.2), we worked on the following translators:

• We continued our work on the FLAC tool, which translates a FIACRE program into a LOTOS program automatically, for verification using CADP. In 2011, 2 bugs reported by users of FLAC were corrected. Those corrections led to revisions 74 and 75 of the FLAC code, which is available on the development forge dedicated to FIACRE compilers [1]. We collected new examples of FIACRE code to enhance our test suite, which now comprises 79 examples.

• BPEL (Business Process Execution Language) [61] is a language inspired by the π-calculus [67] and standardized by the OASIS consortium (led by IBM and Microsoft) to describe the orchestrations of Web services. BPEL depends on other W3C standard XML-related languages: XML Schema for data types, XPath for data expressions, and WSDL for declaring the interfaces (communication links and link functions) of a Web service.

Following interest expressed by research teams at MIT and the Polytechnic University of Bucharest, we designed translation rules from BPEL to LOTOS NT in order to formally verify BPEL services with CADP. We began to develop an automated translator.

In 2011, following a remark by Charles Pecheur (Université Catholique de Louvain, Belgium) who spotted an error in the translation of BPEL processes into LOTOS NT, we corrected the translation of exception handling so that it no longer interferes with the atomicity mechanism. The complete translation algorithm is given in Damien Thivolle’s PhD thesis [23]. We pursued the implementation of our BPEL to LOTOS NT translator and finalized the translation of XML Schema types and WSDL definitions.

• We considered the π-calculus [67], a process algebra based on mobile communication. We proposed a general method for translating the finite control fragment of the π-calculus (obtained by forbidding recursive invocations of an agent through parallel composition operators) into LOTOS NT. The mobile communication is encoded using the data types of LOTOS NT, each channel name

http://gforge.enseeiht.fr/projects/fiacre-compil
being represented as a value of an enumerated data type. The binary synchronization of \( \pi \)-calculus is enforced by associating a LOTOS NT gate to each parallel composition operator present in the \( \pi \)-calculus specification and by tagging each synchronization with the unique identifiers of the sender and receiver agents. The translation preserves the operational semantics by mapping each transition of a \( \pi \)-calculus agent to a single transition of the resulting LOTOS NT term.

In 2011, we have extended the \( \pi \)-calculus with data-handling features, with the goal of widening its possible application domains. This was done by extending the language grammar and the translation to support typed variables and data expressions. As language for describing data, we chose LOTOS NT: indeed, the data types and functions used in the \( \pi \)-calculus specification can be described in LOTOS NT and directly incorporated to the LOTOS NT code produced by translation. This results in an applied \( \pi \)-calculus, such as the variant of the calculus proposed in [45] for the verification of security properties.

The \textsc{Pic2Lnt} translator was extended accordingly. It now consists of 2,100 lines of \textsc{Syntax} code, 3,100 lines of LOTOS NT code, and 500 lines of C code. The tool was tested on 234 examples of \( \pi \)-calculus specifications, including most of the examples provided in the Mobility Workbench distribution.

### 6.3. Case Studies and Practical Applications

#### 6.3.1. Mutual Exclusion Protocols

**Participants:** Radu Mateescu, Wendelin Serwe.

Mutual exclusion protocols are an essential building block of concurrent systems to ensure proper use of shared resources in the presence of concurrent accesses. Many variants of mutual exclusion protocols exist for shared memory, such as Peterson’s or Dekker’s well-known protocols. Although the functional correctness of these protocols has been studied extensively, relatively little attention has been paid to their performance aspects.

In 2011, we considered a set of 27 mutual exclusion protocols for up to sixteen processes with a shared memory and coherent local caches. We specified each protocol in LOTOS NT, using a set of generic modules to describe shared variables, the cache protocol, and the overall architectures (in total, 13,600 lines of LOTOS NT code). Then, we compositionally added Markov delays modeling the latencies of read/write accesses on shared variables, so as to obtain the Interactive Markov Chain (IMC) corresponding to each protocol (up to 1.6 billion states and 2.7 billion transitions for the black-white bakery protocol [73] for four processes).

We verified functional properties using the same set of MCL [18] formulas for each protocol (in total, 380 lines of MCL). The mutual exclusion property was easy to express as an MCL formula, but other properties (livelock and starvation freedom, independent progress, and unbounded overtaking) turned out to be quite involved because they belong to the \( \mu \)-calculus fragment of alternation depth two; fortunately, we succeeded in expressing them using the infinite looping operator of MCL, which can be checked in linear time. In particular, it was challenging to express these properties using a parameter \( N \) for the number of processes.

Finally, using the performance evaluation tools of CADP (i.e., BCG\_STEADY for small numbers of processes and CUNCTATOR for larger numbers of processes), we computed the steady-state throughputs of critical section accesses by varying several parameters (relative speeds of processes, ratio between the time spent in critical and non-critical sections, etc.).

These experiments enabled us to compare the protocols according to their efficiency (steady-state throughputs) and study also their scalability for an increasing number of processors. We observed that symmetric protocols are more robust when the difference in execution speed between processes is large, which confirms the importance of the symmetry requirement originally formulated by Dijkstra [56]. The quantitative results corroborated those of functional verification: the presence of (asymmetric) starvation of processes, detected using temporal formulas, was clearly reflected in their steady-state throughputs. Our results also corroborate experimental measures found in the literature [48].
6.3.2. The Platform 2012 Architecture

**Participant:** Wendelin Serwe.

In the context of the MULTIVAL contract (see § 7.1), STMicroelectronics studied PLATFORM 2012, a many-core programmable multi-cluster platform fabric targets a range of emerging video, imaging, and next-generation immersive multimodal applications. Configurability options include the number of clusters, the number and type of processing elements (PE) per cluster, specialization of the architecture and instruction-set of the PEs, and finally, support of hardware-accelerated PEs. The platform is supported by a rich programming environment which embodies a range of platform programming models.

In 2011 we focused on the DTD (Dynamic Task Dispatcher) hardware block that assigns a set of application tasks on a set of PEs. It is called dynamic because each task itself might add tasks to the set of those to be dispatched by the DTD. The DTD is synthesized from a C++ model, optimized to generate an efficient hardware block. Due to the intrinsic complexity of the DTD, STMicroelectronics was interested in the co-simulation of this C++ code with a formal model of the DTD.

In a first step, we generalized the LOTOS NT model of the DTD developed in 2010 to allow the handling of an arbitrary number of PEs (1, 200 lines of LOTOS NT). We also modeled as LOTOS NT processes the different sets of tasks corresponding to various applications. To express the operations provided by the DTD, we had to include a call-stack in the model of each PE, as a means of circumventing the static-control constraints of CESAR forbidding recursion over parallel composition. STMicroelectronics judged LOTOS NT to be essential in modeling the DTD, because using LOTOS instead would have been extremely difficult, requiring complex continuations with numerous parameters. We also wrote twelve scenarios (1, 000 lines of LOTOS NT) describing applications to be dispatched by the DTD. For each scenario and for up to six PEs, we generated the corresponding LTS (up to 100 million states and 500 million transitions).

Then, for each generated LTS, we verified several properties, such as the correctness of assertions inserted in the model (by checking the set of transition labels), the termination of the scenario, or that each task is executed exactly once. We expressed the latter properties using the MCL language [18] and verified them using the EVALUATOR 4 model checker. This allowed us to point out a difference between our implementation and the one from the architect, highlighting a high sensibility on the order of terms in an equation, revealing an under-specified mechanism. We also verified the correctness of a complex optimization.

Having gained confidence in the LOTOS NT model, we applied the EXEC/CESAR framework to co-simulate the C++ and LOTOS NT models of the DTD, a challenge being the connection of the asynchronous LOTOS NT model with the synchronous C++ model, because one step of the C++ model corresponds, in general, to several transitions of the LOTOS NT model.

This case study enabled us to discover and correct a few bugs in CADP and led to a publication in an international conference [39].

6.3.3. The Self-configuration Protocol

**Participant:** Gwen Salaün.

Cloud computing emerged a few years ago as a major topic in modern programming. It leverages hosting platforms based on virtualization, and promises to deliver resources and applications that are faster and cheaper with a new software licensing and billing model based on the pay-per-use concept.

Distributed applications in the cloud are composed of a set of virtual machines (VMs) running a set of interconnected software components. However, the task of configuring distributed applications is complex as each VM includes many parameters either for local configuration (e.g., pool size, authentication data) or remote interconnection (e.g., IP address and port to access a server). Existing deployment solutions are often specific to certain applications and rarely take into account these configuration parameters, which are usually managed by dedicated scripts that do not work fully automatically.
Together with Xavier Etchevers, Thierry Coupaye (Orange labs), Fabienne Boyer, and Noël de Palma (INRIA Grenoble), we worked on the verification of an innovative self-configuration protocol [34] that automates the configuration of distributed applications in the cloud without requiring any centralized server nor a scripting effort. The high degree of parallelism involved in this protocol making its design difficult and error-prone, we decided to specify the protocol using LOTOS NT and to verify it with CADP. So doing, we detected a major bug, which was corrected in the reference Java implementation. The LOTOS NT specification also served as a workbench to experiment with several possible communication models, which helped us to avoid an erroneous design.

These results have been published in [44].

6.3.4. Realizability of Choreographies

Participants: Matthias Güdemann, Gwen Salaün.

The specification and the analysis of interactions among distributed components play an important role in service-oriented computing. In order to facilitate the integration of independently developed components (named peers) that may reside in different organizations, it is necessary to provide a global contract that the peers participating in a service composition should adhere to. Such a contract is called choreography. One important problem in a top-down development process is figuring out whether a choreography specification can be implemented by a set of peers that communicate via message passing. Given a choreography specification, it would be desirable to generate peers automatically by projecting the global choreography specification to each peer ignoring all messages that are not sent or received by that peer. However, generation of peers that precisely implement a choreography specification is not always possible, i.e., there are choreographies that are not implementable by a set of distributed peers. This problem is known as realizability.

In 2011, we considered the following aspects of the realizability problem:

- In collaboration with Gregor Gössler (INRIA Grenoble) we studied the realizability of choreographies for peers interacting asynchronously through message buffers. Although this problem is generally undecidable for unbounded buffers, we proposed techniques to check whether peers interacting asynchronously with finite buffers can realize a choreography, and if so, for which buffer sizes. These results have been published in [36].
- In collaboration with Pascal Poizat (LRI, Orsay), we proposed an approach to check the realizability of choreographies using the interaction model of BPMN (Business Process Modeling Notation) 2.0. While being a standard for the abstract specification of business workflows and collaboration between services, BPMN has only been recently extended into BPMN 2.0 to support choreographies. Our approach is based on a model transformation into LOTOS NT and the use of equivalence checking. We implemented a prototype of our approach using the EclipseBPMN 2.0 editor and CADP. These results have been published in [43].
- In collaboration with Meriem Ouederni (LINA, Nantes), we studied the automatic synthesis of monitors to enforce realizability, using CADP to check equivalence between the choreography and an automatically obtained distributed implementation.

6.3.5. Other Case Studies

Participants: Hubert Garavel, Frédéric Lang, Radu Mateescu, Gwen Salaün, Wendelin Serwe, Damien Thivolle.

- In the context of the TOPCASED project (see § 7.2), we studied how CADP can be used to verify avionics protocols. In 2011, we prepared two lectures summarizing our prior results on four avionic protocols, namely a ground/plane communication protocol based on TFTP (Trivial File Transfer Protocol) [14], the BITE (Built In Test Equipment)/CMS (Central Maintenance Function), the ATC (Air Traffic Control) system, and the AEN (Air Traffic System Facilities Notification).
- Our prior work (2009–2010) with Fabienne Boyer and Olivier Gruber (Université Joseph Fourier Grenoble) on modeling and verification using LOTOS NT and CADP of the SYNERGY reconfiguration protocol led to a publication in an international conference [31].
Other teams also used the CADP toolbox for various case studies. To cite only recent work not already described in previous VASY activity reports, we can mention:

- behavior analysis of malware by rewriting-based abstraction [51];
- safety verification of fault-tolerant distributed components [46];
- verification of mobile ad hoc networks [57];
- model checking ERLANG applications [59];
- model-checking dataflow in service compositions [63];
- verification of a key chain based TTP transparent CEM protocol [65];
- semantics tuning of UML/SYSML [69];
- atomicity maintenance in EPCREPORT of ALE [70];
- cost analysis of semantic composability validation [71];
- rigorous development of prompting dialogs [75];
- scalably verifiable cache coherence [76].
6. New Results

6.1. Algebraic methods for geometric problems

6.1.1. New bivariate system solver and topology of algebraic curves

We present in [22] a new approach for solving polynomial systems of two bivariate polynomials with rational coefficients. The tools used in our algorithm are classical (subresultants, Groebner basis, triangular systems, regular chains, RUR (rational univariate representations), modular computation) but they are combined in a new way. We first use a classical approach based on subresultant sequences for decomposing a system into subsystems according to the number of roots (counted with multiplicities) in vertical lines. We then show how the resulting triangular subsystems can be efficiently solved by computing lexicographic Gröbner bases and Rational Univariate Representations (RURs) of these systems. We eventually show how this approach can be performed using modular arithmetic, while remaining deterministic, yielding an algorithm that can take advantage of a parallel implementation. We apply our solver to the problem of computing the topology of algebraic curves using the algorithm Isotop [31]. We show that, on generic curves, our algorithm performs similarly as classical resultant-based algorithms and, on non-generic curves, it performs significantly better than all non-GPU based implementations (it outperforms the curve arrangement of CGAL with factors up to several hundreds). Preliminary experiments also hint that the recent GPU-based approach of Berberich et al. [29] and the multi-thread version of our implementation perform similarly on a standard machine, although our implementation naturally depends on the number of threads.

We also started to work on a generalization of these results to real algebraic curves embedded in dimension 3 and higher.

6.1.2. Counting the number of embeddings of a given rigid graph

We addressed the problem of counting the number of embeddings of a given rigid graph (a graph with edges labeled by distances). Such graphs are still not well understood and appear in several applications such as robot kinematics and structural biology. By modeling the problem as a sparse system of polynomial equations, we could bound the maximal number of embeddings from above with the algebraic mixed volume theory, and bound it from below with stochastic optimization methods. This work submitted in 2010 was accepted and published in the proceedings of the IFToMM 2011 conference [21].

6.1.3. Description of singularities of a parameterized mechanism

Kinematic design can be seen as an application of rigid graph theory. In collaboration with the IRCCyN laboratory, we worked on the design of parallel mechanisms. We studied in particular the cable robots, a new kind of architecture, which is difficult to understand. The problem is to describe the set of parameters such that the robot doesn’t break or lose control. Using tools from algebra, we can describe rigorously the working space of a simple planar cable robot [19]. Work on this subject is promising, and some theory developed for rigid graphs could give other interesting results in kinematic design.

6.1.4. Distance between 3-dimensional terrains

We addressed the problem of computing efficiently the distance between two piecewise-linear bivariate functions \( f \) and \( g \) defined over a common domain \( M \). We focus on the distance induced by the \( L_2 \)-norm, that is \( \| f - g \|_2 = \sqrt{\int_M (f - g)^2} \). If \( f \) is defined by linear interpolation over a triangulation of \( M \) with \( n \) triangles, while \( g \) is defined over another such triangulation, the obvious naïve algorithm requires \( \Theta(n^2) \) arithmetic operations to compute this distance. We show that it is possible to compute it in \( O(n \log (n^3)) \) arithmetic operations, by reducing the algebraic problem to multi-point evaluation of a certain type of polynomials [24].
6.1.5. Invariant-based predicate evaluation strategies

We have worked on formalizing polynomial evaluation strategies of geometric predicates using algebraic invariant theory. Let $\mathcal{P}$ be a typical predicate that one encounters in (non-linear) computational geometry. The general approach has three main steps:

1. Identify the symmetries of the problem, i.e. the transformations on the entries $X$ of $\mathcal{P}$ that leave invariant the output of the predicate. These transformations can be modeled by the action $\psi$ of a group $G$ operating on $X$.

2. Use appropriate techniques or known theorems to obtain polynomial invariants for $\psi$. In particular, we have investigated the use of an effective invariant construction method due to Grosshans et al. based on a symbolic representation of invariants.

3. Build a polynomial evaluation strategy for $\mathcal{P}$ by determining those orbits of $\psi$ that are discriminated by the invariants obtained above.

We have applied this general approach to two problems: determining the number of real lines piercing four given lines and determining when two quadrics have no real point in common. For the first problem we essentially reproduce the results obtained previously by Devillers et al. through a direct manipulation of equations. For the second we only have partial results so far.

This work is part of Guillaume Batog’s PhD thesis (defended in December 2011) [12].

6.2. 3D visibility, theory and applications

6.2.1. Calibration for linear cameras

The linear camera is a fairly general geometric model of imaging devices proposed by Jean Ponce and based on linear line congruences, two-dimensional sections of the Klein quadric (a classical model for the space of lines). In a previous work, in collaboration with Jean Ponce, we explored properties of this model. We established in particular the equivalence of linear camera with another model proposed by Tomas Pajdla (which we generalized along the way). The complementarity of these models allowed to extend standard computer vision techniques, such as stereo-reconstruction, to any imaging system modeled by a linear camera. We went one step further and explored how the notion of “calibration” extends to linear cameras. We enriched the usual “intrinsic” parameters used for central cameras by additional parameters that encode the geometry of linear line congruences. This required to investigate the Euclidean aspects of linear line congruences, objects that are usually studied in a projective setting.

This work is part of Guillaume Batog’s PhD thesis (defended in December 2011) [12] and a journal version is in preparation.

6.3. Discrete and computational geometry

6.3.1. On Point-sets that Support Planar Graphs

A set of points is said universal if it supports a crossing-free drawing of any planar graph. For a planar graph with $n$ vertices, if bends on edges of the drawing are permitted, universal point-sets of size $n$ are known, but only if the bend-points are in arbitrary positions. If the locations of the bend-points must also be specified as part of the point-set, no result was known, and we prove that any planar graph with $n$ vertices can be drawn on a universal set $\mathcal{S}$ of $O(n^2 / \log n)$ points with at most one bend per edge and with the vertices and the bend points in $\mathcal{S}$. If two bends per edge are allowed, we show that $O(n \log n)$ points are sufficient, and if three bends per edge are allowed, $\Theta(n)$ points are sufficient. When no bends on edges are permitted, no universal point-set of size $o(n^2)$ is known for the class of planar graphs. We show that a set of $n$ points in balanced biconvex position supports the class of maximum-degree-3 series-parallel lattices [20].
6.3.2. Helly numbers of acyclic families

The nerve of a family of sets is a simplicial complex that records the intersection pattern of its subfamilies. Nerves are widely used in computational geometry and topology, because the nerve theorem guarantees that the nerve of a family of geometric objects has the same topology as the union of the objects, if they form a good cover.

We relaxed the good cover assumption to the case where each subfamily intersects in a disjoint union of possibly several homology cells, and we proved a generalization of the nerve theorem in this framework, using spectral sequences from algebraic topology. We then deduced a new topological Helly-type theorem that unifies previous results of Amenta, Kalai and Meshulam, and Matoušek. This Helly-type theorem is applied to (re)prove, in a unified way, bounds on Helly numbers of sets of lines in geometric transversal theory [25].

6.4. National Initiatives

6.4.1. ANR Blanc - PRESAGE

This project brings together computational geometers (from the VEGAS and GEOMETRICA projects of INRIA) and probabilistic geometers (from Universities of Rouen, Orléans and Poitiers) to tackle new probabilistic geometry problems arising from the design and analysis of geometric algorithms and data structures. We focus on properties of discrete structures induced by or underlying random continuous geometric objects.

This is a four year project, with a total budget of 400kE, that will start on Dec. 31st, 2011. It is coordinated by X. Goaoc (VEGAS).

6.4.2. PEPS Rupture - INS2I: Manifold

The aim of this project is to initiate a collaboration to investigate algebraic/numeric methods for the analysis of manifolds, considering also singularities, that arise in robotics and biological models. Researchers specialized in interval analysis (LINA) and symbolic methods (LORIA) evaluate the relevance of their approaches to applications in robotics (IRCCyN) and biology (LINA). The outcome of this evaluation will be a proposal for a hybridization of both methods that will be worked out in a longer term project.

This is a one year project with a budget of 10 kE. Two one day workshops have been funded with invitations of potential partners for a followup.

6.5. International Initiatives

6.5.1. Visits of International Scientists

- William J. Lenhart, Williams College (USA), one year (sabbatical) from September 2011.
- Andreas Holmsen, KAIST (South Korea), June, 1 week.
- Martin Tancer, Charles University (Prague), April 1 week.
- Pavel Patak, Charles University (Prague), April 1 week.
- Zuzana Safernova, Charles University (Prague), April 1 week.
- Jinsan Cheng, Chinese academy of science (Beijing), November, 1 week.
- Luis Peñaranda, University of Athens, December, 1 week.

6.5.2. Participation In International Programs

- Sylvain Petitjean collaborates with Pr. Gert Vegter of the University of Groningen on “Certified Geometric Approximation”. This collaboration is funded by the Netherlands Organization for Scientific Research (NWO) - 2008–2012.
6. New Results

6.1. Using symmetries in SMT

Participants: David Déharbe, Pascal Fontaine, Bruno Woltzenlogel Paleo.

Methods exploiting problem symmetries have been very successful in several areas including constraint programming and SAT solving. We propose a similar technique for enhancing the performance of SMT-solvers by detecting symmetries in the input formulas and using them to prune the search space of the SMT algorithm. This technique is based on the concept of (syntactic) invariance by permutation of constants. An algorithm for solving SMT by taking advantage of such symmetries is presented. The implementation of this algorithm in the SMT-solver veriT results in an impressive improvement of veriT's performances on the SMT-LIB benchmarks that places it ahead of the winners of the last editions of the SMT-COMP contest in the QF_UF category.

This technique has immediately been adopted by the SMT community. For instance, we are aware that Z3 (Microsoft) and CVC3 (University of New-York and University of Iowa) implemented this technique for the 2011 competition.

6.2. Compression of SMT proofs

Participants: Pascal Fontaine, Stephan Merz, Bruno Woltzenlogel Paleo.

Integrating an SMT solver in a certified environment such as an LF-style proof assistant requires the solver to output proofs. Unfortunately, those proofs may be quite large, and the overhead of rechecking the proof may account for a significant fraction of the proof time. In previous work, we proposed a technique for reducing the sizes of propositional proofs based on the analysis of resolution graphs, which were justified in an algebra of resolution. Unfortunately, the complexity of these techniques turned out to be prohibitive. In a paper published at CADE 2011 [11], we give practical algorithms for more restricted compression techniques and validate them on standard benchmarks. Our algorithms significantly improve state-of-the-art proof compression algorithms and achieve better reduction of proof sizes, often by 30%.

6.3. Combination of decision procedures

Participant: Pascal Fontaine.

We investigate the theoretical limits of combining decision procedures and reasoners, as these are important for the development of the veriT solver (see section 5.1). It has long been known that it is possible to extend any decidable language (subject to a minor requirement on cardinalities) with predicates described by a Bernays-Schönfinkel-Ramsey theory (BSR). A formula belongs to the BSR decidable fragment if it is a conjunction of universal, function-free formulas. As a consequence of this theoretical result, it is possible to extend a decidable quantifier-free language with sets and set operators, relations, orders and similar concepts. This can be used to significantly extend the expressivity of SMT solvers. In previous work, we had generalized this result to the decidable first-order class of monadic predicate logic, and to the two-variable fragment. In 2011, in cooperation with Carlos Areces from Universidad Nacional de Córdoba, Argentina, we showed that two other important decidable fragments (namely the Ackermann fragment, and several guarded fragments) are also easily combinable. This result was presented at the FroCoS Conference 2011 [8], as well as at the SMT’2011 workshop (joint with the Conference on Computer Aided Verification, CAV 2011).

6.4. Encoding TLA+ proof obligations for SMT solvers

Participants: Stephan Merz, Hernán-Pablo Vanzetto.
The TLA\(^+\) proof system TLAPS (see 5.2 ) is being developed within a project at the MSR-INRIA Joint Centre in which we participate. The original release of TLAPS contained an SMT backend that handled quantifier-free proof obligations in linear arithmetic and that was occasionally useful, given that the other backends perform quite poorly on formulas involving arithmetic. However, TLA\(^+\) proof obligations usually mix arithmetic with other theories, in particular set theory, functions, records, and tuples. We propose a new encoding of TLA\(^+\) sequents in SMT-LIB, the generic input language of SMT solvers. The main challenge has been to design a sound translation from untyped TLA\(^+\) to the multi-sorted first-order logic that underlies SMT-LIB. We have developed a type system and a type inference algorithm that assigns SMT-LIB sorts to symbols and terms in the input formula, based on “typing assumptions” among the hypotheses present in the proof obligation.

The translation has been validated over several existing examples, yielding significant reductions in proof sizes. For example, the new backend can automatically verify the main invariant of a parameterized version of the Bakery algorithm, which previously required a few hundred lines of interactive proof. Similarly, an existing proof about a security architecture \footnote{33} has been reduced by about 90\%. The backend has been integrated in TLAPS and has been presented at a workshop \footnote{19}.

6.5. \textbf{Model checking within SimGrid}

\textbf{Participants:} Stephan Merz, Martin Quinson [of project team AlGorille], Cristián Rosa.

For several years we have cooperated with Martin Quinson from the AlGorille project team on adding model checking capabilities to the simulation platform SimGrid for message-passing distributed C programs. The expected benefit of such an integration is that programmers can complement simulation runs by exhaustive state space exploration in order to detect errors such as race conditions that would be hard to reproduce by testing. Indeed, a simulation platform provides a controlled execution environment that mediates interactions between processes, and between processes and the environment, and thus provides the basic functionality for implementing a model checker. The principal challenge is the state explosion problem, as a naive approach to the systematic generation of all possible process interleavings would be infeasible beyond the most trivial programs. Moreover, it is impractical to store the set of global system states that have already been visited: the programs under analysis are arbitrary C programs with full access to the heap, making it difficult and costly to store global states and to determine if two states are equal.

We have implemented a stateless model checker within the SimGrid platform, for verifying safety properties of distributed C programs that communicate by message passing. The visible actions correspond to the communication events, at which points programs can be interrupted by the simulation core. In order to mitigate state explosion, the exploration relies on Dynamic Partial-Order Reduction (DPOR) that avoids exploring redundant interleavings corresponding to the same global happens-before relation. We have identified four primitive communication actions, in terms of which the different message-passing libraries provided by SimGrid can be implemented, and have proved independence theorems for these primitives that underly our DPOR exploration algorithm. We thus obtain a small kernel that supports different communication APIs; nevertheless, practical evaluations yield similar reductions as those obtained by Li et al. \footnote{30} for a much more detailed analysis of a fragment of the MPI library.

The model checker SimGridMC is now part of the SimGrid platform and allows programmers to either perform simulation or model checking runs based on the same source code. It has allowed us to discover a non-trivial bug in an implementation of the Chord algorithm for realizing a distributed hashtable over a P2P network. A conference paper has been published at FORTE 2011 \footnote{13}. Cristián Rosa successfully defended his PhD thesis \footnote{7} in October 2011, which also proposes efficient techniques for parallelizing simulation runs in SimGrid. Marion Guthmuller has explored extensions of our model checking algorithm for verifying liveness properties, and has started working on her PhD thesis in this area in the fall of 2011.

6.6. \textbf{A new version of PlusCal}

\textbf{Participants:} Sabina Akhtar, Stephan Merz, Martin Quinson [of project team AlGorille].
In cooperation with Martin Quinson of the AlGorille team of INRIA Nancy we have defined and implemented a high-level language for the description of concurrent and distributed algorithms. Our work is inspired by Lamport’s PlusCal [29], but extends it for the modeling and verification of distributed algorithms. In particular, processes can be nested and variables are properly scoped; this is useful for modeling concurrent execution at different levels of a hierarchy (such as threads versus processes).

In 2011, the main effort has gone into designing partial-order reduction techniques for model checking PlusCal algorithms, which exploit the locality information present in the models. In particular, we have defined predicates that ensure the independence of two (blocks of) statements and adapted the TLC model checker to implement static partial-order reduction. Sabina Akhtar prepares her PhD thesis manuscript, and the thesis defense is planned for spring 2012.

6.7. Verification of distributed algorithms in the Heard-Of model

Participants: Henri Debrat, Stephan Merz.

Distributed algorithms are often quite subtle, both in the way they operate and in the assumptions required for their correctness. Formal models are important for unambiguously understanding the hypotheses and the properties of a distributed algorithm. We focus on the verification of round-based algorithms for fault-tolerant distributed systems expressed in the Heard-Of model of Charron-Bost and Schiper [26], for which we had already proved a reduction theorem in previous work.

In 2011, we have extended our previous results to the case of Byzantine errors where values may be received that do not correspond to those that should have been computed by the sender process (for example because of an intermittent fault in the sender process, a malicious process, or a value-changing error in the transmission channel). We have formalized a corresponding extension of the Heard-Of model in Isabelle/HOL, and have verified three Byzantine Consensus algorithms (EIG, ATE and UTE) within this framework. These results have been presented at SSS 2011 [9].

6.8. Modeling and verifying the Pastry routing protocol

Participants: Tianxiang Lu, Stephan Merz.

As a significant case study for the techniques that we are developing within VeriDis, we are modeling and verifying the routing protocol of the Pastry algorithm [25] for maintaining a distributed hash table in a peer-to-peer network. As part of his PhD work (under the joint supervision of Stephan Merz and Christoph Weidenbach from MPI-INF Saarbrücken), Tianxiang Lu has developed a TLA+ model of the Pastry routing protocol, which has uncovered several issues in the existing presentations of the protocol in the literature, and in particular a loophole in the join protocol that had been fixed by the algorithm designers in a technical report that appeared after the publication of the original protocol.

In 2011, we have worked towards a correctness proof of the routing protocol. We have in particular identified a number of candidate invariants that have been validated by extensive model checking over finite instances and for which we have formally proved that their validity would imply the correctness of the protocol. Our proofs are carried out in TLAPS (section 5.2 ) and represent a sizable case study for the different proof tools of the proof system. Our results have been presented at FORTE 2011 [12].

6.9. Incremental development of distributed algorithms

Participants: Dominique Méry, Manamiary Andriamiarina.

The development of distributed algorithms and, more generally, of distributed systems, is a complex, delicate, and challenging process. The approach based on refinement helps to gain formality by using a proof assistant, and proposes to apply a design methodology that starts from the most abstract model and leads, in an incremental way, to the most concrete model, for producing a distributed solution. Our works help to formalize pre-existing algorithms, develop new algorithms, as well as develop models for distributed systems.
Our research, carried out with Mohammed Mosbah and Mohammed Tounsi from the LABRI laboratory, was supported by the ANR project RIMEL until 2010 and we are maintaining a joint project B2VISIDIA with LABRI on these topics. More concretely, we aim at an integration of the correct-by-construction refinement-based approach into the local computation programming model. The team of LABRI develops an environment called VISIDIA that provides a toolset for developing distributed algorithms expressed as a set of rewriting rules of graph structures. The simulation of rewriting rules is based on synchronization algorithms and we have developed these algorithms by refinement.

Synchronization algorithms [14] are mandatory for simulating local computation models of distributed algorithms. Therefore, correctness of these algorithms becomes crucial, because it gives confidence that local computations are simulated as designed and do not behave harmfully. However, these algorithms are often very complex to prove correct since they integrate both distributed and probabilistic aspects. We derive proofs of synchronization algorithms upon which the correct-by-construction paradigm depends; the latter is supported by a progressive and incremental process controlled by the refinement techniques. We illustrate our approach by examples such as the Handshake and the LC1 algorithms. These algorithms are designed for an asynchronous distributed network of anonymous processes that communicate by message passing.

A second contribution is related to the integration of probabilistic arguments when reasoning about the design of distributed programs. We particularly focus [20] on probabilistic aspects of distributed algorithms related to termination, e.g. the choice between two delays in the case of communication protocols like IEEE 1394 (FireWire), or the choice between several colors for vertex coloring algorithms. We have in particular applied this approach to developing probabilistic distributed graph coloring algorithms (also called vertex coloring algorithms), based on an algorithm developed by Métivier et al. [32], using the Event B and probabilistic Event B methods.

A third contribution takes into account the modification of links between nodes in a graph modelling a network. We present [15] an incremental formal development of the Dynamic Source Routing (DSR) protocol in Event-B. DSR is a reactive routing protocol, which finds a route for a destination on demand, whenever communication is needed. Route discovery is an important task of any routing algorithm and its formal specification is a challenging problem in itself. The specification is performed in a stepwise manner by introducing more advanced routing components between the abstract specification and topology. It is verified through a series of refinements. The specification includes safety properties as a set of invariants, and liveness properties that characterize when the system reaches stable states. We establish these properties by proof of invariants, event refinement and deadlock freedom. The consequence of this incremental approach helps us achieve a high degree of automatization. Our approach can be useful for formalizing and developing other kinds of reactive routing protocols such as AODV.

6.10. Bounding message length in attacks against security protocols

Participant: Marie Duflot-Kremer.

Security protocols are short programs that describe communication between two or more parties in order to achieve security goals. Despite the apparent simplicity of such protocols, their verification is a difficult problem and has been shown to be undecidable in general. This undecidability comes from the fact that the set of executions to be considered is of infinite depth (an infinite number of protocol sessions can be run) and infinitely branching (the intruder can generate an unbounded number of distinct messages). Several attempts have been made to tackle each of these sources of undecidability. Together with Myrto Arapinis, we have shown [22] that, under a syntactic and reasonable condition of “well-formedness” on the protocol, we can get rid of the infinitely branching part. More precisely we proved that as far as the secrecy property is considered and for a well-formed protocol, we just need to consider well-typed attacks, with a strong typing system. This result directly implies that the messages to be considered are of bounded length. We are currently working on a journal version of this result that extends the set of security properties to which the result is applicable, in particular including authentication properties.
6.11. Formally verified decision procedures for finite automata

Participants: Stephan Merz, Julien Perugini, Hernán Ponce de Leon, Pierre Savonitto.

Decision problems in the theory of finite automata underly verification algorithms in model checking and decision procedures for fragments of arithmetic. We are interested in developing a certified library of automata-theoretic constructions within a trusted interactive proof assistant such as Isabelle. In 2011, two student projects addressed such problems.

Julien Perugini and Pierre Savonitto formalized a decision procedure for the universality problem of finite automata based on the antichain technique suggested by Doyen et al. [27] and verified its correctness in Isabelle/HOL. They then verified a list-based implementation of that algorithm, using the Isabelle Collections Framework, which provides pre-proved data structures for generating executable implementations. Future work should address efficiency issues by adopting better suited data structures.

During his internship, Hernán Ponce de Leon formalized and verified an automaton-based decision procedure for Presburger arithmetic over the integers, based on a previous encoding of a similar procedure restricted to natural numbers.
6. New Results

6.1. Verification

6.1.1. Analysis of partially observed recursive discrete-event systems

Participants: Sébastien Chédor, Thierry Jéron, Hervé Marchand, Christophe Morvan.

Monitoring of recursive discrete-event systems under partial observation is an important issue with major applications such as the diagnosability of faulty behaviors and the detection of information flow. We consider regular discrete-event systems, that is recursive discrete-event systems definable by deterministic graph grammars. This setting is expressive enough to capture classical models of recursive systems such as the pushdown systems. Hence they are infinite-state in general and standard powerset constructions for monitoring do not apply anymore. We exhibit computable conditions on these grammars together with non-trivial transformations of graph grammars that enable us to construct a monitor. This construction is applied to diagnose faulty behaviors, to detect information flow in regular discrete-event systems, and to generate tests.

6.1.2. Analysis of timed systems

6.1.2.1. Approximate determinization of timed automata

Participants: Nathalie Bertrand, Thierry Jéron, Amélie Stainer.

Timed automata are frequently used to model real-time systems. Their determinization is a key issue for several validation problems. However, not all timed automata can be determinized, and determinizability itself is undecidable. In [18], we propose a game-based algorithm which, given a timed automaton, tries to produce a language-equivalent deterministic timed automaton, otherwise a deterministic over-approximation. Our method subsumes two recent contributions: it is at once more general than an existing (non terminating) determinization procedure by Baier et al. (2009) and more precise than the approximation algorithm of Krichen and Tripakis (2009). Moreover, an extension of the method allows to deal with invariants and $\epsilon$-transitions, and to consider other useful approximations: under approximation, and combination of under- and over-approximations which are particularly useful in testing (see 6.2.1).

6.1.2.2. Frequency analysis for timed automata

Participants: Nathalie Bertrand, Amélie Stainer.

The languages of infinite timed words accepted by timed automata are traditionally defined using Büchi-like conditions. These acceptance conditions focus on the set of locations visited infinitely often along a run, but completely ignore quantitative timing aspects. In [15] we propose a natural quantitative semantics for timed automata based on the so-called frequency, which measures the proportion of time spent in the accepting states. We study various properties of timed languages accepted with positive frequency, and in particular the emptiness and universality problems.

6.1.3. Petri nets reachability graphs

Participant: Christophe Morvan.

Petri nets are a general model for concurrency, the structure of their reachability graph is mostly unknown. In [19] we have investigated the decidability and complexity status of model-checking problems on unlabelled reachability graphs of Petri nets by considering first-order, modal and pattern-based languages without labels on transitions or atomic propositions on markings. We consider several parameters to separate decidable problems from undecidable ones. These results illustrate the intrinsic complexity of the structure of these graphs.
6.2. Active and passive testing

6.2.1. Off-line test selection with test purposes for non-deterministic timed automata

Participants: Nathalie Bertrand, Thierry Jéron, Amélie Stainer.

In [17], we propose novel off-line test generation techniques for non-deterministic timed automata with inputs and outputs (TAIOs) in the formal framework of the tioco conformance theory. In this context, a first problem is the determinization of TAIOs, which is necessary to foresee next enabled actions, but is in general impossible. The determinization problem is addressed in [18] thanks to an approximate determinization using a game approach (see 6.1.2.1). We adapt this procedure here to over- and under-approximation, in order to preserve tioco and guarantee the soundness of generated test cases. A second problem is test selection for which a precise description of timed behaviors to be tested is carried out by expressive test purposes modeled by a generalization of TAIOs. Finally, using a symbolic co-reachability analysis guided by the test purpose, test cases are generated in the form of TAIOs equipped with verdicts.

6.2.2. Test generation using pushdown automata

Participant: Puneet Bhateja.

IOLTS (input output labeled transition system) is a versatile model and is frequently used in model based testing to model the functional behavior of an IUT (implementation under test). However when a system is tested remotely, its observed behavior can be different from its actual functional behavior. In a previous paper, we defined a notion of remotely observed behavior of an IOLTS in terms of its actual behavior. Paper [14] contributes by proposing a methodology to simulate a PDA (pushdown automaton) from the given IOLTS such that the simulated PDA precisely expresses the remotely observed behavior of the IOLTS. The simulated PDA can be thought of as an automatic test generator for remote testing.

6.2.3. Test case selection in asynchronous testing

Participants: Puneet Bhateja, Thierry Jéron.

Conformance testing has a rich underlying formal theory called IOLTS-based conformance testing. Depending upon whether the implementation-under-test (IUT) interacts with its environment directly, or indirectly through a medium, IOLTS-based conformance testing can be classified as synchronous testing or asynchronous testing, respectively. So far the problem of test case selection has been addressed mostly in the context of synchronous testing. In this work we contribute by addressing this problem in the context of asynchronous testing. Though an asynchronously communicating process can be simulated by a synchronously communicating process, the fact that the simulating process is infinite state even if the simulated process is finite state made the problem challenging.

6.2.4. A tagging protocol for asynchronous testing

Participant: Puneet Bhateja.

Conformance testing has a rich underlying theory popularly called IOCO-test theory. In the realm of IOCO-test theory, this paper addresses the issue of testing a component of an asynchronously communicating distributed system. Testing a system which communicates asynchronously (i.e., through some medium) with its environment is more difficult than testing a system which communicates synchronously (i.e., directly without any medium). What impedes asynchronous testing is that the actual behavior of the implementation under test (IUT) appears distorted and infinite to the tester. This impediment consequently renders the problem of generating a complete test suite, from the given specification of the IUT, infeasible. To this end, paper [13] proposes a tagging protocol which when implemented by the asynchronously communicating distributed system will enable the generation of a complete test suite, from the specification of any of its component. Further, this paper describes how to generate the test suite from the given specification of the component.

6.2.5. Abstracting time and data for conformance testing of real-time systems

Participants: Thierry Jéron, Hervé Marchand.
Current approaches to model-based conformance testing of real-time systems are mostly based either on finite state machines/transition systems or on timed automata. However, most real-time systems manipulate data while being subjected to time constraints. The usual solution consists in enumerating data values (in finite domains) while treating time symbolically, thus leading to the classical state explosion problem. Paper [12] with W.L. Andrade and P. Machado (Fed. Univ. Campina Grande) proposes a new model of real-time systems as an extension of both symbolic transition systems and timed automata, in order to handle both data and time requirements symbolically. We then adapt the tioco conformance testing theory to deal with this model and describe a test case generation process based on a combination of symbolic execution and constraint solving for the data part and symbolic analysis for timed aspects.

6.2.6. Ensuring security properties

6.2.6.1. Runtime enforcement monitors: composition, synthesis, and enforcement abilities

Participant: Yliès Falcone.

Runtime enforcement is a powerful technique to ensure that a program will respect a given set of properties. In [9] we extend previous work on this topic in several directions. Firstly, we propose a generic notion of enforcement monitors based on a memory device and finite sets of control states and enforcement operations. Moreover, we specify their enforcement abilities w.r.t. the general Safety-Progress classification of properties. Furthermore, we propose a systematic technique to produce a monitor from the automaton recognizing a given safety, guarantee, obligation or response property. Finally, we show that this notion of enforcement monitors is more amenable to implementation and encompasses previous runtime enforcement mechanisms.

6.2.6.2. What can you verify and enforce at runtime?

Participant: Yliès Falcone.

The underlying property, its definition and representation play a major role when monitoring a system. Having a suitable and convenient framework to express properties is thus a concern for runtime analysis. It is desirable to delineate in this framework the sets of properties for which runtime analysis approaches can be applied to. [8] presents a unified view of runtime verification and enforcement of properties in the Safety-Progress classification. Firstly, we extend the Safety-Progress classification of properties in a runtime context. Secondly, we characterize the set of properties which can be verified (monitorable properties) and enforced (enforceable properties) at runtime. We propose in particular an alternative definition of "property monitoring" to the one classically used in this context. Finally, for the delineated sets of properties, we define specialized verification and enforcement monitors.

6.3. Control synthesis

6.3.1. Controllers for probabilistic systems

Participant: Nathalie Bertrand.

Partially Observable Markov Decision Processes (POMDP for short) have been extensively studied in several research communities, among which AI and model-checking. In [16] we address the problem of the minimal information a user needs at runtime to achieve a simple goal, modeled as reaching an objective with probability one. More precisely, to achieve her goal, the user can either choose at each step to use partial information only, or pay a fixed cost and receive full information. The natural question is then to minimize the cost the user needs to fulfill its objective. This optimization question gives rise to two different problems, whether we consider to minimize the worst case cost, or the average cost. On the one hand, concerning the worst case cost, we show that efficient techniques from the model checking community can be adapted to compute the optimal worst case cost and give optimal strategies for the users. On the other hand, we show that the optimal average price (a question typically considered in the AI community) cannot be computed in general, nor can it be approximated in polynomial time even up to a large approximation factor.
6.3.2. **Supervisory control for synchronous systems**

6.3.2.1. **Controller synthesis and programming language**

**Participant:** Hervé Marchand.

In [24], we define a mixed imperative/declarative programming language: declarative contracts are enforced upon imperatively described behaviors. We rely on the notion of Discrete Controller Synthesis (DCS), a formal technique stemming from control theory and the supervisory control of discrete event systems. We target the application domain of adaptive and reconfigurable computing systems: our language can serve programming closed-loop adaptation controllers, enabling flexible execution of functionalities w.r.t. changing resource and environment conditions. We give a synthetic presentation of the language, its semantics and compilation, and we illustrate its use with the example of a robot system.

6.3.2.2. **Symbolic supervisory control of infinite transition systems under partial observation using abstract interpretation**

**Participant:** Hervé Marchand.

In [11], we propose algorithms for the synthesis of state-feedback controllers with partial observation of infinite state discrete event systems modelled by Symbolic Transition Systems. We provide models of safe memoryless controllers both for potentially deadlocking and deadlock free controlled systems. The termination of the algorithms solving these problems is ensured using abstract interpretation techniques which provide an overapproximation of the transitions to disable. We then extend our algorithms to controllers with memory and to online controllers. We also propose improvements in the synthesis of controllers in the finite case which, to our knowledge, provide more permissive solutions than what was previously proposed in the literature. Our tool SMACS gives an empirical validation of our methods by showing their feasibility, usability and efficiency.

6.3.2.3. **Decentralized control of infinite systems**

**Participant:** Hervé Marchand.

In [10], we propose algorithms for the synthesis of decentralized state-feedback controllers with partial observation of infinite state systems, which are modeled by Symbolic Transition Systems. We first consider the computation of safe controllers ensuring the avoidance of a set of forbidden states and then extend this result to the deadlock free case. The termination of the algorithms solving these problems is ensured by the use of abstract interpretation techniques, but at the price of overapproximations, in particular, in the computation of the states which must be avoided. We then extend our algorithms to the case where the system to be controlled is given by a collection of subsystems (modules). This structure is exploited to locally compute a controller for each module. Our tool SMACS gives an empirical evaluation of our methods by showing their feasibility, usability and efficiency.

6.3.2.4. **Polychronous controller synthesis from MARTE CCSL timing specifications**

**Participant:** Hervé Marchand.

The UML Profile for Modeling and Analysis of Real-Time and Embedded systems (MARTE) defines a mathematically expressive model of time, the Clock Constraint Specification Language (CCSL), to specify timed annotations on UML diagrams and thus provides them with formally defined timed interpretations. Thanks to its expressive capability, the CCSL allows for the specification of static and dynamic properties, of deterministic and non-deterministic behaviors, or of systems with multiple clock domains. Code generation from such multicolocated specifications (for the purpose of synthesizing a simulator, for instance) is known to be a difficult issue. We address it in [23] by using the approach of controller synthesis. In our framework, a timed CCSL specification is regarded as a property whose satisfaction should be enforced for any UML diagram carrying it as annotation. To do so, CCSL statements are first translated into dynamical polynomial systems. Such systems can be manipulated using the model-checker Sigali to synthesize an executable property (a controller) which enforces the satisfaction of the specified timing constraints on the UML diagram with which it is executed.
6.3.3. Control of distributed systems

**Participant:** Hervé Marchand.

In this work, we consider the control of distributed systems composed of subsystems communicating asynchronously; the aim is to build local controllers that restrict the behavior of a distributed system in order to satisfy a global state avoidance property. We model our distributed systems as communicating finite state machines with reliable unbounded FIFO queues between subsystems. Local controllers can only observe the behavior of their local subsystem and do not see the queue contents. To refine their control policy, the controllers can use the FIFO queues to communicate by piggybacking extra information (some timestamps and their state estimates) to the messages sent by the subsystems [21]. We provide an algorithm that computes, for each local subsystem (and thus for each controller), during the execution of the system, an estimate of the current global state of the distributed system. The local estimate is updated at each message reception. We then define synthesis algorithms allowing to compute the local controllers. Our method relies on the computation of (co)reachable states. Since the reachability problem is undecidable in our model, we use abstract interpretation techniques to obtain regular overapproximations of the possible FIFO queue contents, and hence of the possible current global states. An implementation of our algorithms provides an empirical evaluation of our method [22].
5. New Results

5.1. Snell envelope with small probability criteria

We present in [33] a new algorithm to compute the Snell envelope in the specific case where the criteria to optimize is associated with a small probability or a rare event. This new approach combines the Stochastic Mesh approach of Broadie and Glasserman with a particle approximation scheme based on a specific change of measure designed to concentrate the computational effort in regions pointed out by the criteria. The theoretical analysis of this new algorithm provides non asymptotic convergence estimates. Finally, the numerical tests confirm the practical interest of this approach.

5.2. On the concentration properties of Interacting particle processes

In [34], we present some new concentration inequalities for Feynman-Kac particle processes. We analyze different types of stochastic particle models, including particle profile occupation measures, genealogical tree based evolution models, particle free energies, as well as backward Markov chain particle models. We illustrate these results with a series of topics related to computational physics and biology, stochastic optimization, signal processing and bayesian statistics, and many other probabilistic machine learning algorithms. Special emphasis is given to the stochastic modeling and the quantitative performance analysis of a series of advanced Monte Carlo methods, including particle filters, genetic type island models, Markov bridge models, interacting particle Markov chain Monte Carlo methodologies.

5.3. A Robbins-Monro procedure for estimation in semiparametric regression models

The paper [39] is devoted to the parametric estimation of a shift together with the nonparametric estimation of a regression function in a semiparametric regression model. We implement a Robbins-Monro procedure very efficient and easy to handle. On the one hand, we propose a stochastic algorithm similar to that of Robbins-Monro in order to estimate the shift parameter. A preliminary evaluation of the regression function is not necessary for estimating the shift parameter. On the other hand, we make use of a recursive Nadaraya-Watson estimator for the estimation of the regression function. This kernel estimator takes in account the previous estimation of the shift parameter. We establish the almost sure convergence for both Robbins-Monro and Nadaraya-Watson estimators. The asymptotic normality of our estimates is also provided.

5.4. Discrete Time Markovian Agents Interacting Through a Potential

A discrete time stochastic model for a multiagent system given in terms of a large collection of interacting Markov chains is studied. The evolution of the interacting particles is described through a time inhomogeneous transition probability kernel that depends on the ‘gradient’ of the potential field. The particles, in turn, dynamically modify the potential field through their cumulative input. Interacting Markov processes of the above form have been suggested as models for active biological transport in response to external stimulus such as a chemical gradient. One of the basic mathematical challenges is to develop a general theory of stability for such interacting Markovian systems and for the corresponding nonlinear Markov processes that arise in the large agent limit. Such a theory would be key to a mathematical understanding of the interactive structure formation that results from the complex feedback between the agents and the potential field. It will also be a crucial ingredient in developing simulation schemes that are faithful to the underlying model over long periods of time. The goal of the work developed in [42] is to study qualitative properties of the above stochastic system as the number of particles (N) and the time parameter (n) approach infinity. In this regard asymptotic properties of a deterministic nonlinear dynamical system, that arises in the propagation of chaos
limit of the stochastic model, play a key role. We show that under suitable conditions this dynamical system has a unique fixed point. This result allows us to study stability properties of the underlying stochastic model. We show that as \( N \to \infty \), the stochastic system is well approximated by the dynamical system, uniformly over time. As a consequence, for an arbitrarily initialized system, as \( N \to \infty \) and \( n \to \infty \), the potential field and the empirical measure of the interacting particles are shown to converge to the unique fixed point of the dynamical system. In general, simulation of such interacting Markovian systems is a computationally daunting task. We propose a particle based approximation for the dynamic potential field which allows for a numerically tractable simulation scheme. It is shown that this simulation scheme well approximates the true physical system, uniformly over an infinite time horizon.

5.5. An Adaptive Interacting Wang-Landau Algorithm for Automatic Density Exploration

While statisticians are well-accustomed to performing exploratory analysis in the modeling stage of an analysis, the notion of conducting preliminary general-purpose exploratory analysis in the Monte Carlo stage (or more generally, the model-fitting stage) of an analysis is an area which we feel deserves much further attention. Towards this aim, the paper [41] proposes a general-purpose algorithm for automatic density exploration. The proposed exploration algorithm combines and expands upon components from various adaptive Markov chain Monte Carlo methods, with the Wang-Landau algorithm at its heart. Additionally, the algorithm is run on interacting parallel chains – a feature which both decreases computational cost as well as stabilizes the algorithm, improving its ability to explore the density. Performance is studied in several applications. Through a Bayesian variable selection example, the authors demonstrate the convergence gains obtained with interacting chains. The ability of the algorithm’s adaptive proposal to induce mode-jumping is illustrated through a trimodal density and a Bayesian mixture modeling application. Lastly, through a 2D Ising model, the authors demonstrate the ability of the algorithm to overcome the high correlations encountered in spatial models.

5.6. A sharp analysis on the asymptotic behavior of the Durbin-Watson statistic for the first-order autoregressive process

The purpose of the paper [40] is to provide a sharp analysis on the asymptotic behavior of the Durbin-Watson statistic. We focus our attention on the first-order autoregressive process where the driven noise is also given by a first-order autoregressive process. We establish the almost sure convergence and the asymptotic normality for both the least squares estimator of the unknown parameter of the autoregressive process as well as for the serial correlation estimator associated to the driven noise. In addition, the almost sure rates of convergence of our estimates are also provided. It allows us to establish the almost sure convergence and the asymptotic normality for the Durbin-Watson statistic. Finally, we propose a new bilateral statistical test for residual autocorrelation.

5.7. Large deviations for Gaussian stationary processes and semi-classical analysis

In [37], we obtain a large deviation principle for quadratic forms of Gaussian stationary processes. It is established by the conjunction of a result of Roch and Silbermann on the spectrum of products of Toeplitz matrices together with the analysis of large deviations carried out by Gamboa, Rouault and the rst author. An alternative proof of the needed result on Toeplitz matrices, based on semi-classical analysis, is also provided.

5.8. Sharp large deviations for the non-stationary Ornstein-Uhlenbeck process

For the Ornstein-Uhlenbeck process, the asymptotic behavior of the maximum likelihood estimator of the drift parameter is totally different in the stable, unstable, and explosive cases. Notwithstanding of this trichotomy, we investigate sharp large deviation principles for this estimator in the three situations. In the explosive case, we exhibit in [38] a very unusual rate function with a shaped flat valley and an abrupt discontinuity point at its minimum.
5.9. Markovian quadratic and superquadratic BSDEs with an unbounded terminal condition

The work in [43] deals with the existence and the uniqueness of solutions to quadratic and superquadratic Markovian backward stochastic differential equations (BSDEs for short) with an unbounded terminal condition. Our results are deeply linked with a strong a priori estimate on $Z$ that takes advantage of the Markovian framework. This estimate allows us to prove the existence of a viscosity solution to a semilinear parabolic partial differential equation with nonlinearity having quadratic or superquadratic growth in the gradient of the solution. This estimate also allows us to give explicit convergence rates for time approximation of quadratic or superquadratic Markovian BSDEs.

5.10. New approach on recursive and non-recursive SIR

In [3], we consider a semiparametric single index regression model involving a p-dimensional quantitative covariable $x$ and a real dependent variable $y$. A dimension reduction is included in this model via an index $x^T \beta$. Sliced inverse regression (SIR) is a well-known method to estimate the direction of the Euclidean parameter $\beta$ which is based on a "slicing step" of $y$ in the population and sample versions. The goal of this paper is twofold. On the one hand, we focus on a recursive version of SIR which is also suitable for multiple indices model. On the other hand, we propose a new method called SIRoneslice when the regression model is a single index model. The SIRoneslice estimator of the direction of $\beta$ is based on the use of only one "optimal" slice chosen among the H slices. Then, we provide its recursive version. We give an asymptotic result for the SIRoneslice approach. Simulation study shows good numerical performances of the SIRoneslice method and clearly exhibits the main advantage of using recursive versions of the SIR and SIRoneslice methods from a computational time point of view. A real dataset is also used to illustrate the approach. The proposed methods and criterion have been implemented in R and the corresponding codes are made available.

5.11. Classification of EEG data by evolutionary algorithm for the study of vigilance states

The objective of this work [18] is to predict the state of vigilance of an individual from the study of its brain activity (EEG signals). The variable to predict is binary (alertness "normal" or "relaxed"). EEG of 44 participants in both states (88 records) were collected with a helmet with 58 electrodes. After a pretreatment step and data validation, a test called "test slope" was chosen. The usual methods of supervised classification (k nearest neighbors, binary classification trees, random forests, and discriminant sparse PLS) were used to provide predictions of the state of participants. The test was then refined using a genetic algorithm, which has built a reliable model (average true classification rate by using CART equal to 86.68 +/- 1.87\%) and to select an electrode from the initial 58.

5.12. Genetic Programming

Recently, it has been stated that the complexity of a solution is a good indicator of the amount of overfitting it incurs. However, measuring the complexity of a program, in Genetic Programming, is not a trivial task. In [22], we study the functional complexity and how it relates with overfitting on symbolic regression problems. We consider two measures of complexity, Slope-based Functional Complexity, inspired by the concept of curvature, and Regularity-based Functional Complexity based on the concept of Holderian regularity. In general, both complexity measures appear to be poor indicators of program overfitting. However, results suggest that Regularity-based Functional Complexity could provide a good indication of overfitting in extreme cases. During the development of applied systems, an important problem that must be addressed is that of choosing the correct tools for a given domain or scenario. This general task has been addressed by the genetic programming (GP) community by attempting to determine the intrinsic difficulty that a problem poses for a GP search. In [21], we present an approach to predict the performance of GP applied to data classification, one of the most common problems in computer science. The novelty of the proposal is to extract statistical descriptors
and complexity descriptors of the problem data, and from these estimate the expected performance of a GP classifier. We derive two types of predictive models: linear regression models and symbolic regression models evolved with GP. The experimental results show that both approaches provide good estimates of classifier performance, using synthetic and real-world problems for validation. In conclusion, this paper shows that it is possible to accurately predict the expected performance of a GP classifier using a set of descriptors that characterize the problem data.

The analysis of image regularity using Holder exponents can be used to characterize singular structures contained within an image, and provide a compact description of local shape and appearance. However, estimating the Holder exponent is not a trivial task and current methods tend to be slow and complex. Therefore, the goal in [35] is to automatically synthesize image operators that can be used to estimate the Holder regularity of an image. We pose this task as an optimization problem and use Genetic Programming (GP) to search for operators that can approximate a traditional estimator, the oscillations method. In our experiments, GP was able to evolve estimators that achieve a low error and a high correlation with the ground truth estimation. Furthermore, most of the GP estimators are faster than the traditional approaches, in some cases their runtime is orders of magnitude smaller. This result allowed us to implement a real-time estimation of the Holder exponent on a live video signal, the first such implementation in current literature. Moreover, the evolved estimators are used to generate local descriptors of salient image regions, a task for which we obtain a stable and robust matching that is comparable with state-of-the-art methods. In conclusion, the evolved estimators produced by GP could help expand the application domain of Holderian regularity within the fields of image analysis and signal processing.

5.13. Relevance of the Holderian regularity-based interpolation for range-Doppler ISAR image post-processing

In ISAR processing, post-processing of the range Doppler image is useful to help the practitioner for ship recognition. Among the image post-processing tools, interpolation methods can be of interest especially when zooming. In [19], we study the relevance of the Holderian regularity-based interpolation. In that case, interpolating consists in adding a new scale in the wavelet transform and the new wavelet coefficients can be estimated from others. In the original method, initially proposed by two of the authors, the image is first interpolated along the rows and then along the columns. Concerning the diagonal pixels, they are estimated as the mean of the adjacent original and interpolated pixels. Here, we propose a variant where the diagonal pixels are estimated by taking into account the local orientation of the image. It has the advantage of conserving local regularity on all interpolated pixels of the image. A comparative study on synthetic data and real range-Doppler images is then carried out with alternative interpolation techniques such as the linear interpolation, the bicubic one, the nearest neighbour interpolation, etc. The simulation results confirm the effectiveness of the approach.

5.14. On-line changepoint detection and parameter estimation with application to genomic data

We propose in [10] an efficient on-line changepoint detection algorithm for an important class of Bayesian product partition models. The algorithm allows to estimate jointly on-line the static parameters of the model using a recursive maximum likelihood estimation strategy. This particle filter type algorithm has a computational complexity which scales linearly both in the number of data and the number of particles. We demonstrate our methodology on a synthetic and two real world datasets from RNA transcript analysis. On simulated data, it is shown that our approach outperforms standard techniques used in this context and hence has the potential to detect novel RNA transcripts.

5.15. Bayesian Sparsity-Path-Analysis of Genetic Association Signal using Generalized t Priors
In [17], we explore the use of generalized t priors on regression coefficients to help understand the nature of association signal within "hit regions" of genome-wide association studies. The particular generalized t distribution we adopt is a Student distribution on the absolute value of its argument. For low degrees of freedom we show that the generalized t exhibits 'sparsity-prior' properties with some attractive features over other common forms of sparse priors and includes the well known double-exponential distribution as the degrees of freedom tends to infinity. We pay particular attention to graphical representations of posterior statistics obtained from sparsity-path-analysis (SPA) where we sweep over the setting of the scale (shrinkage / precision) parameter in the prior to explore the space of posterior models obtained over a range of complexities, from very sparse models with all coefficient distributions heavily concentrated around zero, to models with diffuse priors and coefficients distributed around their maximum likelihood estimates. The SPA plots are akin to LASSO plots of maximum a posteriori (MAP) estimates but they characterize the complete marginal posterior distributions of the coefficients plotted as a function of the precision of the prior. Generating posterior distributions over a range of prior precisions is computationally challenging but naturally amenable to sequential Monte Carlo (SMC) algorithms indexed on the scale parameter. We show how SMC simulation on graphic-processing-units (GPUs) provides very efficient inference for SPA. We also present a scale-mixture representation of the generalized t prior that leads to an EM algorithm to obtain MAP estimates should only these be required.
6. New Results

6.1. Tralics: a LaTeX to XML Translator

Participant: José Grimm.

The major use of Tralics remains the production of the RaWeb (Scientific Annex to the Annual Activity Report of Inria). The software is described in [65], [68], [67], [66]. Other applications of Tralics consist in putting scientific papers on the Web; for instance Cedram (http://www.cedram.org, Centre de diffusion de revues académiques mathématiques), that publishes the Journal de théorie des nombres de Bordeaux, uses Tralics for the abstracts and plans to translate full papers. Tralics is also used by Zentralblatt for converting comments, reviews and abstracts. The Software has been presented at the DML2010 conference [69]. Tralics has been used for the HTML+MathML documentation of the open TURNS software http://trac.openturns.org/.

6.2. Inverse problems for elliptic operators


6.2.1. Boundary value problems for Laplace equation in 3-D

Solving overdetermined Cauchy problems for the Laplace equation on a spherical layer (in 3-D) in order to process incomplete experimental data is a necessary ingredient of the team’s approach to inverse source problems, in particular for applications to EEG since the latter involves propagating the initial conditions from the boundary to the center of the domain where the singularities (i.e., the sources) are sought after. Here, the domain is typically made of several homogeneous layers of different conductivities.

Such problems offer an opportunity to state and solve extremal problems for harmonic fields for which an analog of the Toeplitz operator approach to bounded extremal problems [45] has been obtained in [2]. Still, a best approximation on the subset of a general vector field generated by a harmonic gradient under a $L^2$ norm constraint on the complementary subset can be computed by an inverse spectral equation for some Toeplitz operator. Constructive and numerical aspects of the procedure (harmonic 3-D projection, Kelvin and Riesz transformation, spherical harmonics) and encouraging results have been obtained on numerically simulated data.

Issues of robust interpolation on the sphere from incomplete pointwise data are also under study in order to improve numerical accuracy of our reconstruction schemes. Spherical harmonics, Slepian bases and related special functions are of special interest (thesis of A.-M. Nicu), while splines, spherical wavelets, cubature techniques should be considered as well.

It turns out that Slepian functions are eigenfunctions of truncated Toeplitz operators in the complex plane (the framework of 2-D problems). These properties will be used in order to quantify the robustness properties of our resolution schemes for $L^2$ bounded extremal problems [45], and to establish error estimates.
The analogous problem in $L^p$, $p \neq 2$, is quite important to get tighter control on pointwise approximation. However, it is considerably more difficult. In a collaborative effort with the University of Orleans, within the framework of the ANR project AHPI, we set ourselves the goal of understanding the case $p = \infty$ better. Namely, connections between the $\text{BMO}^2$ distance of a bounded vector field on the sphere to a $\text{BMO}$ harmonic gradient and the spectral properties of a “big” Hankel-like operator acting on $L^2$ harmonic gradients and valued in its orthogonal space are currently being investigated. We obtained a generalization of the Hodge decomposition (we call it the Hardy-Hodge decomposition) for $\mathbb{R}^n$-valued vector fields on $\mathbb{R}^{n-1}$ (resp. $\mathbb{S}^{n-1}$) which stands analog to the Hardy direct sum decomposition in dimension 1. In this decomposition, the analytic part becomes the trace of the gradient of a harmonic function in the half-space (resp. the ball) whose $\text{BMO}$-norm on parallel hyperplanes (resp. concentric spheres) is uniformly bounded. The two main difficulties facing a generalization of Nehari’s theorem are the absence of a constructive derivation of Coifman-Rochberg weak factorizations and Wolff’s phenomenon that a harmonic gradient (even $C^1$-smooth) is not determined by its values on a set of positive measure on $\mathbb{R}^n$ (resp. $\mathbb{S}^n$). This last point is the only obstacle to establish uniqueness and constancy-of-modulus properties. We shall concentrate on these two items in the future.

The above issue is also interesting in $L^p$, $1 < p < \infty$, where it leads to analyze particular solutions to the the $p$-Laplacian on the sphere. This aspect is not pursued in depth at the moment.

### 6.2.2. Sources recovery in 3-D domains, application to MEEG and geophysics

The problem of sources recovery can be handled in 3-D balls by using best rational approximation on 2-D cross sections (disks) from traces of the boundary data on the corresponding circles (see section 4.1). In 3-D, functional or clinical active regions in the cortex are often represented by pointwise sources that have to be localized from measurements on the scalp of a potential satisfying a Laplace equation (EEG, electroencephalography). In the work [7] it was shown how to proceed via best rational approximation on a sequence of 2-D disks cut along the inner sphere, for the case where there are at most 2 sources. A long-haul research on the behaviour of poles of best rational approximants of fixed degree to functions with branch points was completed this year [19], which shows that the technique carries over to finitely many sources.

In this connection, a dedicated software “FindSources3D” (see section 5.7) has been developed, in collaboration with the team Athena.

Further, it appears that in the rational approximation step of these schemes, multiple poles possess a nice behaviour with respect to the branched singularities (see figure 4). This is due to the very basic physical assumptions on the model (for EEG data, one should consider triple poles). Though numerically observed, there is no mathematical justification so far why these multiple poles have such strong accumulation properties, which remains an intriguing observation. This is the topic of [30].

Also, magnetic data from MEG (magneto-encephalography) will soon become available, which should enhance the accuracy of source recovery algorithms.

This approach also appears to be interesting for geophysical issues, namely discretizing the gravitational potential by means of pointwise masses. This is one topic of A.-M. Nicu’s PhD thesis.

Magnetic sources localization from observations of the field away from the support of the magnetization is an issue under investigation in a joint effort with the Math. department of Vanderbilt University and the Earth Sciences department at MIT. The goal is to recover the magnetic properties of rock samples (meteorites) from fine measurements extremely close to the sample that can nowadays be obtained using SQUIDs (supraconducting coil devices).

The magnetization operator is the Riesz potential of the divergence of the magnetization. When the latter has bounded variation, we already described the kernel of this operator (the so-called silent magnetizations or silent source distributions) in terms of measures whose balayage on the boundary of the sample vanishes. This however, is not so very effective, computationally.

---

2 bounded mean oscillation
The case of a thin slab (the magnetization is then modelled as a vector field on a portion of the plane) has proved more amenable. We have shown that that silent sources from above or below can be characterized via the Hardy-Hodge decomposition mentioned in section 6.2.1. The smoothness assumptions have been weakened considerably to accommodate magnetizations that may be any distribution with compact support, more generally any finite sum of partial derivatives of any order of $L^p$ or $BMO$ functions. Silent unidirectional and bi-directional magnetizations demonstrably reduce this way to certain divergence free tangential vector fields. In particular no nonzero compactly supported unidirectional magnetization exists. In the $L^2$ setting, equivalent magnetizations of minimal $L^2$-norm can be computed using the Hardy-Hodge decomposition (which is orthogonal in this case), and an uncertainty principle relating the support of a magnetization and the support of its minimum-norm equivalent magnetization has been obtained. A paper is being written on these results.

Meanwhile, the severe ill-posedness of the reconstruction challenges discrete Fourier methods, one of the main problems being the truncation of the observations outside the range of the SQUID measurements. A next step will be to develop the extrapolation techniques initiated by the project team, using bounded extremal problems, in an attempt to overcome this issue.

### 6.2.3. Boundary value problems for 2-D conductivity equations, application to plasma control

In collaboration with the CMI-LATP (University Marseille I) and in the framework of the ANR AHPI, the team considers 2-D diffusion processes with variable conductivity. In particular its complexified version, the so-called conjugate or real Beltrami equation, was investigated. In the case of a smooth domain, and for Lipschitz conductivity, we analyzed the Dirichlet problem for solutions in Sobolev and then in Hardy classes $H^s$. Their traces merely lie in $L^p (1 < p < \infty)$ of the boundary, a space which is suitable for identification from pointwise measurements. Again these traces turn out to be dense on strict subsets of the boundary. This allows us to state Cauchy problems as bounded extremal issues in $L^p$ classes of generalized analytic functions, in a reminiscent manner of what was done for analytic functions as discussed in section 3.1.1.

This year we generalized the construction to finitely connected Dini-smooth domains and $W^{1,q}$-smooth conductivities, with $q > 2$. The case of an annular geometry is the relevant one for the application to plasma shaping mentioned below. The application that initially motivated this work came from free boundary problems in plasma confinement (in tokamaks) for thermonuclear fusion. This work was initiated in...
collaboration with the Laboratoire J. Dieudonné (University of Nice) and is now the topic of a collaboration with two teams of physicists from the CEA-IRFM (Cadarache).

In the transversal section of a tokamak (which is a disk if the vessel is idealized into a torus), the so-called poloidal flux is subject to some conductivity equation outside the plasma volume for some simple explicit smooth conductivity function, while the boundary of the plasma (in the Tore Supra tokamak) is a level line of this flux [53]. Related magnetic measurements are available on the chamber, which furnish incomplete boundary data from which one wants to recover the inner (plasma) boundary. This free boundary problem (of Bernoulli type) can be handled through the solutions of a family of bounded extremal problems in generalized Hardy classes of solutions to real Beltrami equations, in the annular framework. Such approximation problems also allow us to approach a somewhat dual extrapolation issue, raised by colleagues from the CEA for the purpose of numerical simulation. It consists in recovering magnetic quantities on the outer boundary (the chamber) from an initial guess of what the inner boundary (plasma) is.

In the particular case at hand, the conductivity is $1/x$ and the domain is an annulus embedded in the right half-plane. We obtained a basis of solutions (exponentials times Legendre functions) upon separating variables in toroidal coordinates. This may be viewed as a generalization to the annulus of the Bessel type expansions derived in [21] for simply connected geometries. This provides a computational setting to solve the extremal problems mentioned before, and was the topic of the PhD thesis of Y. Fischer [17], [22]. In the most recent tokamaks, like Jet or ITER, an interesting feature of the level curves of the poloidal flux is the occurrence of a cusp (a saddle point of the poloidal flux, called an X point), and it is desirable to shape the plasma according to a level line passing through this X point for physical reasons related to the efficiency of the energy transfer. We established well-posedness of the Dirichlet problem in weighted $L^p$ classes for harmonic measure on piecewise smooth domains without cusps, thereby laying ground for such a study. This issue is next in line, now that the present approach has been validated numerically on Tore Supra data.

On the half-plane, the conductivity $1/x$ is severely unbounded but the analysis of this test case is quite important for the convergence of extrapolation algorithms to recover magnetic quantities on the chamber. Additive decompositions into Hardy solutions inside the outer boundary and outside the inner boundary, with controlled vanishing on the imaginary axis, have been obtained as part of the PhD work of S. Chaabi. The latter developed this year a multiplicative parameterization of Hardy-smooth solutions by holomorphic functions for the conjugate Beltrami equation, a result which is both subtler and weaker than the classical Stoilow factorization for solutions to the complex Beltrami equation. This factorization is of considerable numerical interest in situations where conductivity is little or not known.

6.3. Rational Approximation for fitting Non Negative EPT densities

**Participants:** Martine Olivi, Bernard Hanzon [Univ. Cork], Conor Sexton [Univ. Cork], Fabien Seyfert.

We explored this year a new application field for our rational approximation methods. We studied the problem of fitting a probability density function from a large set of financial data. The class of density function that we considered is that of non negative EPT (Exponential-Trigonometric-Polynomials) functions that seems to provide a very relevant framework for probabilistic calculations. Parseval’s theorem implies that approximating the rational transform is equivalent to approximating the density itself. During his visit, Conor Sexton (a PHD student of Bernard Hanzon) adapted and ran the RARL2 software on this problem. The results were encouraging albeit the major problem of imposing positivity is still under study.

6.4. Orthogonal rational functions and non-stationary stochastic processes

**Participants:** Laurent Baratchart, Stanislas Kupin [Univ. Bordeaux 1].
The theory of orthogonal polynomials on the unit circle is a most classical piece of analysis which is still the object of intensive studies. The asymptotic behaviour of orthogonal polynomials is of special interest for many issues pertaining to approximation theory and to spectral theory of differential operators. Its connection with prediction theory of stationary stochastic processes has long been known \[64\]. Namely, the \(n\)-th orthonormal polynomial with respect to the spectral measure of the process yields the optimal regression coefficients of a linear one-step ahead predictor from the \(n - 1\)-st last values, in the sense of minimum variance of the error. Likewise, the (inverse of) the dominant coefficient of the polynomial gives the prediction error. In particular, asymptotics for the dominant coefficient determine the asymptically optimal prediction error from the past as time goes large.

As compared to orthogonal polynomials, orthogonal rational functions have not been much considered up to now. They were apparently introduced by Dzrbasjan but the first systematic exposition seems to be the monograph by Bultheel et al. \[57\] where the emphasis is more on the algebraic side of the theory. In fact, the asymptotic analysis of orthogonal rational functions is still in its infancy.

We recently developed an analog of the Kolmogorov-Krein-Szegö theorem \[18\] for orthogonal rational functions which is first of its kind in that it allows for the poles of these functions to approach the unit circle, generalizing previously known results for compactly supported singular set. Dwelling on this asymptotic analysis of orthogonal rational functions, we developed a prediction theory for certain, possibly nonstationary stochastic processes that we call Blaschke varying processes. \[44\]. These are characterized by a spectral calculus where time shift corresponds to multiplication by an elementary Blaschke product (that may depend on the time instant considered). This class of processes contains the familiar Gaussian stationary processes, but it contains many more that exhibit a much more varied behaviour. For instance, the process may be asymptotically deterministic along certain subsequences and nondeterministic along others. The optimal predictor is constructed from the spectral measure via orthogonal rational functions, and its asymptotic behaviour is characterized by the above-mentioned generalization of the Kolmogorov-Krein-Szegö theorem. In the same vein, we also developed prediction theory for another class of nonstationary processes, the so-called Cauchy-processes, that may be characterized as stationary processes feedind in turn a sequence of varying filters of degree 1. Their covariance matrices can be characterized via Nevanlinna-Pick interpolation. The issue of characterizing covariance sequences of Blaschke processes is still open. Their identification raises the problem of constructing optimal Schur rational approximants to a given Schur function.

6.5. Rational and meromorphic approximation

**Participants:** Laurent Baratchart, Herbert Stahl [TFH Berlin], Maxim Yattselev.

We demonstrated in a recent past, under mild smoothness assumptions, the possibility of convergent rational interpolation to Cauchy integrals of complex measures on analytic Jordan arcs and their strong asymptotics \[14\], \[13\]. Subsequently, we started investigating the case of Cauchy integrals on so-called symmetric contours for the logarithmic potential. These correspond to functions with more than two branched singularities, like those arising in the slicing method for source recovery in a sphere when there is more than one source (see section \ref{subsec:6.2.2}). Recently we obtained weak asymptotics in this case, and dwelled on them to elucidate the asymptotic of poles of best \(L^2\) meromorphic approximants of given degree to a function with branched singularities on a curve encompassing them. Namely, the counting measure of the poles converges weak-star, when the degree goes large, to the Green equilibrium distribution of the set with minimum Green capacity inside the curve, outside of which the function is single-valued. The technical core of this contribution is an existence and uniqueness result, along with a differential characterization, of the compact of minimum weighted capacity outside of which the function is single-valued \[19\]. This teams up with results from \[63\] to produce the results.

We presently study strong asymptotics, limiting ourselves at present to a threefold geometry, and to the case of Padé approximants (interpolation at a single point with high order). The result is that uniform convergence can only take place if the weights of the branches of the threefold with respect to the equilibrium distribution are rational. If they are rationally dependent, a spurious pole clusters to certain curves within the domain of analyticity, and if they are rationally independent, exactly one pole exhibits chaotic behaviour in the complex
plane. Moreover, we have shown that the chaotic situation is generic, in a measure theoretic sense, with respect to the location of branchpoints. This generalizes and sharpens results of Suetin for Cauchy integrals on disconnected pieces of a smooth symmetric contour. It is the first time that a branched contour is analyzed with respect to general densities. A paper is being written to report these results.

6.6. Circuit realisations of filter responses: determination of canonical forms and exhaustive computations of constrained realisations

Participants: Smain Amari [Royal Military College, Kingston, Canada], Jean Charles Faugère [EPI SALSA, INRIA Rocquencourt], Giuseppe Macchiarella [Politecnico di Milano, Milan, Italy], Uwe Rosenberg [Design and Project Engineering, Osterholz-Scharmbeck, Germany], Matteo Oldoni [Politecnico di Milano, Milan, Italy], Fabien Seyfert.

We pursued our work on circuit realisations of filter responses with mixed type (inductive or capacitive) coupling elements and constrained topologies [1]. We now focus on the use of resonating couplings in the design of asymmetric filter’s characteristics without the use cross-coupling in order to simplify the practical implementation. In parallel, efforts are being payed to improve the synthesis method for higher order filters, having in mind application to diplexers with high number of symmetrically located transmission zeros.

6.7. Synthesis of compact multiplexers and de-embedding of multiplexers

Participants: Martine Olivi, Fabien Seyfert, Stéphane Bila [Xlim, Limoges, France], Hussein Ezzedin [Xlim, Limoges, France], Damien Pacaud [Thales Alenia Space, Toulouse, France], Giuseppe Macchiarella [Politecnico di Milano, Milan, Italy], Matteo Oldoni [Politecnico di Milano, Milan, Italy].

6.7.1. Synthesis of compact multiplexers

Our work on the synthesis problem for diplexers has continued this year. Based on the polynomial structure highlighted last year [39], a synthesis algorithm was devised and allowed the effective synthesis of multiplexer characteristics (see figure 5). As opposed to other synthesis algorithms [74] the latter only involves polynomial computations on the sub-filters of the overall multiplexer, allowing its application to synthesis problems relative to devices with a numerous number of ports. The latter is based on the recursive solving of an extended Nevanlinna-Pick problem, first introduced in [62]. The convergence of this fixed point procedure is under study as well as its extension to the synthesis of general multiplexer.

Figure 5. Diplexer’s characteristic computed with an algorithm based on the recursive solving of extended Nevanlinna-Pick interpolation problems.
6.7.2. De-embedding of multiplexers

While we presented our work on the de-embedding diplexers at [26] we extended the latter to general multiplexers. The multiplexer de-embedding problem we study is the following. Let $S$ be the measured scattering matrix of a multiplexer composed of a $N$-port junction with response $T$ and $N-1$ filtering devices with responses $F_1, \cdots F_{N-1}$ as depicted by Figure 6.

![Figure 6. Multiplexer made of a junction $T$ and filtering devices $F_1, F_2 \cdots F_N$.](image)

The de-embedding question is the following: given $S$ and $T$, is it possible to derive the $F_k$’s? We derived some rather important and surprising characteristics of this general problem:

- When stating the de-embedding problem in terms of chain rather that scattering matrices, the latter becomes linear,
- For a generic junction the de-embedding problem has a unique solution for $N = 2$ and $N > 3$ while it is degenerated for $N = 3$,
- Degeneracy depends on the junction’s response and can be explained by the occurrence of filtering responses that hide behind the junction. These are responses $F_1, \cdots F_{N-1}$ such that when chained to the junction, the response $T$ remains unchanged.

In the light of these theoretical results the tractability of the practical problem is currently under investigation. For example, it appears that in practice the $T$ responses can not be considered as generic as they are often build by chaining several 3-junctions together. Moreover, while theoretically uniqueness of the de-embedding solution seems to be guaranteed for large values of $N$, the sensitivity of the latter with respect to measurements errors grows dramatically with $N$. Therefore approaches that entail additional hypotheses, such as rationality of the responses $F_k$ are being studied. This work is pursued in collaboration with Thales Alenia Space and the Politecnico di Milano.

6.8. Averaging in control

Participants: Bernard Bonnard, Jean-Baptiste Pomet, Jana Nemcova.

A reference paper on the construction and properties of an “average control system”, has been submitted [27]; it is based on Alex Bombrun’s doctoral work [54] (defended in 2007). It connects properties of convergence of solutions of highly oscillating control systems to those of an average control system, when the frequency of oscillation goes high. Likewise, it details (on a time-interval that goes to infinity) the properties of solutions...
of a conservative system with small controls in relation to those of an average system as the magnitude of control goes to zero. It also gives many properties of this average control system that has “more controls” than the original system, and yields, when this number of new controls is maximal, a Finsler metric on the state manifold. It is however difficult to compute explicitly and is never twice differentiable.

In [33], we study into details this average system arising from low-thrust orbital transfer, in the restricted “meridian” co-planar case, and prove that its trajectories for minimum time never leave the “elliptic domain” where averaging is valid. This gives some ground to using it as a limit to describe transfer from an elliptic orbit to another.

More exploration on this average system and the corresponding Finsler metric is planned.

6.9. Optimal transport

**Participants:** Bernard Bonnard, Ahed Hindawi, Jean-Baptiste Pomet, Ludovic Rifford.

In [23], we provide results on existence and continuity of the optimal transport map between absolutely continuous measures (with respect to Lebesgue measure) for a point-to-point cost coming from controllable linear-quadratic optimal control. This is the simplest case of a cost coming from a system with drift; it bears some interest because the cost can be explicitly computed. A. Hindawi’s PhD also aims at more general systems. It will be defended in 2012.

6.10. Detection of the instability of amplifiers

**Participants:** Laurent Baratchart, Sylvain Chevillard, Martine Olivi, Jean-Baptiste Pomet, Fabien Seyfert.

A new collaboration with the CNES and the University of Bilbao began this year. The goal is to help the development of amplifiers, in particular to detect unstability at an early stage of the design.

Currently, Electrical Engineers from the University of Bilbao, under contract with CNES (the French Space Agency), use heuristics to diagnose instability before the circuit is physically implemented. We intend to set up a rigorously founded algorithm instead, based on properties of transfer function of such amplifiers which belong to particular classes of analytic functions.

We completed the first stage of this collaboration, in that we now have a formal definition of stability within these classes: a stable function is one which, when connected in parallel with a large resistor, yields a $H^\infty$ transfer function (in this context transfer functions might have infinitely many poles). We also obtained a characterization of transfer functions that may actually be realized in terms of elementary electric components (i.e., resistors, self, capacities, LEDs, and transmission lines): they are the rational functions in the variable and the exponentials thereof.

6.11. Tools for numerically guaranteed computations

**Participant:** Sylvain Chevillard.

This work has been performed in collaboration with members of the teams Arénaire in Lyon and Caramel in Nancy. The overall and long-term goal is to enhance the quality of numerical computations. Several aspects are studied:

- A first topic is the development of software code for the multiprecision evaluation of elementary and special functions. Developing such codes is a long and error-prone task. It is hence relevant to automatically generate such codes whenever possible. A first step has been to design an algorithm that automatically generates multiprecision code for the evaluation of constant expressions with an a priori guaranteed error [25]. This is usually necessary for the evaluation, e.g., of the first terms of a Taylor series.
Another topic consists in the design of algorithms that allow developers of double precision mathematical libraries (so-called libm) to certify their library. In the process of developing a libm, one usually replaces the function \( f \) to be evaluated by a good polynomial approximation \( p \). In order to certify the quality of the library, it is then necessary to give a rigorous mathematical proof that the relative error \( \varepsilon = (p - f)/f \) between \( p \) and \( f \) is bounded by a small constant. This turns out to be equivalent to the problem of computing a sharp yet certified upper bound of the supremum norm of \( \varepsilon \). An efficient algorithm has been designed for this purpose [20] (this work is the publication of a work initially begun in the Arénaire team and continued in the Caramel team).

Finally, a more general endeavor is to develop a tool that helps developers of libms in their task. This is performed by the software Sollya, which has originally been developed in the Arénaire team, in collaboration with C. Lauter and M. Joldeş. A new release has been performed this year [32].
5. New Results

5.1. Reactive trajectories for molecular dynamics

**Participants:** Frédéric Cérou, Arnaud Guyader, Florent Malrieu.

*See 3.3 and 4.2.*

This is a collaboration with Tony Lelièvre (CERMICS, Ecole des Ponts ParisTech).

Consider a one-dimensional Brownian motion in a double well potential. It is known that, as its variance goes to zero, the Brownian particle has to wait for a longer and longer time to jump from a well to the other. This metastable behavior is described by the Freidlin–Wentzell theory. We are investigating the length of the paths between the last passage time in the bottom of a well and the hitting time of the other one (reactive trajectory). In the case of an Ornstein–Uhlenbeck process between the wells, we obtained the remarkable result that the time length of the reactive trajectories converges in distribution when the noise intensity goes to zero, to a Gumbel random variable shifted by a deterministic term growing like minus the logarithm of the noise intensity. Our numerical simulations are also in good accordance with this theoretical result. We are also very close to extend this result to more general one dimensional diffusion processes.

5.2. Learning the optimal importance distribution

**Participants:** François Le Gland, Rudy Pastel.

*See 3.3 and 4.2.*

This is a collaboration with Jérôme Morio (ONERA Palaiseau).

As explained in 3.3, multilevel splitting ideas can be useful even to solve some static problems, such as evaluating the (small) probability that a random variable exceeds some (extreme) threshold. Incidentally, it can also be noticed that a population of particles is available at each stage of the algorithm, which is distributed according to the original distribution conditioned on exceeding the current level. Furthermore, this conditional distribution is known to be precisely the optimal importance distribution for evaluating the probability of exceeding the current level. In other words, the optimal importance distribution is learned automatically by the algorithm, as a by-product, and therefore can be used to produce an importance sampling estimate with very low variance. Building on this idea, several other iterative methods have been studied, that learn the optimal importance distribution at each stage of the algorithm, such as nonparametric adaptive importance sampling (NAIS) [69], or the cross–entropy (CE) method [65], [33]. These methods have been applied to a practical example from the aerospace industry, the evaluation of collision probabilities between two satellites, or between a satellite and space debris.

5.3. Impact of dimension in particle filtering and the Laplace method

**Participants:** François Le Gland, Paul Bui–Quang.

*See 3.1.*

This is a collaboration with Christian Musso (ONERA Palaiseau).
Particle filtering is a widely used Monte Carlo method to approximate the posterior probability distribution in non-linear filtering, with an error scaling as $1/\sqrt{N}$ in terms of the sample size $N$, but otherwise independently of the underlying state dimension. However, it has been observed for a long time in practice that particle filtering can be quite inefficient when the dimension of the system is high. In a simple static linear Gaussian model, it has been possible indeed to check that the error on the estimation of the predicted likelihood, a quantitative indicator of the consistency between the prior distribution and the likelihood function, increases exponentially with the dimension [30]. This preliminary result has been extended to a non-linear / non-Gaussian model, using the Laplace method [23]. The Laplace method, which approximates multidimensional integrals accurately, has also been used to compute the asymptotic variance of the importance weights, as the sample size $N$ goes to infinity, and to analyze its dependence on the dimension.

5.4. Stochastic wind–wave modelling

Participant: Valérie Monbet.

This is a collaboration with Pierre Ailliot (UBO) and Christophe Maisondieu (IFREMER)

Directional wave spectra generally exhibit several peaks due to the coexistence of wind sea generated by local wind conditions and swells originating from distant weather systems. This paper proposes a new algorithm for partitioning such spectra and retrieving the various systems which compose a complex sea-state. It is based on a sequential Monte–Carlo algorithm which allows to follow the time evolution of the various systems. The main particularity of the algorithm is that the dimension of the hidden state can change from time to time and a model selection step is included. The proposed methodology is validated on both synthetic and real spectra and the results are compared with a method commonly used in the literature.

5.5. Sequential data assimilation: ensemble Kalman filter vs. particle filter

Participants: François Le Gland, Valérie Monbet.

See 6.4.

Surprisingly, very little was known about the asymptotic behaviour of the ensemble Kalman filter [38], [39], [40], whereas on the other hand, the asymptotic behaviour of many different classes of particle filters is well understood, as the number of particles goes to infinity. Interpreting the ensemble elements as a population of particles with mean–field interactions, and not only as an instrumental device producing an estimation of the hidden state as the ensemble mean value, it has been possible to prove the convergence of the ensemble Kalman filter, with a rate of order $1/\sqrt{N}$, as the number $N$ of ensemble elements increases to infinity [25]. In addition, the limit of the empirical distribution of the ensemble elements has been exhibited, which differs from the usual Bayesian filter. The next step has been to prove (by induction) the asymptotic normality of the estimation error, i.e. to prove a central limit theorem for the ensemble Kalman filter.
6. New Results

6.1. Numerical schemes and algorithms for fluid mechanics.

Participants: Rémi Abgrall [Corresponding member], Guillaume Baurin, Pietro Marco Congedo, Cécile Dobrzynski, Marc Duruflé, Dante De Santis, Algiane Froehly, Gianluca Geraci, Robin Huart, Arnaud Krust, Cédric Lachat, Mario Ricchiuto, Birte Schmidtman, Héloïse Beaugendre, Sébastien Blaise.

6.1.1. Residual distribution schemes

This year, many developments have been conducted and implemented in the Realfluids and Slowswss software after the initial ideas discussed in [55] and in [64], [61], [65], which have opened up many doors.

First of all, the parallel three dimensional high order extension of the scheme of [55] has been finally validated on several external aerodynamics configurations [3], including the classical ONERA M6 wing case on a large mesh containing $5.5 \times 10^6$ vertices (simulation run on 256 processors), and a business jet configuration in supersonic conditions, on a mesh obtained by the GAMMA3 EPI.

Meanwhile, the improvement of the treatment of viscous terms has been investigated within the PhD theses of G. Baurin and D. DeSantis [46], [19]. The validation on laminar flows of a classical formulation based on a Petrov-Galerkin approach [60] [17] has shown its limitations. An improved formulation, based on a recovery of the solution gradient, has been proposed and tested. In both the second and third order cases, while showing the improvement in accuracy for steady state laminar flows, the results also show a slow iterative convergence, and a systematic small accuracy drop when the cell Reynolds number is of order one. These issues are currently under investigations, while the current formulation is being enhanced by adding a Spalart Almaras turbulent model. Contributions to these activities come from the PhD of Guillaume Baurin, who has implemented the third order version of our methods in a real industrial platform (N3S Natur of SAFRAN developed by Incka), and from the PhD of Dante DeSantis who is developing the turbulent implementation in Realfluids within the EU project IDIHOM.

Meanwhile, we are refining and validating the extension of the schemes to elements using improved approximations based on Bézier and NURBS polynomials. The initial two-dimensional implementation [56], [59] of third and fourth order schemes on curved meshes is now being enhanced by adding a local mesh refinement procedure and is also currently being extended to three space dimensions. Contributions to this topic come from the PhD of Algiane Froehly.

R. Abgrall has extended the RDS formalism to Lagrangian hydrodynamics. The results are comparable to what can be obtained for more standard methods, a publication is in preparation.

Concerning time dependent flows, the ideas of [61], [64] have led to two main lines of developments. On one hand, the unconditionally second order and stable space-time approach of [61] has been further validated [14] and extended to higher orders of accuracy [51]. The main advantage of this technique is its ability to preserve monotonicity unconditionally w.r.t. the time step. This has interesting applications in shallow water flows [36] in which the schemes previously developed [65] did allow to preserve the positivity of the water depth, however with an inefficient implicit procedure constrained by an explicit-type time step restriction.

In parallel, the genuinely explicit formulation of [64] has been combined with the positivity preserving approach of [65] to obtain a genuinely explicit positivity preserving scheme for shallow water simulations [16]. With a time step restriction quite close to that necessary for the scheme of [65], the approach proposed allows a very efficient explicit time stepping with a tenfold reduction of the computational time for the same accuracy level.
These developments are implemented in the SLOWS platform and are thoroughly summarized in the manuscript [2]. We now dispose of a spectrum of numerical tools allowing either classical temporal integration based on implicit multistep schemes, or on unconditionally stable and positivity preserving space-time schemes, or on a genuinely explicit approach. Current developments aim at extending these tools to arbitrary accuracy, and at developing hybrid implicit/explicit approaches.


In this work, Héloïse Beaugendre, Boniface Nkonga and Christelle Wervaecke proposed a strongly coupled numerical formulation for the Spalart-Allmaras model, in the framework of Stabilized Finite Element Methods. Computations are performed for compressible Newtonian fluids (2D and 3D) on unstructured grids of high aspect ratio. Results are compared with experimental data and also with solutions obtained by different numerical strategies. The additional transport equations for subscale model are often numerically weakly coupled to Navier-Stokes equations through operator splitting. These variables are strongly coupled for the transport process within a stabilized finite element formulation. The stabilization tensor is defined, such as to reduce mesh dependencies and to still be consistent at the asymptotic of highly anisotropic meshes. Indeed, this tensor involves a measure of the local length scale $h$ which can be difficult to define in the case of a stretched element. In this work, the local length scale is implicitly given by the inverse of the absolute flux Jacobian matrix as proposed in Barth (1998) and more recently in Abgrall (2006). The stabilized finite element strategy is also suitable for complex geometries and the resulting schemes have a compact stencil which we exploit for efficient parallel strategies combining domain decomposition and message passing tools (MPI).

6.1.3. Uncertainty quantification

R. Abgrall and P.M. Congedo have made a detailed comparison between the semi-intrusive method developed last year with more classical non intrusive polynomial chaos methods, and Monte Carlo results. The effectiveness of this method is illustrated for a modified version of Kraichnan-Orszag three-mode problem where a discontinuous pdf is associated to the stochastic variable, and for a nozzle flow with shocks. The results have been analyzed in terms of accuracy and probability measure flexibility. Finally, the importance of the probabilistic reconstruction in the stochastic space is shown up on an example where the exact solution is computable, the viscous Burgers equation. These results have been reported in [25], [47]. Following this studies, two contributions have been obtained within the context of Gianluca Geraci’s thesis. First one is an adaptive strategy, inspired by the Harten multi-resolution framework that has been developed in order to compute efficiently statistics. This preliminary work aims to show the potentialities of this approach in order to evaluate the possibility to include this strategy in the semi-intrusive method developed in the recent years. We obtained [34] well-converged results with a lower computational cost due to a reduction of the numerical evaluations. Second contribution [48] is a study concerning the Sparse Grid techniques coupled with Polynomial Chaos for multi-dimensional stochastic problems. Sparse grid techniques have been used to compute the multi-dimensional integrals needed to evaluate the coefficients of the polynomial expansion. We also investigated the possibility to reduce the number of random variables by means of an ANOVA analysis. P.M. Congedo investigated the possibility to perform a stochastic inverse analysis by using an hybrid method within a Polynomial Chaos/Genetic Algorithms framework. This strategy has been applied on the numerical simulation of a dense gas shock-tube. Previous theoretical and numerical studies have shown that a rarefaction shock wave (RSW) is relatively weak and that the prediction of its occurrence and intensity are highly sensitive to uncertainties on the initial flow conditions and on the fluid thermodynamic model. The objective of this work has been to introduce an innovative, flexible and efficient algorithm combining computational fluid dynamics (CFD), uncertainty quantification (UQ) tools and metamodel-based optimization in order to obtain a reliable estimate for the RSW probability of occurrence and to prescribe the experimental accuracy requirements ensuring the reproducibility of the measurements with sufficient confidence.

Uncertainty quantification tools have been used to perform some applicative studies on epistemic uncertainties, in particular on some complex equations of state [8], and some turbulence [12] models. We have also started
considering the influence of model parameters uncertainties in free surface models for long-waves such as tsunamis [52], coupling the numerics developed in the team for shallow-water flows [5.3] and the tools available for uncertainty quantifications. This is certainly a field of application where these developments will demonstrate very useful.

Within the associated team AQUARIUS activities (collaboration with Stanford University), two efficient global strategy for robust optimization have been developed. First one is based on the extension of simplex stochastic collocation to the optimization space, while the second one consists in an hybrid strategy using ANOVA decomposition. The Simplex Stochastic Collocation (SSC) method has been developed for adaptive uncertainty quantification (UQ) in computational problems with random inputs. In this work [30], we showed how this formulation based on Simplex space representation, discretization of non-hypercube probability spaces and adaptive refinements can be easily coupled with a well-known optimization method, i.e. Nelder-Mead algorithm, also known as Downhill Simplex Method. Numerical results showed that this method is very efficient for mono-objective optimization and minimizes global number of deterministic evaluations in order to determine optimal design. This method has been then applied to a realistic problem of robust optimization of a two-component race-car airfoil.

We proposed also an efficient strategy [29] for robust optimization when a large number of uncertainties is taken into account. ANOVA analysis is used in order to perform a variance-based decomposition and to reduce stochastic dimension. A massive use of metamodels allows reconstructing response surfaces for sensitivity indexes and fitness function in the design variables plan. Proposed strategy has been applied to the robust optimization of a turbine cascade for thermodynamically complex flows.

6.1.4. Multiphase flows

Starting from [58], R. Abgrall and H. Kumar are developing a method that is able to compute multiphase flows when the interfacial area takes any value. In the previous version, either we could take into account infinite interfacial areas or pure interface problems. M.G. Rodio is developing, along similar lines, a scheme for the Navier Stokes equations.

6.1.5. Numerical schemes for advanced materials

Two parallel lines of work on developments of numerical models for advanced materials have seen important developments this year.

On one hand, Rémi Abgrall and Pierre-Henri Maire (CEA Cesta) are extending the Lagrangian method developed a couple years ago and currently implemented in the CHIC code to elastodynamics. The stress tensor is no longer diagonal and here we consider the Wilkins model. The main difficulty is to understand the role of the second principle and how to deal with the von Mises criteria.

In parallel, Mario Ricchiuto and the group led by Gérard Vignoles at LCTS (UMR-5801 LCTS) have been developing a finite element numerical model of the evolution of the liquid oxide evolution during the healing-phase taking place in the silicon-based composite materials similar to those used in SAFRAN’s new aero-engines [50]. This micro-model is meant to be used as a numerical closure for the LCTS’ structural mechanics solver [31], allowing to obtain a faithful description of the material’s behavior, including the effects of the healing process.

6.1.6. Discontinuous Galerkin schemes, New elements in DG schemes

Rémi Abgrall and Pierre-Henri Maire (CEA CESTA), with François Vilar (PhD at CELIA funded by a CEA grant started in October 2009), are working on fully Lagrangian schemes within the Discontinuous Galerkin schemes. The idea is to start from the formulation of the Euler equation in full Lagrange coordinates: the spatial derivative are written in Lagrangian coordinates. The mesh element are now curved and we are working on the geometrical conservation law. The application to several standard test case indicate the potential of the method.

6.1.7. Penalization techniques with unstructured adapted meshes

Penalization methods are an efficient alternative to explicitly impose boundary conditions but their accuracy is generally of first order. In this work we combine the easiness of penalization techniques with the precision of unstructured anisotropic mesh adaptation. Level sets are used to describe the geometry so that geometrical and topological changes due to physics are straightforward to follow. Navier-Stokes simulations are performed and a new way to impose a slipping wall boundary condition is proposed.

6.1.8. Mesh adaptation

A work on high order mesh generation has been pushed further. Starting with a $P^1$ (triangle) mesh and some information on the boundary (control point), we are able to generate a valid third order curved mesh. The algorithm is based on edge swaps and is similar to a boundary enforcement procedures. This method is very robust but not efficient of the boundary layer. Indeed the edge swaps destroy a part of the boundary layer. To solve this problem, we investigate the use of linear elasticity to curve a $P^1$ mesh.

Moreover, we started to make high order mesh adaptation. That means we are able to refine high order meshes where the error is maximum and so we generate non uniform meshes of order $k$ with $k > 2$. We compute Euler compressible simulations on those meshes to validate the mesh adaptation strategy.

In parallel to these developments, we have started work on a generalized formal approach to obtain discrete adjoint equations for residual based and Petrov-Galerkin finite element schemes. We have shown that these discrete adjoint equations can now be used as a local error estimator for mesh refinement, giving to these methods the same potential for adaptation of Galerkin schemes.

6.2. High performance simulation for plasma physics

Participants: Rémi Abgrall, Robin Huart, Xavier Lacoste, François Pellegrini, Pierre Ramet [Corresponding member].

In the Realfluids code, the Rusanov scheme for Ideal MHD has already shown its ability to capture discontinuities and its robustness many times in 2D problems. But other spatial schemes could be interesting for applications in tokamak experiments, since we may not encounter strong shocks in these cases. Hence, according to the type of the problem, coupled schemes could be used. We already developed the 4 well known base RD schemes: Narrow, LDA, Rusanov and SU (a RD version of the SUPG scheme). Coupling may not be challenging, a working shock sensor is already implemented for stabilized methods. Very high order of accuracy (at least 3rd order) should be reachable in all cases, the main parts of this work have already been done for several types of elements. The non-dimensionalized equations of resistive MHD (with viscosity and heat transfer) have been added to the code with a Continuous Galerkin discretization. Also, 2nd order implicit and explicit methods were developed in all cases. Once we succeed in ensuring a very good iterative convergence, taking into account the hyperbolic divergence cleaning technique in an unsteady context, we will be able to simulate plasma instabilities. This is really the key issue for now. These results will be presented in the PhD defense of R. Huart planned at January 2012.

The JOREK code is now able to use several hundred of processors routinely. Simulations of ELMs are produced taking into account the X-point geometry with both closed and open field lines. But a higher toroidal resolution is required for the resolution of the fine scale filaments that form during the ELM instability. The complexity of the tokamak’s geometry and the fine mesh that is required leads to prohibitive memory requirements. In the current release, the memory scaling is not satisfactory: as one increases the number of processes for a given problem size, the memory footprint on each process does not reduce as much as one can expect.

In the context of the new ANR proposal (ANEMOS project), we are working to reduce memory consumers in the JOREK code. Compression techniques can be foreseen to reduce the footprint of the matrix without having to pay large computation expenses. Moreover, the storage of the factorized preconditioning matrix inside the direct solver takes also a large amount of memory. We have defined and developed a general programming interface for sparse linear solvers for which we also provided some test programs and documentation. Our goal is to normalize the application programming interface of
sparse linear solvers and to provide some very simple ways of doing some fastidious tasks such as parallel matrix assembly for instance. This interface has been validated in RealfluïDS and JOREK for HIPS and PaStiX. Using this common interface, we are looking for a fair distribution of data over the parallel processes in order to reduce memory consumption. The effective parallelization of this assembly step is one of the main bottlenecks up to now, as far as memory usage is concerned. The GMRES driver is also a large consumer in terms of memory and we plan to consider an up-to-date parallel implementation of this step.

6.3. Algorithms and high-performance solvers

Participants: Astrid Casadei, Cécile Dobryniski, Sébastien Fourestier, Damien Genêt, Hervé Guillard ([Pumas]), Laurent Hascoët ([Tropics]), Cédric Lachat, Xavier Lacoste, François Pellegrini, Pierre Ramet [Corresponding member].

6.3.1. Parallel domain decomposition and sparse matrix reordering

Most of the work carried out within the Scotch project (see section 5.7) has been carried out in the context of the PhD of Sébastien Fourestier.

The first axis concerns dynamic repartitioning and remapping. A new set of sequential routines has been devised, which offers new features: mapping (including plain partitioning) with fixed vertices, remapping, and remapping with fixed vertices. All of the above developments are about to be released in the major release 6.0 of Scotch. The porting of the remapping algorithms in parallel is being carried out, and will be part of release 6.1.

A work carried out in the Joint Laboratory for Petascale Computing (JLPC) between INRIA and UIUC resulted in the inclusion of Scotch as a load balancer in the Charm++ parallel environment. A jointly written conference paper has been submitted on this subject. Another potential use of the remapping features of Scotch concerns multi-phase mapping. Experiments are being carried out at UIUC regarding the use of Scotch as a multi-phase mapper for the OpenAtom scientific code.

6.3.2. Parallel mesh adaptation

This research topic deals with the design of efficient and scalable software tools for parallel dynamic remeshing. This is a joint work with Cécile Dobryniski, in the context of the PhD of Cédric Lachat, funded by a CORDI-S grant managed by the PUMAS team.

PaMPA (see Section 5.11) is a middleware library dedicated to the management of distributed meshes.

The software development of PaMPA is going on. The internal data structure for representing meshes has been frozen, and developments are in progress. The first developments aimed at proving the efficiency of the planned API for handling distributed meshes. A simple P1 FEM Laplacian solver has been written over PaMPA by the PUMAS team to demonstrate how to iterate over PaMPA entities (elements and nodes) and access values borne by the entities, so as to perform FEM computations. These features are available in version 0.1, which has not yet been diffused to other interested parties. Several new potential users are already willing to try out this version, e.g. ONERA.

PaMPA is already used as the data structure manager for two solvers being developed at INRIA: the Plato solver being developed by the PUMAS team, and the Aerosol new generation fluid dynamics solver being developed in the context of the PhD of Damien Genêt. The interaction with these users allows us to refine the interface to match their needs.

This work now focuses on the core of the PhD of Cédric Lachat: interfacing PaMPA with MMG3D to demonstrate the ability of PaMPA to perform parallel mesh adaptation.
6.3.3. **High-performance direct solvers on multi-platforms**

New supercomputers incorporate many microprocessors which include themselves one or many computational cores. These new architectures induce strongly hierarchical topologies. These are called NUMA architectures. In the context of distributed NUMA architectures, in collaboration with the INRIA RUNTIME team, we study optimization strategies to improve the scheduling of communications, threads and I/O. Sparse direct solvers are a basic building block of many numerical simulation algorithms. We have developed dynamic scheduling designed for NUMA architectures in the PaStiX solver. The data structures of the solver, as well as the patterns of communication have been modified to meet the needs of these architectures and dynamic scheduling. We are also interested in the dynamic adaptation of the computation grain to use efficiently multi-core architectures and shared memory. Experiments on several numerical test cases have been performed to prove the efficiency of the approach on different architectures.

In collaboration with the ICL team from the University of Tennessee, and the RUNTIME team from INRIA, we are evaluating the way to replace the scheduling driver of the PaStiX solver by one of the generic frameworks, DAGuE (see [http://icl.cs.utk.edu/dague/overview/] ) or StarPU (see [http://runtime.bordeaux.inria.fr/StarPU ]), to execute the task graph corresponding to a sparse factorization. This work now focuses on the core of the PhD of Xavier Lacoste, the aim is to study and design algorithms and parallel programming models for implementing direct methods for the solution of sparse linear systems on emerging computer equipped with GPU accelerators. This is a joint work with the MUMPS team and algorithms will have to be adapted in order to exhibit parallelism that can be more suitable for the dynamic scheduling of computational tasks on such heterogeneous architectures.

6.3.4. **Hybrid direct-iterative solver based on a Schur complement approach.**

In HIPS, we propose several algorithmic variants to solve the Schur complement system that can be adapted to the geometry of the problem: typically some strategies are more suitable for systems coming from a 2D problem discretisation and others for a 3D problem; the choice of the method also depends on the numerical difficulty of the problem. We have a parallel version of HIPS that provides full iterative methods as well as hybrid methods that mixes a direct factorization inside the domain and an iterative method in the Schur complement.

Graphs or meshes partitioners are now able to deal with problems that have more than several billion of unknowns. Solving linear systems is clearly the limiting step to reach this challenge in numerical simulations. During her PhD, Astrid Casadei will have to propose solutions to get an efficient algorithmic coupling of direct and iterative methods that allow a powerful management of whole the levels of parallelism. As a preliminary study, we focus on memory issues to build a Schur complement in our direct solver. During factorization step, memory overhead may occur for two reasons. The first one is due to the fan-in approach, that is to say the local storage of non-local contributions. The second overhead is due to the coupling matrices (between direct part and Schur complement), which remain allocated during the whole computation and are freed only at the end. Our first ideas to reduce memory consumption was to postpone the allocation of each block and, thanks to a right-looking algorithm, a column-block may be freed as soon as it has been treated. However, many blocks may be allocated very quickly, and a solution would be to use a left-looking scheme when dealing with local contributions. Thus, we introduce a mixed version : a right-looking algorithm is used, except for local contributions in the direct part where a left-looking scheme is applied. Some experiments have been performed and first results show that some substantial memory reductions can be achieved.
BIPOP Project-Team

6. New Results

6.1. Modeling

6.1.1. Simulation of electrical circuits as nonsmooth dynamical systems

Participants: Vincent Acary, Olivier Bonnefon, Bernard Brogliato.

DC-DC converters are usually difficult to simulate with classical tools like SPICE because of the highly nonlinear behaviour of some components and the frequent occurrence of intrinsically generated switching events.

The simulation of such circuits modelled as nonsmooth systems has been successfully achieved with a clear advantage over several SPICE simulators and a simulator belonging to the hybrid modelling approach [1][48].

6.1.2. Spiking neuronal networks dynamics

Participant: Arnaud Tonnelier.

Precise spatiotemporal sequences of spikes are observed in many neural systems and are thought to be involved in the neural processing of sensory stimuli. In [58] we examine the capability of spiking neural networks to propagate stably spatiotemporal sequences of spikes. We derive some analytical results for the wave speed and show that the stability of simple waves is determined by the Schur criteria. The transmission of a sequence of several spikes is related to the existence of stable composite waves, i.e. the existence of stable spatiotemporal periodic traveling waves. We show that the stability of composite waves is related to the roots of a system of multivariate polynomials.

A fundamental aspect that shapes the properties of traveling waves in networks is the underlying lattice-structure of the space. Discreteness has a strong effect on propagating activity patterns and, for instance, anisotropy or propagation failure can be observed. Numerical simulations and analytical calculations have been carried out to characterize more precisely these properties [47].

6.1.3. Computational Toxicology

Participant: Arnaud Tonnelier.

It is now well recognized that toxicology has entered a new era. Previously mainly based on animal testing, toxicology is now turning to in vitro and in silico experiments. To assess the risk of chemicals but also to gather and to interpret the massive amounts of experimental data generated by modern toxicology, the development of mathematical and computational tools are essential. An important element in risk assessment of chemicals is the human bioaccumulative potential. We developed a predictive tool for human bioaccumulation assessment using a physiologically based toxicokinetic model [28].

6.1.4. High-order models of mechanical rods

Participants: Florence Bertails-Descoubes, Romain Casati.

Reduced-coordinates models for rods such as the articulated rigid body model or the super-helix model [50] are able to capture the bending and twisting deformations of thin elastic rods while strictly and robustly avoiding stretching deformations. In this work we are exploring new reduced-coordinates models based on a higher-order geometry. Typically, elements are defined by a polynomial curvature function of the arc length, of degree $d \geq 1$. The main difficulty compared to the super-helix model (where $d = 0$) is that the kinematics has no longer a closed form. We have already investigated the clothoidal case ($d = 1$) in the 2d case [51], relying on Romberg numerical integration, and a general approach in 3d based on power series expansion was formulated in the master thesis of R. Casati, for a single element. R. Casati is currently extending the method to a chain of linked elements as well as to an arbitrary degree $d$ of the curvature function.
6.1.5. Inverse modeling of mechanical rods


Controlling the input shape of slender structures such as rods is desirable in many design applications (such as hairstyling, reverse engineering, etc.), but solving the corresponding inverse problem is not straightforward. In [29] we started to extend to 3d our 2d method introduced in [8] for automatically converting a smooth sketched curve into a dynamic curve at stable equilibrium under gravity. The main challenge in 3d amounts to converting an input curve into a continuous piecewise helix. Using a least-squares optimization approach is a natural option, however it may suffer from both robustness and computational issues due to the presence of multiple local minima in the objective function. To overcome these issues, we have recently proposed to reformulate the problem as a geometric interpolation problem. In this new method, only tangents are strictly interpolated while points are displaced in an optimal way so as to lie in a feasible configuration, i.e., a configuration that is compatible with the interpolation by a helix. Our method proves to be much more robust and faster compared to the global optimization approach. We plan to publish these results in 2012.

6.1.6. Multiple impacts modelling

Participants: Bernard Brogliato, Hongjian Zhang, Ngoc-Son Nguyen.

The work consists of studying two systems: the rocking block and tapered chains of balls, using the Darboux-Keller model of multiple impacts previously developed. The objectives are threefold: 1) show that the model predicts well the motion by careful comparisons with experimental data found in the literature, 2) study the system’s dynamics and extract critical kinetic angles that allow the engineer to predict the system’s gross motion, 3) develop numerical code inside the SICONOS platform that incorporates the model of multiple impact. Results are in [42].

6.1.7. Simulating contact with Coulomb friction in fiber assemblies

Participants: Florence Bertails-Descoubes, Gilles Daviet.

We have developed a new frictional contact solver in [21] which is able to robustly and efficiently handle large fiber problems composed of thousands self-contacting rods with exact Coulomb friction. The solver relies on a Gauss-Seidel iterative approach, where each local one-contact solver is based on a hybrid strategy. The solution to the one-contact problem is first searched for using a nonsmooth Newton method based upon a generalized Fischer-Burmeister formulation. This primary solver manages to solve the local problem in 99.9% of the cases. When the solver fails to converge to an acceptable solution, the method switches to a more costly but exact solver, based on the $\alpha$-formulation introduced in [39]. This hybrid strategy experimentally allows us to always find a solution to the local problem, which greatly contributes to improve the robustness of the global solver. We have compared our new solver against other solvers of the literature (e.g., damped Newton solvers relying on the Alart-Curnier function) and observed a noticeable gain, both in terms of robustness and computational efficiency.

6.2. Optimization

6.2.1. Nonsmooth analysis and optimization on matrix manifolds

Participant: Jérôme Malick.

Optimization on matrix manifolds is an emerging fields of research in optimization, driven by applications in robotics. We have contributed on two different projects.

- **Numerical efficiency of optimization methods.** Newton method on manifolds would require to compute a geodesic (that is to solve a ODE). It is clear that replacing classical differential-geometric objects with certain approximations, resulting in faster and possibly more robust algorithms. With our colleague P.-A. Absil from the Department of Mathematical Engineering of the École Polytechnique de Louvain (Belgique), we propose in [16] a way to construct “retractions” (a key step when applying optimization algorithms on matrix manifolds) by projecting onto the submanifold. We show
that the operation remains a retraction if the projection is generalized to a projection-like procedure that consists of coming back to the submanifold along “admissible” directions. This theory offers a framework in which previously-proposed retractions can be analyzed, as well as a toolbox for constructing new ones. Illustrations are given for projection-like procedures on some specific manifolds for which we have an explicit, easy-to-compute expression.

- **Towards the application of matrix optimization techniques to spectral manifolds.** Spectral sets are sets of matrices that depend only on the constraints on the eigenvalues: \( S \) is a spectral set if \( S = \lambda^{-1}(C) \) with \( C \) a subset of \( \mathbb{R}^n \). A spectral set \( S \) inherits from properties of the underlying set \( C \), such as convexity. We prove in [46] that the spectral sets associated to smooth manifolds in \( \mathbb{R}^n \) (having some local symmetry) are themselves manifolds in the space of matrices. This result looks simple but generalizes several useful particular cases, and was extremely difficult to prove: we brace together tools from nonsmooth analysis, differential geometry, group theory and spectral analysis.

### 6.2.2. Semidefinite programming and combinatorial optimization

**Participants:** Nathan Krislock, Jérôme Malick.

We have worked with Frederic Roupin (Prof. at Paris XIII) on the use of semidefinite programming to solve combinatorial optimization problems to optimality. Within exact resolution schemes (branch-and-bound), “good” bounds are those with a “good” balance between tightness and computing times.

We proposed a new family of semidefinite bounds for 0-1 quadratic problems with linear or quadratic constraints [26], [54]. An interesting feature is that the final accuracy level is controlled by a real parameter acting like a cursor. This gives ways to trade computing time for a (small) deterioration of the quality of the usual semidefinite bounds, in view of enhancing this efficiency in exact resolution schemes. Extensive numerical comparisons and tests showed the superior quality of our bounds on standard test-problems (unconstrained 0-1 quadratic problems, heaviest k-subgraph problems, and graph bisection problems).

We have embedded the new bounds within branch-and-bound algorithms to solve 2 standard combinatorial optimization problems to optimality.

- **Heaviest k-subgraph problems.** Our algorithm [26] takes advantage of the new bounds to prune very well in the search tree. Its performances are then comparable with the best method (based on convex quadratic relaxation using CPLEX as an engine). In practice, our method works particularly fine on the most difficult instances (with a large number of vertices, small density and small k).

- **Max-cut.** We are working on extending our algorithm to max-cut problems [53]. It dynamically manages polyedral and semidefinite relaxations to outperform the state-of-the-art solver ([56]) on the large test-problems.

### 6.2.3. Marginal prices in electricity production

**Participants:** Claude Lemaréchal, Jérôme Malick, Welington Oliveira, Sofia Zaourar.

Two subjects were involved this year in our ongoing collaboration with EdF.

- **Stabilizing prices.** Unit-commitment optimization problems in electricity production are large-scale, nonconvex and heterogeneous, but they are decomposable by Lagrangian duality. Realistic modeling of technical production constraints makes the dual objective function computed inexact though. An inexact version of the bundle method has been dedicated to tackle this difficulty [52]. However, the computed optimal dual variables show a noisy and unstable behaviour, that could prevent their use as price indicator. We propose a simple and controllable way to stabilize the dual optimal solutions, by penalizing the total variation of the prices [59]. Our illustrations on the daily electricity production optimization of EDF show a striking stabilization at a negligible cost.

- **Accelerating** the solution phase by the so-called disaggregation technique [49], using the fact that (see Activity Report of 2010) the dual objective function is the sum of two terms: one coming from primal cost, one coming from valorization of constraints (plus possibly a third term when price stabilization is present). The resulting CPU time is drastically improved, sometimes divided by 10.
6.3. Control

6.3.1. Digital sliding mode control

Participants: Vincent Acary, Bernard Brogliato.

The problem of digital sliding mode controllers is a long-standing issue not yet satisfactorily solved. We propose in ideas which are inspired from the numerical methods of contact mechanics [2] and which permit a) to suppress the numerical chattering, b) to obtain a smooth stabilization on the sliding surfaces. The work is continued together with Yury Orlov in more general cases where the system is acted upon by disturbances and a disturbance estimation is added [19].

6.3.2. Discrete-time discontinuous systems

Participants: Vincent Acary, Bernard Brogliato, Carmina Georgescu, Scott Greenhalg, Thorsten Schindler.

We focus on some classes of discontinuous dynamical systems like relay systems, linear complementarity systems. The objectives are to show that the time-stepping numerical schemes like Moreau’s algorithm can be used to successfully simulate such systems (like in the case of biological systems like gene networks), and also to study the properties of these schemes for finite nonzero time steps (like preservation of dissipativity properties). Results are in [23], [35]. Further work deals with timestepping schemes for nonsmooth dynamical systems. So far, these schemes are locally of order one both in smooth and nonsmooth segments. This is inefficient for applications with few events like circuit breakers. To consistently improve the behavior during smooth episodes, the traditional schemes are being embedded in time discontinuous Galerkin methods. After establishing the correct mathematical setting, a Petrov-Galerkin distributional differential inclusion is outlined. The bouncing ball example illustrates its capabilities.

6.4. Locomotion analysis

6.4.1. Synchronous imitation of human motion by a humanoid robot

Participants: Mehdi Benallegue, Pierre-Brice Wieber.

Interactions between humans and robots require that each one is able to understand and interpret each other’s actions. From the point of view of the robot, this means: (i) to move in a way that can be naturally interpreted by humans and (ii) to be able to understand the humans’ actions. Studies in Neuroscience in the case of interactions between humans indicate that these two abilities might be tightly linked in the human’s brain: we understand actions when we map the observed action onto our motor representation of the same action [57]. In this work, we consider that the “motor representation” of a task is the control law, and “mapping an observed action” means finding the corresponding control parameter, in an observer-based approach.

Considering a correspondence between two different control laws can be seen a modeling error. This modeling error can be seen as an unknown arbitrary perturbation on the modeled system, or an unknown input sent to the observed system. We developed an observer that can cancel the effects of unknown inputs on the dynamics of discrete time linear systems with unknown inputs. To do so, the observer has to satisfy a delayed invertibility condition and use delayed outputs. In other words, the observer has to wait for several measurement after a given instant to collect enough data to reconstruct the state at that instant.

6.4.2. Hierarchic QP solver

We are working in collaboration with the LAAS-CNRS and the CEA-LIST on solving multi-objective Quadratic Programs with Lexicographic ordering: Hierarchic QPs [25]. The focus this year has been on the regularization of the problem when the Quadratic Program approaches singularities. There is indeed a problem of discontinuity of the solution when reaching such singularities, what’s not a rare event in robotic applications. This discontinuity has been related to the fact that Lexicographically ordered QPs correspond to the limit of weighed multi-objective QPs when weights go singular, and that the regularization is itself a weighting of objectives which goes to a limit when approaching singularities, and those two limit processes interfere. The solution we found so far is to cancel the first limiting process and move back from strict Hierarchic QPs to weighted QPs staying at a small distance from singularity [33]. But this solution is not really satisfying and we have to find a better one.

6.4.3. Numerical modeling of muscle contraction under FES


We have been working in collaboration with the EPI DEMAR in Montpellier on modeling muscle contraction under Functional Electric Stimulation (FES). With respect to the literature in the domain, our contributions are mostly linked to the model of the contractile element, through the introduction of the recruitment at the fibre scale, formalizing the link between FES parameters, recruitment and Calcium signal paths. The resulting controlled model is able to reproduce both short term (twitch) and long term (tetanus) responses. It also matches some of the main properties of the dynamic behaviour of muscles, such as the Hill force-velocity relationship or the instantaneous stiffness of the Mirsky-Parmley model. The specific contribution of the BIPOP team has been on the numerical implementation of the contraction model as a Linear Complementarity Problem (LCP) allowing fast and precise numerical simulations [22].

6.4.4. Modeling of human balance in public transports


In our ongoing collaboration with the IFSTTAR (previously INRETS) on modeling human balance in public transports, we have aggregated biomechanical studies and numerical models proposed in robotics, and compared how they match or mismatch in situations of strong perturbations requiring a step to recover balance. We began developing a specific Model Predictive Control scheme for the prediction of recovery step locations with adaptive step timings, reproducing various balance recovery strategies as observed in humans. Initial results for stepping predictions have been validated against a balance recovery scenario found in the literature [45].

6.4.5. Model Predictive Control for Biped Walking


We improved our Linear MPC-based walking motion generator by incorporating explicitly the robot’s kinematic constraints: polyhedral constraints on the position of the CoM ensure the kinematic feasibility of the generated walking motions for arbitrary vertical motions of the CoM. This more precise kinematic model within the LMPC allowed considering toe rotations in a safer way, considerably improving energy efficiency, naturalness of the motion, and maximal speed.

We proposed a formulation of dynamic constraints for 3D motion through simple bounds on the variables, leading to faster resolution of the corresponding optimization problem. This allowed generating three-dimensional walking on non-planar ground in real-time. Thanks to specifically enforcing leg compliance, this scheme managed additionally to reproduce the natural profiles of the CoM and the contact forces observed in human walking.

We finally refined our numerical scheme for solving Linear MPC problems in walking motion generation. We switched the underlying QP solver to enable reductions of the number of iterations through warmstart, non-empty initial active sets, and obtaining feasible solutions at each iterations, what considerably reduced the computation time, allowing 1 ms feedback loops [30], [32].
6.5. Software development

6.5.1. MECHE toolbox
Participants: Florence Bertails-Descoubes, Gilles Daviet.

The main tool developed in 2011 in the MECHE software was the hybrid iterative solver for Coulomb friction, published in [21]. In 2011, the MECHE software was extensively used to validate this new solver on large data consisting of thousands interacting fibers (subject to tens of thousands frictional contacts). Code parallelization and optimization were performed so as to speed up computations.

6.5.2. Platform development: Siconos
Participants: Vincent Acary, Olivier Bonnefon, Maurice Brémond, Franck Pérignon.

The main achievements for the SICONOS platform are:
1. Automatic serialization of the whole set of classes in SICONOS
2. Improvements and development of a full auto-generated Python wrapper in the SICONOS/Front-End
3. Development of the Siconos/Multi-body library and validation on industrial examples (C60 circuit breaker of Schneider Electric)
4. New algorithms for the resolution of the discrete frictional contact problem
5. Development of \((\theta/\gamma)\)-schemes for first order dynamical systems
6. Development of routines for sliding mode control

6.5.3. AMELIF framework
Participants: Pierre-Brice Wieber, François Keith, Jory Lafaye.

The main improvements to the AMELIF framework developed this year are:

- A new package specific to torque control has been developed, that contains the algorithms required to realize a given motion with a humanoid robot: estimation of contact forces and torque computation (feedforward), feedback methods ensuring the contact force convergence. These algorithms have been tested for two humanoid platforms: the robot Romeo and the robot HRP-2.
- The dynamics algorithm has been improved, based on the expertise coming from the HuMaNS toolbox. Besides, inverse dynamic algorithms and Runge Kutta integration methods have been added.
- Finally, the bridge with the stack-of-tasks framework [55], that computes the inverse kinematics and the inverse dynamics of humanoid systems, has been enhanced to handle the binding with Python. With this framework, it is possible to use the Model Predictive Control algorithm aforementioned and to simulate the behaviour of a humanoid in a dynamic simulation realized by AMELIF. Tests are still in progress.
6. New Results

6.1. Robustness and Tolerance

6.1.1. Cubic B-spline approximation by curve unclamping

Participants: Xiao-Diao Chen, Weiyin Ma, Jean-Claude Paul.

A new approach for cubic B-spline curve approximation is presented. The method produces an approximation cubic B-spline curve tangent to a given curve at a set of selected positions, called tangent points, in a piecewise manner starting from a seed segment. A heuristic method is provided to select the tangent points. The first segment of the approximation cubic B-spline curve can be obtained using an inner point interpolation method, least-squares method or geometric Hermite method as a seed segment. The approximation curve is further extended to other tangent points one by one by curve unclamping. New tangent points can also be added, if necessary, by using the concept of the minimum shape deformation angle of an inner point for better approximation. Numerical examples show that the new method is effective in approximating a given curve and is efficient in computation [2].

6.1.2. Computing the Hausdorff distance between two B-spline curves

Participants: Xiao-Diao Chen, Weiyin Ma, Gang Xu, Jean-Claude Paul.

This paper presents a geometric pruning method for computing the Hausdorff distance between two B-spline curves. It presents a heuristic method for obtaining the one-sided Hausdorff distance in some interval as a lower bound of the Hausdorff distance, which is also possibly the exact Hausdorff distance. Then, an estimation of the upper bound of the Hausdorff distance in an sub-interval is given, which is used to eliminate the sub-intervals whose upper bounds are smaller than the present lower bound. The conditions whether the Hausdorff distance occurs at an end point of the two curves are also provided. These conditions are used to turn the Hausdorff distance computation problem between two curves into a minimum or maximum distance computation problem between a point and a curve, which can be solved well. A pruning technique based on several other elimination criteria is utilized to improve the efficiency of the new method. Numerical examples illustrate the efficiency and the robustness of the new method [3].

6.1.3. Surface area estimation of digitized 3D objects using quasi-Monte Carlo methods

Participants: Yu-Shen Liu, Jing Yi, Hu Zhang, Guo-Qin Zheng, Jean-Claude Paul.

A novel and efficient quasi-Monte Carlo method for estimating the surface area of digitized 3D objects in the volumetric representation is presented. It operates directly on the original digitized objects without any surface reconstruction procedure. Based on the Cauchy-Crofton formula from integral geometry, the method estimates the surface area of a volumetric object by counting the number of intersection points between the object’s boundary surface and a set of uniformly distributed lines generated with low-discrepancy sequences. Using a clustering technique, we also propose an effective algorithm for computing the intersection of a line with the boundary surface of volumetric objects. A number of digitized objects are used to evaluate the performance of the new method for surface area measurement [5].

6.1.4. Reverse Engineering for NC Machining Simulation

Participants: Nabil Anwer, Yi-Jun Yang, Haibi Zhao, Olivier Coma, Jean-Claude Paul.
Reverse engineering for NC Machining simulation is becoming an important component of NC simulation and verification. Design engineers need more accurate and complete CAD model of the simulated machined part for finite element analysis or parametric feature-based modeling for design modification or update. The as-cut or in process geometry should be correctly accessed in the CAD/CAM environment at any stage of the machining process. Few commercial software are addressing the reverse engineering issue and provide robust solutions. Until now, in process CAD models for NC simulation have been created with many drawbacks and inaccurate methods are proposed. Reverse engineering for NC machining simulation based on polyhedral in-process geometry is addressed. Two complementary approaches are presented here. An enriched representation embedded in “Spring technologies Reverse engineering” or SRE file format enables to convert the polyhedral model to STEP file and a discrete shape recognition and segmentation approach provides a promising issue thanks to discrete differential geometry [43].

6.1.5. Projection of curves on B-spline surfaces using quadratic reparameterization

**Participants:** Yi-Jun Yang, Wei Zeng, Hui Zhang, Jun-Hai Yong, Jean-Claude Paul.

Curves on surfaces play an important role in computer aided geometric design. In this paper, we present a hyperbola approximation method based on the quadratic reparameterization of Bezier surfaces, which generates reasonable low degree curves lying completely on the surfaces by using iso-parameter curves of the reparameterized surfaces. The Hausdorff distance between the projected curve and the original curve is controlled under the user-specified distance tolerance. The projected curve is $G^1$ continuous, where $T$ is the user-specified angle tolerance. Examples are given to show the performance of our algorithm [64].

6.1.6. A kind of parametric transform for trimmed surfaces

**Participants:** Sheng Yang, Jun-Hai Yong.

An approach for a kind of parametric transform is presented for trimmed parametric surfaces. Firstly, the characteristics of trimmed surface before and after parametric transform are evaluated. Then, an algorithm is proposed to adjust the geometric and topological data of a trimmed surface in order to achieve consistency. At last, the trimmed sphere surface is taken as example to further illustrate the algorithm [63].

6.1.7. A cell-based algorithm for evaluating directional distances in GIS

**Participants:** Sheng Yang, Jun-Hai Yong, Jia-Guang Sun, He-Jin Gu, Jean-Claude Paul.

Directional distance is commonly used in geographical information systems as a measure of openness. In previous works, the sweep line method and the interval tree method have been employed to evaluate the directional distances on vector maps. Both methods require rotating original maps and study points in every direction of interest. In this article, we propose a cell-based algorithm that pre-processes a map only once; that is, it subdivides the map into a group of uniform-sized cells and records each borderline of the map into the cells traversed by its corresponding line segment. Based on the pre-processing result, the neighbouring borderlines of a study point can be directly obtained through the neighbouring cells of the point, and the borderlines in a definite direction can be simply acquired through the cells traversed by the half line as well. As a result, the processing step does not need to enumerate all the borderlines of the map when determining whether a point is on a borderline or finding the nearest intersection between a half line and the borderlines. Furthermore, we implement the algorithm for determining fetch length in coastal environment. Once the pre-processing is done, the algorithm can work in a complex archipelago environment such as to calculate the fetch lengths in multiple directions, to determine the inclusion property of a point, and to deal with the singularity of a study point on a borderline [61].

6.1.8. A point-in-polygon method based on a quasi-closest point

**Participants:** Sheng Yang, Jun-Hai Yong, Jia-Guang Sun, He-Jin Gu, Jean-Claude Paul.
This paper presents a numerically stable solution to a point-in-polygon problem by combining the orientation method and the uniform subdivision technique. We define first a quasi-closest point that can be locally found through the uniform subdivision cells, and then we provide the criteria for determining whether a point lies inside a polygon according to the quasi-closest point. For a large number of points to be tested against the same polygon, the criteria are employed to determine the inclusion property of an empty cell as well as a test point. The experimental tests show that the new method resolves the singularity of a test point on an edge without loss of efficiency. The GIS case study also demonstrates the capability of the method to identify which region contains a test point in a map [62].

6.1.9. Using diffusion distances for flexible molecular shape comparison
Participants: Yu-Shen Liu, Qi Li, Guo-Qin Zheng, Karthik Ramani, William Benjamin.

Background: Many molecules are flexible and undergo significant shape deformation as part of their function, and yet most existing molecular shape comparison (MSC) methods treat them as rigid bodies, which may lead to incorrect shape recognition. In this paper, we present a new shape descriptor, named Diffusion Distance Shape Descriptor (DDSD), for comparing 3D shapes of flexible molecules. The diffusion distance in our work is considered as an average length of paths connecting two landmark points on the molecular shape in a sense of inner distances. The diffusion distance is robust to flexible shape deformation, in particular to topological changes, and it reflects well the molecular structure and deformation without explicit decomposition. Our DDSD is stored as a histogram which is a probability distribution of diffusion distances between all sample point pairs on the molecular surface. Finally, the problem of flexible MSC is reduced to comparison of DDSD histograms [50].

6.1.10. A fast sweeping method for computing geodesics on triangular manifolds
Participants: Song-Gang Xu, Yun-Xiang Zhang.

A wide range of applications in computer intelligence and computer graphics require computing geodesics accurately and efficiently. The fast marching method (FMM) is widely used to solve this problem, of which the complexity is $O(N \log N)$, where $N$ is the total number of nodes on the manifold. A fast sweeping method (FSM) is proposed and applied on arbitrary triangular manifolds of which the complexity is reduced to $O(N)$. By traversing the undigraph, four orderings are built to produce two groups of interfering waves, which cover all directions of characteristics. The correctness of this method is proved by analyzing the coverage of characteristics. The convergence and error estimation are also presented.

6.1.11. A torus patch approximation approach for point projection on implicit surface
Participants: Xiao-Ming Liu, Lei Yang.

Point projection on an implicit surface is essential for geometric modeling and graphics applications of it. This paper presents a method for computing the principle curvatures and principle directions of an implicit surface. Using the principle curvatures and principle directions, we construct a torus patch to approximate the implicit surface locally. The torus patch is second order osculating to the implicit surface. By taking advantage of the approximation torus patch, this paper develops a second order geometric iterative algorithm for point projection on the implicit surface. Experiments illustrate the efficiency and less dependency on initial values of our algorithm [52].

6.1.12. Shape similarity assessment approach for CAD models based on graph edit distance
Participants: Bin Wang, Dong Li.

This paper proposes a new shape similarity assessment approach for CAD models in Boundary Representation (Brep) based on graph edit distance. A suboptimal computational procedure is performed to find the best alignment between local structures sets of attributed graphs derived from models. Assuming that only a minority of local structures characterize the functionality, we figure out the weight of every local structure in the query model through a training phase, and then evaluate the similarity between two models by calculating the weighted graph edit distance of corresponding attributed graphs. Experiment results show that our method provides solid retrieval performance on a real-world CAD model database [56].
6.1.13. Reconstructing 3D objects from 2D sectional views of engineering drawings using volume-based method

Participants: Yamei Wen, Hui Zhang, Zhongmian Yu, Jia-Guang Sun, Jean-Claude Paul.

Sectional views are widely used in engineering practice due to their clear and concise expression. However, it is difficult for computers to understand because of the large numbers of omitted entities and their diversified representations. This paper aims at reconstructing 3D models from 2D sectional views by improving the traditional volume based method. First, we present a two-stage loop searching algorithm to extract desired loops from sectional views. Then, sub-objects are identified by the hint-based feature identification algorithm with an intuitive loop-matching criterion. After that, a model-directed algorithm is proposed to guide the generation of sub-objects which are assembled together to form the final objects. The algorithm can handle full sections, partial sections and offset sections, as well as orthographic views. Multiple sectional views are supported in our algorithm. Moreover, the domain of objects is extended to inclined quadric surfaces and intersecting quadric surfaces with higher order curves. Experiment results show its practicability [60].


Participants: Wen-Ke Wang, Hui Zhang, Xiao-Ming Liu, Jean-Claude Paul.

This paper presents a necessary and sufficient condition to judge whether two cubic Beziér curves are coincident. For two cubic Beziér curves whose control points are not collinear, they are coincident if and only if their corresponding control points are coincident or one curve is the reversal of another curve. However, this is not true for the degree that is higher than 3. This paper provides a set of counterexamples of degree 4 [28].

6.1.15. Registration of point clouds using sample-sphere and adaptive distance restriction

Participants: Yu Meng, Hui Zhang.

Registration of point clouds is a fundamental problem in shape acquisition and shape modeling. In this paper, a novel technique, the sample-sphere method, is proposed to register a pair of point clouds in arbitrary initial positions. This method roughly aligns point clouds by matching pairs of triplets of points, which are approximately congruent under rigid transformation. For a given triplet of points, this method can find all its approximately congruent triplets in $O(kn \log n)$ time, where $n$ is the number of points in the point cloud, and $k$ is a constant depending only on a given tolerance to the rotation error. By employing the techniques of wide bases and largest common point set (LCP), our method is resilient to noise and outliers. Another contribution of this paper is proposing an adaptive distance restriction to improve ICP (iterative closest point) algorithm, which is a classical method to refine rough alignments. With this restriction, the improved ICP is able to reject unreasonable corresponding point pairs during each iteration, so it can precisely align the point clouds which have large non-overlapping regions [24].

6.1.16. Torus/torus intersection

Participants: Xiao-Ming Liu, Chang-Yuan Liu, Jun-Hai Yong, Jean-Claude Paul.

This paper presents a new algorithm for torus/torus intersection. The pre-image of the intersection in the parametric space of one torus is represented by an implicit equation. The pre-image is divided into one-valued function curve segments by characteristic points. The topological feature of these characteristic points is analyzed to obtain the structure of the pre-image. Intersection curves satisfying required precision are found by a self-adaptive refinement method. Experiment results are presented to illustrate the stability and efficiency of the method [21].

6.1.17. Polyline approach for approximating Hausdorff distance between planar free-form curves

Participants: Yan-Bing Bai, Jun-Hai Yong, Jean-Claude Paul, Xiao-Ming Liu.
This paper presents a practical polyline approach for approximating Hausdorff distance between planar free-form curves. After the input curves are approximated with polylines using recursively splitting method, the precise Hausdorff distance between polylines is computed as approximation of Hausdorff distance between free-form curves, and the error of approximation is controllable. The computation of Hausdorff distance between polylines is based on an incremental algorithm that computes directed Hausdorff distance from a line segment to a polyline. Furthermore, not every segment on polylines contributes to the final Hausdorff distance. Based on the bound properties of Hausdorff distance and continuity of polylines, two pruning strategies are applied in order to prune useless segments. R-Tree structure is employed as well to accelerate the pruning process. We experimented our algorithm on sets of Bezier, B-Spline and NURBS curves respectively, and there are 95% segments pruned on approximating polylines in average. Two comparisons are also presented: One is with an algorithm computing directed Hausdorff distance on polylines by building Voronoi diagram of segments, the other is with equation solving method for computing Hausdorff distance between free-from curves [1].

6.1.18. Algorithm for orthogonal projection of parametric curves onto B-spline surfaces

Participants: Hai-Chuan Song, Xiao-Ming Liu, Jean-Claude Paul.

This paper proposes an algorithm for calculating the orthogonal projection of parametric curves onto B-spline surfaces. It consists of a second order tracing method with which we construct a polyline to approximate the pre-image curve of the orthogonal projection curve in the parametric domain of the base surface. The final 3D approximate curve is obtained by mapping the approximate polyline onto the base surface. The Hausdorff distance between the exact orthogonal projection curve and the approximate curve is controlled under the user-specified distance tolerance. And the continuity of the approximate curve is $\varepsilon_T G^1$, where $\varepsilon_T$ is the user-specified angle tolerance. Experiments demonstrate that our algorithm is faster than the existing first order algorithms [8].

6.1.19. Converting sectional views to three orthographic views to reconstruct 3D models

Participants: Fengqing Ding, Hui Zhang, Yamei Wen.

Compared with the CSG-based approach, the Brep-based approach has several advantages to construct 3D models from 2D engineering drawings, such as the structure is simpler and the domain of objects that can be handled is wider. However, this approach cannot handle sectional views directly. In this paper, a new method of converting sectional views to three orthographic views is presented. Firstly, the views which have the same projection direction are merged into one view. If the number of views is two, then a new view will be added according to the coordinate relations. Secondly, elements which have been omitted in sectional views are recovered according to the matching information of the existing edges. Finally, the existing Brep-based approach is used to reconstruct the 3D models. The algorithm can handle full sections, broken-out sections, offset sections as well as two orthographic views. The algorithm has been validated by experiments [15].

6.1.20. Computing the Minimum Distance between Two Tori

Participants: Xiao-Ming Liu, Chang-Yuan Liu, Qiang Hu, Jun-Hai Yong.

The minimal distance computing between two tori is the basis of their collision detection and intersection. A method is proposed for discriminating three types of position relationship (i.e., inclusion, disjunction and intersection) between two tori, and for computing their minimal distance. This paper proves that the Hausdorff distance between two circles in three-dimensional space can be obtained by computing their collinear normal points, which can be calculated by solving an equation of degree 8. With classification and comparison of the collinear normal points, the minimum distance and the Hausdorff distance between these two circles are obtained. In addition, this paper proves that the position relationship between two tori relates to not only the minimum distance but also the directed Hausdorff distance between their central circles. And then the minimum distance between two tori is calculated. Numerical results are presented to illustrate the stability and efficiency of the method [20].
6.1.21. Computing the Inner Distances of Volumetric Models for Articulated Shape Description with a Visibility Graph  

Participants: Yu-Shen Liu, Karthik Ramani, Min Liu.

A new visibility graph based algorithm is presented for computing the inner distances of a 3D shape represented by a volumetric model. The inner distance is defined as the length of the shortest path between landmark points within the shape. The inner distance is robust to articulation and can reflect well the deformation of a shape structure without an explicit decomposition. Our method is based on the visibility graph approach. To check the visibility between pairwise points, we propose a novel, fast, and robust visibility checking algorithm based on a clustering technique which operates directly on the volumetric model without any surface reconstruction procedure, where an octree is used for accelerating the computation. The inner distance can be used as a replacement for other distance measures to build a more accurate description for complex shapes, especially for those with articulated parts [22].

6.1.22. An extended schema and its production rule-based algorithms for assembly data exchange using IGES  

Participants: Kai-Mo Hu, Bin Wang, Yong Liu, Jing Huang, Jun-Hai Yong.

Assembly data exchange and reuse play an important role in CAD and CAM in shortening the product development cycle. However, current CAD systems cannot transfer mating conditions via neutral file format, and their exported IGES files are heterogeneous. In this paper, a schema for the full data exchange of assemblies is presented based on IGES. We first design algorithms for the pre-and-post processors of parts based on solid model, in which the topologies are explicitly specified and will be referred by mating conditions, and then extend the IGES schema by introducing the Associativity Definition Entity and Associativity Instance Entity defined in IGES standard, so as to represent mating conditions. Finally, a production rule-based method is proposed to analyze and design the data exchange algorithms for assemblies. Within this schema, the heterogeneous representations of assemblies exported from different CAD systems can be processed appropriately, and the mating conditions can be properly exchanged. Experiments on the prototype system verify the robustness, correctness, and flexibility of our schema [49].

6.1.23. Robust shape normalization of 3D articulated, volumetric models  

Participants: Chao Wang, Yu-Shen Liu, Min Liu, Jun-Hai Yong, Jean-Claude Paul.

3D shape normalization is a common task in various computer graphics and pattern recognition applications. It aims to normalize different objects into a canonical coordinate frame with respect to rigid transformations containing translation, rotation and scaling in order to guarantee a unique representation. However, the conventional normalization approaches do not perform well when dealing with 3D articulated objects. To address this issue, we introduce a new method for normalizing a 3D articulated object in the volumetric form. We use techniques from robust statistics to guide the classical normalization computation. The key idea is to estimate the initial normalization by using implicit shape representation, which produces a novel articulation insensitive weight function to reduce the influence of articulated deformation. We also propose and prove the articulation insensitivity of implicit shape representation. The final solution is found by means of iteratively reweighted least squares. Our method is robust to articulated deformation without any explicit shape decomposition. The experimental results and some applications are presented for demonstrating the effectiveness of our method [27].


Participants: Zhi Chen, Bin Wang, Norbert Muller, Hui Zhang, Jun-Hai Yong.
The summation of floating-point numbers is ubiquitous in computer systems, while computation implemented in fixed length floating-point arithmetic may lead to inaccurate result due to rounding error. This paper presents an efficient algorithm which produces a faithful result by combining splitting the mantissa and error-free accumulation. Each summand is split into several parts with limited significant bits, which ensures these parts can be accumulated without rounding error under certain conditions. In the implementation, we discuss how to get exponent of floating-point number quickly, which is key to decide how to split summand. Our method works on computers complying with IEEE 754 standard. The running time of our algorithm is proportional to the size of the input data, according to both analysis and numerical tests [12].

6.1.25. 3DMolNavi: A navigation system for flexible molecular shape retrieval based on histogram and dimensionality reduction
Participants: Yu-Shen Liu, Meng Wang, Jean-Claude Paul.

3DMolNavi is a web-based visualized navigation system developed for intuitively exploring flexible molecular shape retrieval. This system is based on the histogram of Inner Distance Shape Signature (IDSS) for fast retrieving molecules that are similar to a query molecule, and uses dimensionality reduction to navigate the retrieved results in 2D and 3D spaces [23].

Participants: Hui Kong, Hehua Zhang, Xiaoyu Song, Ming Gu, Jiaguang Sun.

Geometric algorithms are widely used in many scientific fields like computer vision, computer graphics. To guarantee the correctness of these algorithms, it’s important to apply formal method to them. In this paper, we propose an approach to proving the correctness of geometric algorithms. The main contribution of the paper is that a set of proof decomposition rules is proposed which can help improve the automation of the proof of geometric algorithms. We choose TLA+2, a structural specification and proof language, as our experiment environment. The case study on a classical convex hull algorithm shows the usability of the method [36].

6.1.27. Multi-resolution mesh fitting by B-spline surfaces for reverse engineering
Participants: Sen Zhang, Zhigang Li, Hui Zhang.

This paper presents a new multi-resolution mesh fitting algorithm, extending the adaptive patch-based fitting scheme where each underlying quadrilateral is recursively subdivided into four sub-patches. In this paper, the $G^1$ continuity constraints, which mainly consist of perpendicular constraints and twist compatibility constraints, are deduced for B-spline patches. In order to construct a unique B-spline patch for each quadrilateral, the mesh vertices are applied in a least-square approximation, and the energy functions associated with a patch are minimized. In contrast to the original algorithm, this paper fits the mesh into B-spline patches instead of Bézier patches with $G^1$ continuity. The B-spline patches make the algorithm have more free control points to be used for optimizing the shape of the quadrilateral patches to achieve higher flexible patch control and less recursive times [42].

6.1.28. An example-driven symbol recognition approach based on key features in engineering drawings
Participants: Tiantian Guo, Hui Zhang, Yamei Wen.

In this paper, we present an example-driven symbol recognition algorithm based on its key features in CAD engineering drawings. When user provides an example of a specific symbol, the input symbol is analyzed and its features are extracted automatically. Based on the relation representation, the constrained tree with key feature priority can be established for this type of symbol. By this means, the symbol library can be built and expanded automatically in order to handle variety engineering drawings. In the next stage of the recognition processes, we first locate the key feature nodes in drawings, and then find other elements around which satisfy the topology structure of constrained tree. If all the elements and constraints in the tree are found, the symbol object will be recognized. Because of the accurate position, unnecessary matching calculations are greatly reduced. Experimental results validate that our approach is effective [33].
6.1.29. Automatic generation of canonical views for CAD models

Participants: Kaimo Hu, Bin Wang, Bin Yuan, Junhai Yong.

Selecting the best views for 3D objects is useful for many applications. However, with the existing methods applied in CAD models, the results neither exhibit the 3D structures of the models fairly nor conform to human’s browsing habits. In this paper, we present a robust method to generate the canonical views of CAD models, and the above problem is solved by considering the geometry and visual salient features simultaneously. We first demonstrate that for a CAD model, the three coordinate axes can be approximately determined by the scaled normals of its faces, such that the pose can be robustly normalized. A graph-based algorithm is also designed to accelerate the searching process. Then, a convex hull based method is applied to infer the upright orientation. Finally, four isometric views are selected as candidates, and the one whose depth image owns the most visual features is selected. Experiments on the Engineering Shape Benchmark (ESB) show that the views generated by our method are pleasant, informative and representative. We also apply our method in the calculation of model rectilinearity, and the results demonstrate its high performance [34].

6.2. Geometric Uncertainties

6.2.1. $G^n$ blending multiple surfaces in polar coordinates


This paper proposes a method of $G^n$ blending multiple parametric surfaces in polar coordinates. It models the geometric continuity conditions of parametric surfaces in polar coordinates and presents a mechanism of converting a Cartesian parametric surface into its polar coordinate form. The basic idea is first to re-parameterize the parametric blendees into the form of polar coordinates. Then they are blended simultaneously by a basis function in the complex domain. To extend its compatibility, we also propose a method of converting polar coordinate blending surface into NURBS patches. One application of this technique is to fill N-sided holes. Examples are presented to show its feasibility and practicability [6].

6.2.2. Filling n-sided regions with $G^1$ triangular Coons B-spline patches


Filling n-sided regions is an essential operation in shape and surface modeling. Positional and tangential continuities are highly required in designing and manufacturing. We propose a method for filling n-sided regions with untrimmed triangular Coons B-spline patches, preserving $G^1$ continuity exactly. The algorithm first computes a central point, a central normal, the central, and the corner derivative vectors. Then the region is split into n triangular areas by connecting the central point to each corner of the boundary. These inner curves and all cross-boundary derivatives are computed fulfilling $G^1$ compatibility conditions. And finally, the triangular patches are generated in the Coons B-spline form, one boundary of which is regressed to the central vertex. Neither positional nor tangential error is introduced by this method. And only one-degree elevation is needed [54].

6.2.3. $G^n$ Filling orbicular N-sided holes using periodic B-spline surfaces


The orbicular N-sided hole filling problem is usually introduced by filleting an end-point of a part with large radius. The existing methods based on quadrilateral partition or constrained-optimization can rarely generate high-order continuous blending surfaces under these circumstances. This paper first reparameterizes the boundary of the specified orbicular N-sided hole to ensure the compatibility of neighboring cross-boundary derivatives on the connecting points, preserving their $G^n$ continuity. Then we compute the control points of the periodic B-spline surface using the sufficient $G^n$ continuity condition on the pole and the algorithm of extending parametric surfaces. This method generates single blending surface, which can be converted into standard B spline surface by adding knots without introducing errors. It only elevates the degree of the boundary by n. The construction method is simple and efficient, without iteration nor large-scale matrix solving. It achieves $G^n$ continuity under compatible conditions. The blending examples underline its feasibility and practicability [55].
6.2.4. A Thin-plate CAD Mesh Model Splitting Approach Based on Fitting Primitives  
**Participants:** Chun Geng, Hiromasa Suzuki, Dong-Ming Yan, Takasi Michikawa, Yuichi Sato, Masayoshi Hashima, Eiji Ohta.

Extracting structural information from mesh models is crucial for Simulation Driven Design (SDD) in industrial applications. Focusing on thinplate CAD mesh models (the most commonly used parts in electronic products such as PCs, mobile phones and so on), we present an algorithm based on primitive fitting for segmenting thinplate CAD mesh models into parts of three different types, two of which are extruding surfaces and the other is a lateral surface. This method can be used for solid model reconstruction in the SDD process. Our approach involves two steps. First, a completely automatic method for accurate primitive fitting on CAD meshes is proposed based on the hierarchical primitive fitting framework. In the second step, a novel procedure is proposed for splitting thinplate CAD mesh models by detecting parallel extruding surfaces and lateral surfaces. The method presented here has been proved to work smoothly in applications of real product design [46].

6.2.5. A face-based shape matching method for IGES surface model  
**Participants:** Kaimo Hu, Bin Wang, Yi Gao, Qiming Yuan, Junhai Yong.

IGES is a widely used standard for mechanical data exchange. In this paper, we present a new method for the retrieval task of IGES surface model. Based on this method, a novel distinctive face selection strategy is proposed and evaluated. In the training database, each model is treated as a set of disordered faces, and their features are extracted and stored respectively. The Discounted Cumulative Gain (DCG) value of each face is then calculated and stored for later utilization. To retrieve models in the testing database, we first forecast each face’s DCG value by searching its most similar face’s DCG value in training database, and then the top k faces with highest forecasted DCGs are selected as query input. A greedy algorithm is finally applied to get the total similarity. Experimental results show that our algorithm is superior or at least comparable to some of the most powerful methods in finding parts with similar functionality in most cases [48].

6.2.6. Generating B-spline curves based on control-point interpolation  
**Participants:** Jing Liu, Kan-Le Shi, Jun-Hai Yong, He-Jin Gu.

Generating smooth B-spline curves is a fundamental operation of computer aided geometric design. This paper presents a method to calculate unknown control points using specified control points and knots to generate a smooth B-spline curve. It is based on the basis-function-maximum-value parameterization introduced in this paper. This method first parameterizes all control points; then regards given control points as data points to create a fit curve by interpolation; and finally obtains the unknown control points by evaluating the corresponding parameters directly, which ensures the continuity and smoothness of the generated B-spline curve. The examples in the last section illustrate the feasibility of this method [51].

6.2.7. A new functionality-based benchmark for basic CAD Model retrieval  
**Participants:** Kaimo Hu, Bin Wang, Yi Gao, Dong Li, Junhai Yong.

In this paper, we propose a new functionality-based benchmark for CAD model retrieval. Our benchmark contains 1968 frequently-used CAD models which are divided into training set and test set. The models are carefully classified by their functionalities in industry. Eight different shape descriptors are then compared using four famous evaluation measurements. The results show that models having the same functionalities do not necessarily share the same or similar shapes, hence the functionality-based retrieval methods are encouraged, which we believe will be of great help for the improvements of design reusability. Some possible future work for 3D model retrieval in mechanical domain are also proposed based on the observation of our experiments [47].

6.2.8. Shape similarity assessment approach for CAD models based on graph edit distance  
**Participants:** Bin Wang, Dong Li, Kaimo Hu, Hui Zhang.
This paper proposes a new shape similarity assessment approach for CAD models in Boundary Representation (Brep) based on graph edit distance. A suboptimal computational procedure is performed to find the best alignment between local structures sets of attributed graphs derived from models. Assuming that only a minority of local structures characterize the functionality, we figure out the weight of every local structure in the query model through a training phase, and then evaluate the similarity between two models by calculating the weighted graph edit distance of corresponding attributed graphs. Experiment results show that our method provides solid retrieval performance on a real-world CAD model database [57].

6.2.9. The transition between sharp and rounded features and the manipulation of incompatible boundary in filling n-sided holes

Participants: Kan-Le Shi, Jun-Hai Yong, Peng Liu, Jia-Guang Sun, Jean-Claude Paul.

N-sided hole filling plays an important role in vertex blending. Piegl and Tiller presented an algorithm to interpolate the given boundary and cross-boundary derivatives in B-spline form. To deal with the incompatible cases that their algorithm cannot handle, we propose an extension method to manipulate the transition between sharp and rounded features. The algorithm first patches n crescent-shaped extended surfaces to the boundary with $G^2$ continuity to handle incompatibility problem in the corners. Then, we compute the inner curves and the corresponding cross-boundary derivatives fulfilling tangent and twist compatibilities. The generated B-spline Coons patches are $G^1$-continuously connected exactly, and have $\varepsilon$-$G^1$ continuity with the extended surfaces. Our method improves the continuity-quality of the shape and reduces the count of the inserted knots. It can be applied to all $G^0$-continuous boundary conditions without any restrictions imposed on the boundary or cross-boundary derivatives. It generates better shapes than some popular industrial modeling systems on these incompatible occasions. Some examples underline its feasibility [53].

6.2.10. Epsilon-$G^2$ B-spline surface interpolation

Participants: Kan-Le Shi, Jun-Hai Yong, Jia-Guang Sun, Jean-Claude Paul.

This paper proposes a method to construct a B-spline surface that interpolates the specified four groups of boundary derivative curves in the B-spline form. The discontinuity can be bounded by an arbitrary geometric invariant as the tolerance. The method first handles the six types of the compatibility problems by continuity-preserving reparameterization, knot-insertion and local control-point tuning. The transformed boundary conditions are then parametrically compatible, so the Coons strategy can be applied to construct the final interpolant. Not only can it be used in the reliable geometric modeling, but the approach also can be applied to many other algorithms that require compatibility guarantee [26].

6.2.11. $G^2$ B-spline interpolation to a closed mesh


This paper focuses on interpolating vertices and normal vectors of a closed quad-dominant mesh $1G^2$-continuously using regular Coons B-spline surfaces, which are popular in industrial CAD/CAM systems. We first decompose all non-quadrangular facets into quadrilaterals. The tangential and second-order derivative vectors are then estimated on each vertex of the quads. A least-square adjustment algorithm based on the homogeneous form of $G^2$ continuity condition is applied to achieve curvature continuity. Afterwards, the boundary curves, the first- and the second-order cross-boundary derivative curves are constructed fulfilling $G^2$ continuity and compatibility conditions. Coons B-spline patches are finally generated using these curves as boundary conditions. In this paper, the upper bound of the rank of $G^2$ continuity condition matrices is also strictly proved to be $2n-3$, and the method of tangent-vector estimation is improved to avoid petal-shaped patches in interpolating solids of revolution [7].

6.2.12. A new method for identifying and validating features from 2D sectional views

Participants: Yamei Wen, Hui Zhang, Jiaguang Sun, Jean-Claude Paul.
Feature identification is one of the key steps for 3D solids reconstruction from 2D vector engineering drawings using the volume-based method. In this paper, we propose a novel method to identify and validate features from sectional views. First, features are classified as explicit features (EPFs) and implicit features (IPFs), which are then identified in an order of priority using heuristic hints. We show that the problem of constructing EPFs can be formulated as a 0-1 integer linear program (ILP), and the IPFs are generated based on the understanding of semantic information of omitted projections in sectional views. Then, the Loop-Relation Graph (LRG) is introduced as a multi-connected-subgraph representation for describing the relations between loops and features. According to the LRG, a reasoning technique based on confidence is implemented to interactively validate features. This method can recover features without complete projections, and the level of understanding sectional views is improved. Full sections, partial sections, offset sections as well as revolved sections can be handled by our method. Several examples are provided to demonstrate the practicability of our approach.

6.2.13. Ridge extraction of a smooth 2-manifold surface based on vector field
Participants: Wujun Che, Xiaopeng Zhang, Yi-Kuan Zhang, Jean-Claude Paul.

This paper presents a general scheme to compute ridges on a smooth 2-manifold surface from the standpoint of a vector field. A ridge field is introduced. Starting with an initial ridge, which may or may not be umbilical, a ridge line is then traced by calculating an associated integral curve of this field in conjunction with a new projection procedure to prevent it from diverging. This projection is the first that can optimize a ridge guess to lie on a ridge line uniquely and accurately. In order to follow this scheme, we not only develop practical ridge formulae but also address their corresponding computational procedures for an analytical surface patch, especially for an implicit surface. In contrast to other existing methods, our new approach is mathematically sound and characterized by considering the full geometric structures and topological patterns of ridges on a generic smooth surface. The resulting ridges are accurate in the numerical sense and meet the requirement of high accuracy with complete topology. Although the objective of this paper is to develop a mathematically sound framework for ridges on a smooth surface, we give a comprehensive review of relevant works on both meshes and smooth surfaces for readers.

6.2.14. Manifold-ranking based retrieval using $k$-regular nearest neighbor graph
Participants: Bin Wang, Feng Pan, Kaimo Hu, Jean-Claude Paul.

Manifold-ranking is a powerful method in semi-supervised learning, and its performance heavily depends on the quality of the constructed graph. In this paper, we propose a novel graph structure named $k$-regular nearest neighbor ($k$-RNN) graph as well as its constructing algorithm, and apply the new graph structure in the framework of manifold-ranking based retrieval. We show that the manifold-ranking algorithm based on our proposed graph structure performs better than that of the existing graph structures such as $k$-nearest neighbor ($k$-NN) graph and connected graph in image retrieval, 2D data clustering as well as 3D model retrieval. In addition, the automatic sample reweighting and graph updating algorithms are presented for the relevance feedback of our algorithm. Experiments demonstrate that the proposed algorithm outperforms the state-of-the-art algorithms.

6.2.15. Efficient computation of clipped Voronoi diagram for mesh generation
Participants: Dong-Ming Yan, Wenping Wang, Bruno Lévy, Yang Liu.

The Voronoi diagram is a fundamental geometric structure widely used in various fields, especially in computer graphics and geometry computing. For a set of points in a compact domain (i.e. a bounded and closed 2D region or 3D volume), some Voronoi cells of their Voronoi diagram are infinite or partially outside of the domain, but in practice only the parts of the cells inside the domain are needed, as when computing the centroidal Voronoi tessellation. Such a Voronoi diagram confined to a compact domain is called a clipped Voronoi diagram. We present an efficient algorithm for computing the clipped Voronoi diagram for a set of sites with respect to a compact 2D region or a 3D volume. We also apply the proposed method to optimal mesh generation based on the centroidal Voronoi tessellation.
6.2.16. Automatic Generation of Canonical Views for CAD Models

Participants: Kaimo Hu, Bin Wang, Bin Yuan, Junhai Yong.

Selecting the best views for 3D objects is useful for many applications. However, with the existing methods applied in CAD models, the results neither exhibit the 3D structures of the models fairly nor conform to human’s browsing habits. In this paper, we present a robust method to generate the canonical views of CAD models, and the above problem is solved by considering the geometry and visual salient features simultaneously. We first demonstrate that for a CAD model, the three coordinate axes can be approximately determined by the scaled normals of its faces, such that the pose can be robustly normalized. A graph-based algorithm is also designed to accelerate the searching process. Then, a convex hull based method is applied to infer the upright orientation. Finally, four isometric views are selected as candidates, and the one whose depth image owns the most visual features is selected. Experiments on the Engineering Shape Benchmark (ESB) show that the views generated by our method are pleasant, informative and representative. We also apply our method in the calculation of model rectilinearity, and the results demonstrate its high performance \[ 35 \].

6.2.17. Mechanical Parts Retrieval Based on Typical Face Matching

Participants: Yi Gao, Bin Wang, Kaimo Hu, Junhai Yong.

This paper presents a face-based retrieval algorithm to search mechanical parts with similar partial features. The method makes it easier to retrieve models with partial features so as to support early stage reusability. In the training phase, all the faces in the database are trained and assigned with a value indicating their distinction. Trivial faces and atypical ones are removed in this phase to improve online retrieval efficiency. In the query phase, we evaluate the distinction of the input faces by aligning them with faces in the database. A greedy algorithm is finally applied to match the input faces and the faces in the database according to their similarity order. Experimental results show that our method can provide a favorable performance when applied to retrieve the models with common partial features comparing to some other mesh-based methods \[ 18 \].

6.2.18. Continuity Transition with a Single Regular Curved-Knot B-Spline Surface

Participants: Kan-Le Shi, Jun-Hai Yong, Jean-Claude Paul, Jia-Guang Sun.

We propose a canonical form of the curved-knot B-spline surface called the regular curved-knot B-spline. On one hand it allows the transition of the knot vectors so that the continuity configurations of the two opposite boundaries can be different. On the other hand, the regular form achieves the simplicity in storage, evaluation and the construction algorithms, and that makes it possible to be applied in the industrial geometric modeling systems. The applications: bridging, multi-sided hole filling and irregular feature modeling, show that it is well suited for modeling complicated objects, such as a transition between sharp and rounded features. Compared with patching numbers of B-splines, it not only increases the inter-surface continuity of the shape, but also reduces the complexity of algorithms \[ 25 \].

6.2.19. Meshless quadrangulation by global parameterization

Participants: Er Li, Bruno Lévy, Xiaopeng Zhang, Wu-Jun Che, Weiming Dong, Jean-Claude Paul.

Point cloud is a basic description of discrete shape information. Parameterization of unorganized points is important for shape analysis and shape reconstruction of natural objects. In this paper we present a new algorithm for global parameterization of an unorganized point cloud and its application to the meshing of the cloud. Our method is guided by principal directions so as to preserve the intrinsic geometric properties. After initial estimation of principal directions, we develop a kNN(k-nearest neighbor) graph-based method to get a smooth direction field. Then the point cloud is cut to be topologically equivalent to a disk. The global parameterization is computed and its gradients align well with the guided direction field. A mixed integer solver is used to guarantee a seamless parameterization across the cut lines. The resultant parameterization can be used to triangulate and quadrangulate the point cloud simultaneously in a fully automatic manner, where the shape of the data is of any genus \[ 19 \].
6.3. Computer Graphics

6.3.1. Multi-Image Based Photon Tracing for Interactive Global Illumination of Dynamic Scenes  
Participants: Chunhui Yao, Bin Wang, Bin Chan, Junhai Yong, Jean-Claude Paul.  

Image space photon mapping has the advantage of simple implementation on GPU without pre-computation of complex acceleration structures. However, existing approaches use only a single image for tracing caustic photons, so they are limited to computing only a part of the global illumination effects for very simple scenes. In this paper we fully extend the image space approach by using multiple environment maps for photon mapping computation to achieve interactive global illumination of dynamic complex scenes. The two key problems due to the introduction of multiple images are 1) selecting the images to ensure adequate scene coverage; and 2) reliably computing raygeometry intersections with multiple images. We present effective solutions to these problems and show that, with multiple environment maps, the image-space photon mapping approach can achieve interactive global illumination of dynamic complex scenes. The advantages of the method are demonstrated by comparison with other existing interactive global illumination methods [10].

6.3.2. Quality Solid Texture Synthesis using Position and Index Histogram Matching  
Participants: Jiating Chen, Bin Wang.  

The synthesis quality is one of the most important aspects in solid texture synthesis algorithms. In recent years several methods are proposed to generate high quality solid textures. However, these existing methods often suffer from the synthesis artifacts such as blurring, missing texture structures, introducing aberrant voxel colors, and so on. In this paper, we introduce a novel algorithm for synthesizing high quality solid textures from 2D exemplars. We first analyze the relevant factors for further improvements of the synthesis quality, and then adopt an optimization framework with the k-coherence search and the discrete solver for solid texture synthesis. The texture optimization approach is integrated with two new kinds of histogram matching methods, position and index histogram matching, which effectively cause the global statistics of the synthesized solid textures to match those of the exemplars. Experimental results show that our algorithm outperforms or at least is comparable to the previous solid texture synthesis algorithms in terms of the synthesis quality [44].

6.3.3. Real-time rendering of heterogeneous translucent objects with arbitrary shapes  
Participants: Yajun Wang, Jiaping Wang, Nicolas Holzschuch, Kartic Subr, Jun-Hai Yong, Baining Guo.  

We present a real-time algorithm for rendering translucent objects of arbitrary shapes. We approximate the scattering of light inside the objects using the diffusion equation, which we solve on-the-fly using the GPU. Our algorithm is general enough to handle arbitrary geometry, heterogeneous materials, deformable objects and modifications of lighting, all in real-time. In a pre-processing step, we discretize the object into a regular 4-connected structure (QuadGraph). Due to its regular connectivity, this structure is easily packed into a texture and stored on the GPU. At runtime, we use the QuadGraph stored on the GPU to solve the diffusion equation, in real-time, taking into account the varying input conditions: Incoming light, object material and geometry. We handle deformable objects, provided the deformation does not change the topological structure of the objects [58].

6.3.4. Fast Local Color Transfer via Dominant Colors Mapping  
Participants: Weiming Dong, Guanbo Bao, Xiaopeng Zhang, Jean-Claude Paul.  

We present a novel algorithm to address the above issues. Our method establishes a tight connection between the local color statistics of the source and target images. All the obvious color features can be presented in the result [45].

6.3.5. Real-time watercolor illustrations and animation on GPU  
This paper presents a real-time approach to render 3D scenes with the effects of watercolor on GPU. Most processes of the approach are implemented with image-space techniques. Our algorithm renders detail layer, ambient layer and stroke layer separately, and then combines them into final result. During the rendering processes, we use screen space ambient occlusion and shadow mapping to compute shadow in much shorter time, and we use image filter approach to simulate important effects of watercolor. Because our approach is mainly implemented with image-space techniques, it is convenient to use GPU to accelerate the rendering processes and finally our approach achieves real-time speed \[ 59 \].

6.3.6. A Hierarchical Grid Based Framework for Fast Collision Detection

We present a novel hierarchical grid based method for fast collision detection (CD) for deformable models on GPU architecture. A two-level grid is employed to accommodate the non-uniform distribution of practical scene geometry. A bottom-to-top method is implemented to assign the triangles into the hierarchical grid without any iteration while a deferred scheme is introduced to efficiently update the data structure. To address the issue of load balancing, which greatly influences the performance in SIMD parallelism, a propagation scheme which utilizes a parallel scan and a segmented scan is presented, distributing workloads evenly across all concurrent threads. The proposed method supports both discrete collision detection (DCD) and continuous collision detection (CCD) with self-collision. Some typical benchmarks are tested to verify the effectiveness of our method. The results highlight our speedups over prior algorithms on different commodity GPUs \[ 17 \].

6.3.7. Improved Stochastic Progressive Photon Mapping with Metropolis Sampling
Participants: Jiating Chen, Bin Wang, Junhai Yong.

This paper presents an improvement to the stochastic progressive photon mapping (SPPM), a method for robustly simulating complex global illumination with distributed ray tracing effects. Normally, similar to photon mapping and other particle tracing algorithms, SPPM would become inefficient when the photons are poorly distributed. An inordinate amount of photons are required to reduce the error caused by noise and bias to acceptable levels. In order to optimize the distribution of photons, we propose an extension of SPPM with a Metropolis-Hastings algorithm, effectively exploiting local coherence among the light paths that contribute to the rendered image. A well-designed scalar contribution function is introduced as our Metropolis sampling strategy, targeting at specific parts of image areas with large error to improve the efficiency of the radiance estimator. Experimental results demonstrate that the new Metropolis sampling based approach maintains the robustness of the standard SPPM method, while significantly improving the rendering efficiency for a wide range of scenes with complex lighting \[ 14 \].

6.3.8. Efficient Depth-of-Field Rendering with Adaptive Sampling and Multiscale Reconstruction

Depth-of-field is one of the most crucial rendering effects for synthesizing photorealistic images. Unfortunately, this effect is also extremely costly. It can take hundreds to thousands of samples to achieve noise-free results using Monte Carlo integration. This paper introduces an efficient adaptive depth-of-field rendering algorithm that achieves noise-free results using significantly fewer samples. Our algorithm consists of two main phases: adaptive sampling and image reconstruction. In the adaptive sampling phase, the adaptive sample density is determined by a ‘blur-size’ map and ‘pixel-variance’ map computed in the initialization. In the image reconstruction phase, based on the blur-size map, we use a novel multiscale reconstruction filter to dramatically reduce the noise in the defocused areas where the sampled radiance has high variance. Because of the efficiency of this new filter, only a few samples are required. With the combination of the adaptive sampler and the multiscale filter, our algorithm renders near-reference quality depth-of-field images with significantly fewer samples than previous techniques \[ 13 \].

6.3.9. Real-Time Illumination of Complex Lights Using Divided Sampling
Participants: Chunhui Yao, Bin Wang, Junhai Yong.
Existing methods for real-time illumination of complex lights, either require long time pre-computation, or only focus on some special types of illumination. Because computing different kinds of illumination requires different sampling strategies, this paper introduces a novel efficient framework for rendering illumination of complex light sources, in which the pre-computation is not necessary. We divide the rendering equation into three parts: high-frequency term, low-frequency term and occlusion term. Each term is computed by a proper sampling strategy. High-frequency term is solved by importance sampling the BRDF, while low-frequency term is computed by importance sampling the light sources. Occlusion term is computed with depth information in screen-space, and the required number of samples is greatly reduced by interleaved sampling. Our framework is easy to implement on GPU and can solve many real-time rendering problems. We take real-time environment-map-lighting as an example for demonstrating applications of this framework. The results show that our technique can handle complete light effects with higher quality than previous works [31].

6.3.10. Fast Multi-Operator Image Resizing and Evaluation
Participants: Weiming Dong, Guanbo Bao, Xiaopeng Zhang, Jean-Claude Paul.

Current multi-operator image resizing methods succeed in generating impressive results by using image similarity measure to guide the resizing process. An optimal operation path is found in the resizing space. However, their slow resizing speed caused by inefficient computation strategy of the bidirectional patch matching becomes a drawback for practical use. In this paper, we present a novel method to address this problem. By combining seam carving with scaling and cropping, our method can realize content-aware image resizing very fast. We define cost functions combining image energy and dominant color descriptor for all the operators to evaluate the damage to both local image content and global visual effect. Therefore our algorithm can automatically find an optimal sequence of operations to resize the image by dynamic programming or greedy algorithm. We also extend our algorithm to indirect image resizing which can protect the aspect ratio of the dominant object in an image [16].

6.3.11. Real-time Volume Caustics with Image-based Photon Tracing
Participants: Yuxiang Wang, Bin Wang, Li Chen.

Rendering of volume caustics in participating media is often expensive, even with different acceleration approaches. Basic volume photon tracing is used to render such effects, but rather slow due to its massive quantity of photons to be traced. In this paper we present an image-based volume photon tracing method for rendering volume caustics at real-time frame rates. Motivated by multi-image based photon tracing, our technique uses multiple depth maps to accelerate the intersection test procedure, achieving a plausible and fast rendering of volume caustics. Each photon dynamically selects the depth map layer for intersection test, and the test converges to an approximate solution using image space methods in a few recursions. This allows us to compute photon distribution in participating media while avoiding massive computation on accurate intersection tests with scene geometry. We demonstrate that our technique, combined with photon splatting techniques, is able to render volume caustics caused by multiple refractions [39].

6.3.12. Parallel Spacial Hashing for Collision Detection of Deformable Surfaces

We present a fast collision detection method for deformable surfaces with parallel spatial hashing on GPU architecture. The efficient update and access of the uniform grid are exploited to accelerate the performance in our method. To deal with the inflexible memory system, which makes the building of stream data a challenging task on GPU, we propose to subdivide the whole workload into irregular segments and design an efficient evaluation algorithm, which employs parallel scan and stream compaction, to build the stream data in parallel. The load balancing is a key aspect that needs to be considered in the SIMD parallelism. We break the heavy and irregular collision computation down into lightweight part and heavyweight part, ensuring the later perfectly run in load balancing manner with each concurrent thread processes just a single collision. In practice, our approach can perform collision detection in tens of milliseconds on a PC with NVIDIA GTX 260 graphics card on benchmarks composed of millions of triangles. The results highlight our speedups over prior CPU-based and GPU-based algorithms [32].
6.3.13. Distribution-Aware Image Color Transfer

Participants: Fuzhang Wu, Weiming Dong, Xing Mei, Xiaopeng Zhang, Xiaohong Jia, Jean-Claude Paul.

Color transfer is a practical image editing technology which is useful in various applications. An ideal color transfer algorithm should keep the scene in the source image and apply the color styles of the reference image. All the dominant color styles of the reference image should be presented in the result especially when there are similar contents in the source and reference images. We propose a robust color transfer framework to address the above issues. Our method can establish a soft connections between the local color statistics of the source and reference images. All the obvious color features can be presented in the result image, as well as the spatial distribution of the reference color pattern [40].

6.3.14. Translucent Material Transfer Based on Single Images

Participants: Chao Li, Weiming Dong, Ning Zhou, Xiaopeng Zhang, Jean-Claude Paul.

Extraction and re-rendering of real materials give large contributions to various image-based applications. As one of the key properties of modeling the appearance of an object, materials mainly focus on the effects caused by light transportation. Therefore, understanding the characteristics of a complex material from a single photograph and transferring it to an object in another image becomes a very challenging problem. In this paper, we present a novel framework to transfer real translucent materials such as fruits and flowers between single images. We define a group of information which can model the attributes during the extraction and transfer process. Once we extract this information from both the source and target images, we can easily produce a realistic photograph of an object with target-like materials and suitable shading effects in the environment of sources [37].

6.3.15. Multicage Image Deformation On GPU

Participants: Weiliang Meng, Xiaopeng Zhang, Weiming Dong, Jean-Claude Paul.

As a linear blending method, cage-based deformation is widely used in various applications of image and geometry processing. In most cases especially in the interactive mode, deformation based on embedded cages does not work well as some of the coefficients are not continual and make the deformation discontinuous, which means existing “spring up” phenomenon. However, it’s common for us to deform the ROI (Region of Interest) while keeping local part untouched or with only small adjustments. In this paper, we design a scheme to solve the above problem. A multicage can be generated manually or automatically, and the image deformation can be adjusted intelligently according to the local cage shape to preserve important details. On the other hand, we don’t need to care about the pixels’ position relative to the multicage. All the pixels go through the same process, and this will save a lot of time. We also design a packing method for cage coordinates to pack all the necessary coefficients into one texture. A vertex shader can be used to accelerate the deformation process, leading to realtime deformation even for large images [38].

6.3.16. Real-time volume caustics with image-based photon tracing

Participants: Yuxiang Wang, Bin Wang, Li Chen.

Rendering of volume caustics in participating media is often expensive, even with different acceleration approaches. Basic volume photon tracing is used to render such effects, but rather slow due to its massive quantity of photons to be traced. In this paper we present an image-based volume photon tracing method for rendering volume caustics at real-time frame rates. Motivated by multi-image based photon tracing, our technique uses multiple depth maps to accelerate the intersection test procedure, achieving a plausible and fast rendering of volume caustics. Each photon dynamically selects the depth map layer for intersection test, and the test converges to an approximate solution using image space methods in a few recursions. This allows us to compute photon distribution in participating media while avoiding massive computation on accurate intersection tests with scene geometry. We demonstrate that our technique, combined with photon splatting techniques, is able to render volume caustics caused by multiple refractions [39].

6.3.17. Interactive Visual Simulation of Dynamic Ink Diffusion Effects

Participants: Shibiao Xu, Xing Mei, Weiming Dong, Zhiyi Zhang, Xiaopeng Zhang.
This paper presents an effective method that simulates the ink diffusion process with visual plausible effects and real-time performance. Our algorithm updates the dynamic ink volume with a hybrid grid-particle representation: the fluid velocity field is calculated with a low-resolution grid structure, while the highly detailed ink effects are controlled and visualized with the particles. We propose an improved ink rendering method with particle sprites and motion blur techniques. The simulation and the rendering processes are efficiently implemented on graphics hardware for interactive frame rates. Compared to traditional simulation methods that treat water and ink as two mixable fluids, our method is simple but effective: it captures various ink effects such as pinned boundary [Chu and Tai 2005] and filament pattern [Shiny et al. 2010] with real-time performance; it allows easy interaction with the artists; it includes basic solid-fluid interaction. We believe that our method is attractive for industrial animation and art design [41].
6. New Results

6.1. Low Mach number flows simulations issue

Participants: Pascal Bruel, Tarik Kousksou.

Since the targeted flows simulations (DNS) are by essence unsteady and at low Mach number, the use of a compressible solve has to be considered with great care. As a preliminary step towards the use of a fully implicit DG approach, we have joined the group of E. Dick at the Ghent University (Belgium) to help studying different aspects linked with low Mach number flow simulations [8], [7], [3]. The question of time consistency of the hyperbolic fluxes schemes for unsteady calculations has been dealt with and some pathological behavior of schemes such as the ones belonging to the AUSM family have been evidenced. A modification of AUSM+-up that satisfies the time-step dependency as well as the suitable scaling property of the pressure-velocity coupling evidenced has been proposed and tested [6].

6.2. Simulations of jets in crossflow

Participants: Pascal Bruel, Tarik Kousksou.

In order to prepare our benchmarking activity, we have been using the LES computer code AVBP (from Cercfas and IFP) in order to simulate a multijet in crossflow configuration corresponding to the MAVERIC flow configuration. These simulations have been done in partnership with Turbomeca. The comparisons between LES and experiments is quite encouraging. An acoustic forcing technique has been used to establish a stationary acoustic wave inside the crossflow. The mass flow rate through the different holes proved to be significantly altered by the presence of the planar acoustic wave [4].

6.3. Discontinuous Galerkin methods for compressible multiphase flows

Participants: Vincent Perrier, Erwin Franquet.

We developed discontinuous Galerkin methods for compressible multiphase flows. This method is based on the method developed in [12] concerning the modelling of multiphase flows, and on the method developed in [36] concerning multiphase flows. In the method developed in [12], the exact expression of the continuous system is unclear, and as a consequence, the cell integral that naturally appear in the discontinuous Galerkin formulation is not clearly defined. In [2], [11], we developed an original analysis of the scheme [12] by mean of a stochastic process, which gave a unified framework for defining both the numerical scheme and the continuous limit. This was then applied to multiphase flows with phase transition in [10], [1]. Our method is currently being extended to interface flows with the maximum preserving limiter developed by [38].
6. New Results

6.1. Mathematical analysis of kinetic models

**Participants:** Aurore Back, Nicolas Besse, Emmanuel Frénod, Mathieu Lutz.

**6.1.1. Asymptotic analysis of gyrokinetic models**

Proceeding [52] presents the method that allows us to get the gyrokinetic Approximation of the Dynamical System satisfied by the trajectory of a particle submitted to a Strong Magnetic Field. The goal of the method is to build a change of coordinates in order to make the fast dynamics of two components of the trajectory to disappear. This change of coordinates is based on a Darboux mathematical Algorithm and on a Lie Transform. It is the first work of a forthcoming series of papers which goal is to make the Geometrical gyrokinetic Approximation a mathematically affordable theory. Review paper [51] presents the results of Two-Scale Convergence Theory and an application to Homogenization of linear Singularly Perturbed Hyperbolic Partial Differential Equations. It consists in the theoretical basis of the Two-Scale Numerical Methods.

**6.1.2. Two-scale convergence with differential form**

In the framework of the thesis of Aurore Back [11], we developed at two-scale convergence theory using the tools of exterior calculus and differential forms. A geometric formulation of the Vlasov-Maxwell equations was introduced and some geometric conservation properties were proved.

**6.1.3. Analysis of multi-water-bag models**

In the case of toroidal geometry, thanks to the strong anisotropy between parallel (to the magnetic field) and transverse direction, we could perform an asymptotic analysis of the eigenvalue problem for the integro-differential gyrowaterbag operator with two independent dimensions, the third being represented by Fourier toroidal modes. This analysis enabled us to reduce the two-dimensional integro-differential operator to a series of one-dimensional integro-differential operators the solution of which enables to obtain as well the poloidal as the radial envelope of the global eigenmodes as well as their local frequency. Note that the terms of the series can be computed numerically independently from each other as the differential variables decouple, which leads to an embarrassingly parallel algorithm. The global dispersion relation for the global eigenfrequency appears as an integral quantification relation involving the local frequency due to the property of conservation of the action. On the other hand the mathematical analysis, in particular the spectral properties, of the obtained operators has been performed. Several mathematical results on the well-posedness of gyrowaterbag models have been obtained as an exact geometric reduction of the Vlasov equation, the notion of solution differing depending on the nature of the problem being considered [14], [15], [16].

6.2. Numerical analysis

**Participants:** Martin Campos Pinto, Nicolas Crouseilles, Michel Mehrenberger, Eric Sonnendrücker.

**6.2.1. Analysis of numerical methods for the Vlasov-Poisson system**

In [47], we derive the order conditions for fourth order time splitting schemes in the case of the 1D Vlasov-Poisson system. Computations to obtain such conditions are motivated by the specific Poisson structure of the Vlasov-Poisson system: this structure is similar to Runge-Kutta-Nyström systems. The obtained conditions are proved to be the same as RKN conditions derived for ODE up to the fourth order. Numerical tests are performed and show the benefit of using high order splitting schemes in that context.
In [37], we prove enhanced error estimates for high order semi-lagrangian discretizations of the Vlasov-Poisson equation. It provides new insights into optimal numerical strategies for the numerical solution of this problem. The new error estimate $O(\min(\frac{\Delta x}{\Delta t}, 1)\Delta x^p + \Delta t^2)$ is based on advanced error estimates for semi-lagrangian schemes, also equal to shifted Strang schemes, for the discretization of the advection equation.

### 6.2.2. Analysis of a new particle method with deformable shapes

Particle methods are known to be simple and efficient in most practical cases, however they suffer from weak convergence properties: they only converge in a strong sense when the particles present an extended overlapping (i.e., when the number of overlapping particles tends to infinity as the mesh size $h$ of their initialization grid tends to 0), and additional constraints such as vanishing moments. In practice, extended particle overlapping can be expensive and it involves an additional parameter to be optimized, such as the overlapping exponent $q < 1$ for which the particles radius behaves like $h^q$. In PIC codes for instance, extended overlapping requires increasing the number of particles per cell together with the number of cells, which determine the radius of the particles. In many practical cases such conditions are not met, which leads to strong oscillations in the solutions. To smooth out the oscillations some methods (like the Denavit redeposition scheme, recently revisited as a Forward semi-Lagrangian scheme) use periodic remappings, but frequent remappings introduce unwanted numerical diffusion which seems to contradict the benefit of using low-diffusion particle schemes. Moreover, the vanishing moment condition prevents high orders to be achieved with positive particles.

In [44] we present a new class of particle methods with deformable shapes for transport problems that converge in the supremum norm without requiring remappings, extended overlapping or vanishing moments for the particles. Indeed, unlike the classical error analysis based on a smoothing kernel argument, our estimates hold for any particle collection with Lipschitz smoothness and compact supports that have the same scale than their initialization grid. Our results are threefold. On the theoretical side we first show that for arbitrarily smooth characteristic flow, high order convergence rates are obtained by deforming the particles with local polynomial mappings. On the practical side we provide an explicit implementation of the first order case: the resulting linearly-transformed particle (LTP) scheme consists of transporting the particle centers along the numerical flow, together with finite difference approximations of the local Jacobian matrices of the flow. For the fully discrete scheme we establish rigorous a priori error estimates and demonstrate the uniform boundedness of the particle overlapping. Finally, we describe an adaptive multilevel version of the LTP scheme that includes a local correction filter for positivity-preserving approximations.

In [45] we apply the LTP method to the 1+1d Vlasov-Poisson problem with a simple deposition scheme and show that deforming the particles helps removing the noise traditionally observed with standard PIC schemes.

### 6.2.3. Two-Scale Asymptotic-Preserving issues

In the submitted paper [48], we build a Two-Scale Macro-Micro decomposition of the Vlasov equation with a strong magnetic field. This consists in writing the solution of this equation as a sum of two oscillating functions with circumscribed oscillations. The first of these functions has a shape which is close to the shape of the Two-Scale limit of the solution and the second one is a correction built to offset this imposed shape. The aim of such a decomposition is to be the starting point for the construction of Two-Scale Asymptotic-Preserving Schemes. The aim of using Two-Scale Asymptotic-Preserving Schemes is first, to deal efficiently with long time scales with solutions having high frequency oscillations and second, to manage the transition between different regimes, in a unified framework.

The aim of a new starting project is to test on a simplified model the Two-Scale Asymptotic-Preserving Schemes. The model, a two dimensional in phase space Vlasov-Poisson equation with small parameter, is used for a long time simulation of a beam in a focusing channel. This work was already done in [71] in the case where the solution is approximated by the two scale limit. The goals are first to improve this approximation, by going further, to the first order one, and secondly, to replace this approximation by an exact decomposition, using the macro-micro framework. This last approach will permit to treat the case of a not necessary small parameter.
In order to accomplish the first task we started to write a PIC code which is to be integrated in SeLaLib.

6.3. Development of numerical methods

**Participants:** Aurore Back, Nicolas Besse, Jean-Philippe Braeunig, Anaïs Crestetto, Emmanuel Frénod, Philippe Helluy, Sever Hirstoaga, Ahmed Ratnani, Eric Sonnendrücker.

6.3.1. Application of isogeometric analysis to plasma physics simulations

Mainly around the PhD thesis of Ahmed Ratnani [13] which has been defended in October 2011, we have been using the concept of isogeometric analysis introduced by Hughes and co-workers [77] which consists in representing the computational domain as well as the numerical solution of the equations with NURBS (Non Uniform Rational B-Splines).

In [26] we introduced a time-domain conforming Finite Element solver using arbitrary order B-Splines as basis functions. The discrete function spaces used in the Finite Element formulation form a De Rham sequence, which has proved to be an important property for numerical Maxwell solvers. In particular, they allow to have a simple relation between spline coefficients of the magnetic and electric field that is independent on order and geometry for one of Ampère’s or Faraday’s law. The other then necessarily involves a discrete Hodge operator which depends on order and geometry. High-order energy conserving leap-frog schemes have been validated with this solver.

In [21], we developed an arbitrary order B-Spline Finite Element solver for the quasi-neutrality equation that is generally coupled to gyrokinetic Vlasov-solvers. Compared to the previous solver used in GYSELA which was spectral in the angular variable and second order Finite Differences in the radial variable this solver can be of high-order in both direction. This enables us for a given accuracy to decrease the number of grid points. Moreover, thanks to the periodicity in the angular variable and the tensor product structure of the problem, we could introduce a fast diagonalization method using a FFT such that the cost of the new solver only marginally depends on the order and is only slightly higher than the cost of the previously used method. Another important new algorithm introduced in this work is the decoupling of the parallel and transverse parts of the equation, by solving successively for the average value along the parallel direction and the remaining part.

In [28] we present an axisymmetric PIC code based on isogeometric analysis, which was initially the IsoPIC project (supported by CEA Gramat) of CEMRACS 2010. The goal of this study is to use it for solving the system of Vlasov-Maxwell equations. The idea is to develop an axisymmetric Finite Element PIC (Particle-In-Cell) code in which specific spline Finite Elements are used to solve the Maxwell equations (in 2D transverse electric mode) and the same spline functions serve as shape function for the particles. The computational domain itself is defined using splines or NURBS. We are in particular interested in the emission of electrons in a diode with hemispherical cathode (thanks to symmetry in θ direction, we can consider the 2D axisymmetric geometry).

6.3.2. Spline discrete differential forms

In [11], [36] we have developed the concept of spline discrete differential forms which can be used to discretized equations defined using the notions of exterior calculus. These have been applied for the numerical solution of the Maxwell and the Vlasov equations. Hodge operators either using a dual grid or a weak formulation are developed and commuting diagram properties are proved.

6.3.3. Drift-kinetic simulations

We implemented the conservative semi-Lagrangian method in the GYSELA code which is based on the classical backward semi-Lagrangian method which is not exactly conservative. We noticed that for the conservative method it is essential that the advection field remains numerically exactly divergence free in order to avoid numerical instabilities. In addition specific limiters for the conservative method were developed and comparison between the backward semi-Lagrangian method, the conservative semi-Lagrangian method with 1D splitting and the same method with an unsplit Finite Volume like formulation in the (r, θ) plane which provides a better conservation of volume [43].
6.3.4. Waterbag simulations

In [31] we apply the multi-water-bag model and the method of moments to the Vlasov-Poisson system in a case where the solution becomes multivalued. The motivation of this study is that the kinetic Vlasov-Poisson model is very expensive to solve numerically. It can be approximated by a multi-water-bag model in order to reduce the complexity. This model amounts to solve a set of Burgers equations, which can be done easily by finite volume methods. However, the physical solution can become multivalued (filamentation appears). In this case, shocks appear in the simulation and we lose information about the filaments. To catch them, we can use a moment method. We describe here the two models and present several numerical experiments.

A linear analysis code CYLGYR based on the gyrowaterbag model in cylindrical geometry has been developed. It enables, starting from a given equilibrium configuration, to obtain the whole set of modes that can exist. It was used to validate the linear phase of the previously developed non linear semi-lagrangian gyrowaterbag code in cylindrical geometry GMWB3D-SL [61]. Excellent agreement for the growth rates (eigenvalues) and the radial envelopes (eigenfunctions) has been obtained for global eigenmodes. On the other hand the linear code CYLGYR gives results in excellent agreement with those given by the linear kinetic code KINEZERO. We are now using as well the linear gyrowaterbag code CYLGYR and the non linear code GMWB3D-SL to compare the quasi-linear and non-linear fluxes and thus measure the validity of the quasi-linear approach for gyrokinetic turbulence. Such comparisons will also be performed with the gyrokinetic code GYSELA in cylindrical geometry.

6.3.5. Validation of the quasi-linear theory

We have developed and optimized a parallel semi-Lagrangian code for the numerical resolution of the Hamiltonian Vlasov-wave model in two phase-space dimensions. Using this code to perform a statistical study on a large number of runs of the system we have showed that the quasi-linear theory, whose aim is to justify the approximation of a self-coherent hamiltonian system like the Vlasov-Poisson model by diffusive self-coherent model of Fokker-Planck type, was valid in the strongly chaotic non linear regime of 1D electrostatic turbulence provided it is regarded from a statistical point of view [17].

6.4. Plasma-wall interactions - application to ELM modes on JET

Participants: Sever Hirstoaga, Giovanni Manfredi.

The aim of [24] was to model the effect of energetic charged particles (generated during violent events known as edge-localized modes) on the divertor plates of a tokamak. We thus have developed a 1D Vlasov-Poisson code with open boundaries for both ions and electrons. This work was already described in last year’s report. In the proceedings [33] we compare the numerical results previously obtained with the Eulerian code (in [24]) to those obtained with two different approches: a PIC code (developed at the University of Innsbruck) and a fluid code (developed at Culham, UK) which is based on a set of Branginskii-type equations. These comparisions show a very good agreement between Eulerian and PIC codes when computing the energy fluxes of both species, while the fluid code overestimates these fluxes. We also note that both the Eulerian and PIC codes yield similar results for the early burst of electrons, whereas the fluid code is not able to reproduce this effect, thus confirming its kinetic origin.

6.5. Full wave modeling of lower hybrid current drive in tokamaks

Participants: Pierre Bertrand, Simon Labrunie, Takashi Hattori, Jean Rodolphe Roche.

This work is performed in collaboration with Yves Peysson (DRFC, CEA Cadarrache).

The aim of this project is to develop a finite element numerical method for the full-wave simulation of electromagnetic wave propagation in a plasma. Full-wave calculations of the LH wave propagation is a challenging issue because of the short wave length with respect to the machine size. In the continuation of the works previously led in cylindrical geometry, a full toroidal description for an arbitrary poloidal cross-section of the plasma has been developed.
Since its wavelength $\lambda$ at the LH frequency is very small compared to the machine size $R$, a conventional full wave description represents a considerable numerical effort. Therefore, the problem is addressed by an appropriate mathematical finite element technique, which incorporates naturally parallel processing capabilities. This is particularly important aspect when simulations for plasmas of large size must be considered. It is based on a mixed augmented variational (weak) formulation taking account of the divergence constraint and essential boundary conditions, which provides an original and efficient scheme to describe in a global manner both propagation and absorption of electromagnetic waves in plasmas.

With such a description, usual limitations of the conventional ray tracing related to the approximation $\lambda << \phi_B << R$, where $\phi_B$ is the size of the beam transverse to the rf power flow direction, may be overcome. Since conditions are corresponding to $\lambda << \phi_B \sim R$, the code under development may be considered as a WKB full wave, dielectric properties being local.

The domain considered is as near as possible of the cavity filled by a tokomak plasma. Toroidal coordinates are introduced. In our approach we consider Fourier decomposition in the angular coordinate to obtain stationary Maxwell equations in a cross-section of the tokamak cavity.

A finite element method is proposed for the simulation of time-harmonic electromagnetic waves in a plasma, which is an anisotropic medium. The approach chosen here is sometimes referred to as full-wave modeling in the literature: the original Maxwell’s equations are used to obtain a second order equation for the time-harmonic electric field. These are written in a weak form using an augmented variational formulation (AVF), which takes into account the divergence. The variational formulation is then discretized using modified Taylor-Hood (nodal) elements.

During 2011 we introduced a new boundary condition in order to take account of the antenna and essential condition are considered in the code "FullWaveFEM" and new real case was considered.

### 6.6. Domain decomposition for the resolution of nonlinear equations

**Participant:** Jean Rodolphe Roche.

This a joint work with Noureddine Alaa, Professor at the Marrakech Cadi Ayyad University.

Strongly problems of parabolic equations have received considerable attentions, and various forms of this problems have been proposed in the literature, especially in the area of reaction-diffusion equations with cross-diffusion, such problems arise from biological, chemical and physical systems. Various methods have been proposed in the mathematical literature to study the existence, uniqueness and compute numerical approximation of solutions for quasi-linear partial differential equation problems. This year our we develop a numerical method to solve periodic non linear parabolic equations based on domain decomposition and optimization interior points method, see [34].

### 6.7. Inverse problem governed by Maxwell equations

**Participant:** Jean Rodolphe Roche.

This work is performed in collaboration with Jose Herskovits Norman of UFRJ, Rio de Janeiro, Antonio André Novotny from the LNCC, Petropolis, both from Brazil and Alfredo Canelas from the University of the Republic, Montevideo, Uruguay.

The industrial technique of electromagnetic casting allows for contactless heating, shaping and controlling of chemical aggressive, hot melts. The main advantage over the conventional crucible shape forming is that the liquid metal does not come into contact with the crucible wall, so there is no danger of contamination. This is very important in the preparation of very pure specimens in metallurgical experiments, as even small traces of impurities, such as carbon and sulphur, can affect the physical properties of the sample. Industrial applications are, for example, electromagnetic shaping of aluminum ingots using soft-contact confinement of the liquid metal, electromagnetic shaping of components of aeronautical engines made of superalloy materials (Ni,Ti, ...), control of the structure solidification.
The electromagnetic casting is based on the repulsive forces that an electromagnetic field produces on the surface of a mass of liquid metal. In the presence of an induced electromagnetic field, the liquid metal changes its shape until an equilibrium relation between the electromagnetic pressure and the surface tension is satisfied. The direct problem in electromagnetic casting consists in determining the equilibrium shape of the liquid metal. In general, this problem can be solved either directly studying the equilibrium equation defined on the surface of the liquid metal, or minimizing an appropriate energy functional. The main advantage of this last method is that the resulting shapes are mechanically stable.

The inverse problem consists in determining the electric currents and the induced exterior field for which the liquid metal takes on a given desired shape. This is a very important problem that one needs to solve in order to define a process of electromagnetic liquid metal forming.

In a previous work we studied the inverse electromagnetic casting problem considering the case where the inductors are made of single solid-core wires with a negligible area of the cross-section. In a second paper we considered the more realistic case where each inductor is a set of bundled insulated strands. In both cases the number of inductors was fixed in advance. This year we aim to overcome this constraint, and look for configurations of inductors considering different topologies with the purpose of obtaining better results. In order to manage this new situation we introduce a new formulation for the inverse problem using a shape functional based on the Kohn-Vogelius criterion. A topology optimization procedure is defined by means of topological derivatives, see [18] and [30]. To take account the free boundary evolution we consider a new level set method adapted to topological first and second order asymptotic topological analysis.

6.8. Diffusion of knowledge and methods towards other fields

Participant: Emmanuel Frénod.

Methods, results and more generally knowledge produced within Calvi team have been applied to environmental sciences In [22] and [50] asymptotic methods initially designed for tokamak plasmas are applied to coastal ocean waters linked phenomena.

[38] deals with the concept of confinement of paralic ecosystems. It improves an existing model in order to account for tide oscillations in any kind of geometry such as a non-rectangular lagoons with a non-flat bottom. The model, that relies on PDEs is then implemented thanks to the finite element method. Numerical results confirm the feasibility of confinement studies thanks to the introduced model.

Methods for mass transfer modeling was applied in the haulage context in [49].
5. New Results

5.1. Contributions earlier to 2011 but only published in 2011

Participants: Gérard Biau, Olivier Catoni, Sébastien Gerchinovitz, Vincent Rivoirard, Gilles Stoltz.

We do not discuss here the contributions provided by [12], [13], [14], [17], [18], [20], [22], [23], [25], since they were achieved in 2009 or earlier (but only published this year due to long queues in publication tracks of journals). [32] was revised but is still under review.

5.2. Sparse regression estimation

Participants: Gérard Biau, Olivier Catoni, Sébastien Gerchinovitz, Vincent Rivoirard, Gilles Stoltz, Jia Yuan Yu.

Sébastien Gerchinovitz and Jia Yuan Yu continued the work initiated by the former in the above-mentioned conference paper [25]; they derived from the sparsity results in individual sequences presented therein the minimax optimal rates of aggregation for individual sequences on $\ell^1$ balls. In particular, they exhibited, in certain cases, a phase transition between the $\ln T$ and the $\sqrt{T}$ behavior of the minimax regret, where $T$ denotes the number of instances. These results and all previous ones are summarized in the PhD thesis [10].

Other results were obtained in a stochastic framework, where input–output pairs are given by i.i.d. variables; they are described in the technical report [30]. Let $(X, Y)$ be a random pair taking values in $\mathbb{R}^p \times \mathbb{R}$. In the so-called single-index model, one has $Y = f^*(\theta^*^T X) + W$, where $f^*$ is an unknown univariate measurable function, $\theta^*$ is an unknown vector in $\mathbb{R}^d$, and $W$ denotes a random noise satisfying $\mathbb{E}[W|X] = 0$. The single-index model is known to offer a flexible way to model a variety of high-dimensional real-world phenomena. However, despite its relative simplicity, this dimension reduction scheme faces severe complications as soon as the underlying dimension becomes larger than the number of observations and this is why this estimation problem was considered from a sparsity perspective using a PAC-Bayesian approach.

Last but not least, we mention the edited book [29], which provides a modern overview on high-dimensional estimation.

5.3. Sequential learning with limited feedback; in particular, bandit problems

Participants: Gilles Stoltz, Jia Yuan Yu.

Some of the results cited below are summarized or stated as open problems in the habilitation thesis [11].

5.3.1. Bandit problems

We achieved three contributions. The first is described in the conference paper [27]: it revisits asymptotically optimal results of Lai and Robbins, Burnetas and Katehakis in a non-asymptotic way. The second is stated in the journal article [19] and is concerned with obtaining fast convergence rates for the regret in case of a continuum of arms (of course under some regularity and topological assumptions on the mean-payoff function $f$).

The third one is detailed in [24] and started from the following observation. Typical results in the bandit literature were of the following form: if the regularity of the mean-payoff function $f$ is known (or if a bound on it is known) then the regret is small. Actually, results were usually taking the following weaker form: when the algorithm is tuned with some parameters, then the regret is small against a certain class of stochastic environments. The question was thus to have an adaptive procedure, that, given one unknown environment (with unknown regularity), ensures that the regret is asymptotically small; even better, the desired aim was to control the regret in some uniform manner (in a distribution-free sense up to the regularity parameters). As described in this conference paper, a solution was achieved in the case of Lipschitz environments.
5.3.2. Approachability in games with partial monitoring

The conference paper [28] explains how we could re-obtain, in a simple, more straightforward, and computationally efficient manner a result proven by Perchet in his PhD thesis: the necessary and sufficient condition for the approachability of a closed convex set under partial monitoring.

5.4. Inference

Participant: Gérard Biau.

5.4.1. Geometric inference

This line of research is in collaboration with the Geometrica project-team (INRIA Saclay). As the latter says:

Due to the fast evolution of data acquisition devices and computational power, scientists in many areas are demanding efficient algorithmic tools for analyzing, manipulating and visualizing more and more complex shapes or complex systems from approximating data. Many of the existing algorithmic solutions which come with little theoretical guarantees provide unsatisfactory and/or unpredictable results. Since these algorithms take as input discrete geometric data, it is mandatory to develop concepts that are rich enough to robustly and correctly approximate continuous shapes and their geometric properties by discrete models. Ensuring the correctness of geometric estimations and approximations on discrete data is a sensitive problem in many applications.

Thus, motivated by a broad range of potential applications in topological and geometric inference, we introduce in [15] a weighted version of the $k$-nearest neighbor density estimator. Various pointwise consistency results of this estimator are established; the proposed method is also implemented to recover level sets in both simulated and real-life data.

Another problem of geometric inference is the following one, studied in [16]. Principal curves are nonlinear generalizations of the notion of first principal component. Roughly, a principal curve is a parameterized curve in $\mathbb{R}^d$ that passes through the “middle” of a data cloud drawn from some unknown probability distribution. Depending on the definition, a principal curve relies on some unknown parameters (number of segments, length, turn...) which have to be properly chosen to recover the shape of the data without interpolating. In this paper, we consider the principal curve problem from an empirical risk minimization perspective and address the parameter selection issue using the point of view of model selection via penalization. We offer oracle inequalities and implement the proposed approaches to recover the hidden structures in both simulated and real-life data.

5.4.2. Statistical inference

We still keep an eye on more traditional mathematical statistics; in particular, the technical report [31] takes place within this field. It shows, for a large class of distributions and large samples, that estimates of the variance $\sigma^2$ and of the standard deviation $\sigma$ are more often Pitman closer to their target than the corresponding shrinkage estimates which improve the mean squared error. The results thus indicate that Pitman closeness criterion, despite its controversial nature, should be regarded as a useful and complementary tool for the evaluation of estimates of $\sigma^2$ and of $\sigma$.

5.5. Statistical inference for biological systems based on a size-structured population

Participant: Vincent Rivoirard.
The journal paper [21] considers the problem of estimating the division rate of a size-structured population in a nonparametric setting. The size of the system evolves according to a transport-fragmentation equation: each individual grows with a given transport rate, and splits into two offsprings of the same size, following a binary fragmentation process with unknown division rate that depends on its size. In contrast to a deterministic inverse problem approach, this paper takes the perspective of statistical inference: the data consists in a large sample of the size of individuals when the evolution of the system is close to its time-asymptotic behavior, so that it can be related to the eigenproblem of the considered transport-fragmentation equation. By estimating statistically each term of the eigenvalue problem and suitably inverting a certain linear operator, it constructs a more realistic estimator of the division rate that achieves the same optimal error bound as in related deterministic inverse problems. The procedure relies on kernel methods with automatic bandwidth selection. It is inspired by model selection and recent results of Goldenschluger and Lepski.
6. New Results

6.1. Optimal control with singular arcs

Participants: Pierre Martinon, Andrei Dmitruk [Moscow State University], Pablo Lotito [U. Tandil, Argentina], Soledad Aronna, Frédéric Bonnans.

These studies enter in the framework of the PhD thesis of S. Aronna, supervised by J.F. Bonnans and P. Lotito, that ended in December 2011.

In the paper [21] we deal with optimal control problems for systems affine in the control variable. We have nonnegativity constraints on the control, and finitely many equality and inequality constraints on the final state. First, we obtain second order necessary optimality conditions. Secondly, we get a second order sufficient condition for the scalar control case. The results use in an essential way the Goh transformation.

In the report [22], we design a shooting algorithm applied to optimal control problems for which all control variables enter linearly in the Hamiltonian. This shooting algorithm is non standard, in particular since there are more equations than unknowns, and extends some previous algorithms designed for specific structures. We start investigating the case having only initial-final state constraints and free control variable, and afterwards we deal with control bounds. The shooting algorithm is locally well-posed and quadratically convergent if the derivative of its associated shooting function is injective at the optimal solution. The main result of this paper is to provide a sufficient condition for this injectivity, that is very close to the second order necessary condition. We prove that this sufficient condition guarantees the stability of the optimal solution under small perturbations and the well-posedness of the shooting algorithm for the perturbed problem. We present numerical tests that validate our method.

In the report [20] we deal with optimal control problems for systems that are affine in one part of the control variables and nonlinear in the rest of the control variables. We have finitely many equality and inequality constraints on the initial and final states. First we obtain second order necessary and sufficient conditions for weak optimality. Afterwards, we propose a shooting algorithm, and we show that the sufficient condition above-mentioned is also sufficient for the injectivity of the shooting function at the solution.

6.2. Characterization of a local quadratic growth of the Hamiltonian for control constrained optimal control problems

Participants: Frédéric Bonnans, Nikolai Osmolovskii [Systems Research Institute, Warsaw].

In the paper [25] we consider an optimal control problem with inequality control constraints given by smooth functions satisfying the hypothesis of linear independence of gradients of active constraints. For this problem, we formulate a generalization of strengthened Legendre condition and prove that this generalization is equivalent to the condition of a local quadratic growth of the Hamiltonian subject to control constraints.

6.3. Hamilton-Jacobi approach for deterministic control problems

Participants: Albert Altarovici, Olivier Bokanowski, Yingda Cheng [University of Texas], Anna Desilles, Nicolas Forcadel, Zhiping Rao, Chi-Wang Shu [Brown University], Hasnaa Zidani.
The paper [30] deals with deterministic optimal control problem with state constraints and non-linear dynamics. It is known for such a problem that the value function is in general discontinuous and its characterization by means of an HJ equation requires some controllability assumptions involving the dynamics and the set of state constraints. Here, we first adopt the viability point of view and look at the value function as its epigraph. Then, we prove that this epigraph can always be described by an auxiliary optimal control problem free of state constraints, and for which the value function is Lipschitz continuous and can be characterized, without any additional assumptions, as the unique viscosity solution of a Hamilton-Jacobi equation. The idea introduced in this paper bypass the regularity issues on the value function of the constrained control problem and leads to a constructive way to compute its epigraph by a large panel of numerical schemes. Our approach can be extended to more general control problems. We study in this paper the extension to the infinite horizon problem as well as for the two-player game setting. Finally, an illustrative numerical example is given to show the relevance of the approach.

In [34], [19] we study an optimal control problem governed by measure driven differential systems and in presence of state constraints. First, under some weak invariance assumptions, we study in [19] the properties of the value function and obtain its characterization by means of an auxiliary control problem of absolutely continuous trajectories. For this, we use some known techniques of reparametrization and graph completion. Then we give a characterization of the value function as the unique constrained viscosity solution of a Hamilton-Jacobi equation with measurable time dependant Hamiltonians.

The general case without assuming any controllability assumption is considered in [34]. We prove that the optimal solutions can still be obtained by solving a reparametrized control problem of absolutely continuous trajectories but with time-dependent state-constraints.

The paper [17] deals with minimal time problems governed by nonlinear systems under general time dependent state constraints and in the two-player games setting. In general, it is known that the characterization of the minimal time function, as well as the study of its regularity properties, is a difficult task in particular when no controllability assumption is made. In addition to these difficulties, we are interested here to the case when the target, the state constraints and the dynamics are allowed to be time-dependent. We introduce a particular reachability control problem, which has a supremum cost function but is free of state constraints. This auxiliary control problem allows to characterize easily the backward reachable sets, and then, the minimal time function, without assuming any controllability assumption. These techniques are linked to the well known level-set approaches. Our results can be used to deal with motion planning problems with obstacle avoidance, see [16].

Several works have been also carried out in the domain of numerical methods of HJB equations. The paper [31] aims at studying a discontinuous Galerkin scheme for front propagation with obstacles. We extend a first work published in [11], to propose a simple and direct discontinuous Galerkin (DG) method adapted to such front propagation problems. We follow the formulation of [12], leading to a level set formulation driven by a Hamilton-Jacobi variational inequality. The DG scheme is motivated by the variational formulation when the equation corresponds to linear convection problems in presence of obstacles. The scheme is then generalized to nonlinear equations, written in an explicit form. Stability analysis are performed for the linear case with Euler forward, a Heun scheme and a Runge-Kutta third order time discretization. Several numerical examples are provided to demonstrate the robustness of the method. Finally, a narrow band approach is considered in order to reduce the computational cost.

### 6.4. Stochastic programming

**Participants:** Frédéric Bonnans, Zhihao Cen, Thibault Christel [Total].

In [29] we consider a model of medium-term commodity contracts management. Randomness takes place only in the prices on which the commodities are exchanged whilst state variable is multi-dimensional. In our previous article, we proposed an algorithm to deal with such problem, based on quantization of random process and a dual dynamic programming type approach. We obtained accurate estimates of the optimal value and a suboptimal strategy from this algorithm. In this paper, we analyse the sensitivity with respect to
parameters driving the price model. We discuss the estimate of marginal price based on the Danskin’s theorem. Finally, some numerical results applied to realistic energy market problems have been performed. Comparisons between results obtained by our algorithm and other classical methods are provided and evidence the accuracy of the estimate of marginal prices.

6.5. Stochastic control

**Participants:** Frédéric Bonnans, Xiaolu Tan [CMAP], Imene Ben Latifa, Mohamed Mnif [ENIT, Tunis].

In [24], we extend a study by Carmona and Touzi on an optimal multiple stopping time problem in a market where the price process is continuous. In this paper, we generalize their results when the price process is allowed to jump. Also, we generalize the problem associated to the valuation of swing options to the context of jump diffusion processes. Then we relate our problem to a sequence of ordinary stopping time problems. We characterize the value function of each ordinary stopping time problem as the unique viscosity solution of the associated Hamilton-Jacobi-Bellman Variational Inequality.

In [27], we consider, in the framework of Galichon, Henry-Labordère and Touzi, the model-free no-arbitrage bound of variance option given the marginal distributions of the underlying asset. We first make some approximations which restrict the computation on a bounded domain. Then we propose a gradient projection algorithm together with a finite difference scheme to approximate the bound. The general convergence result is obtained. We also provide a numerical example on the variance swap option.

6.6. Stochastic control of an hybrid vehicle

**Participants:** Kamal Aouchiche [Renault], Frédéric Bonnans, Giovanni Granato, Hasnaa Zidani.

In the CDC paper [18] we present a stochastic dynamic programming (SDP) algorithm that aims at minimizing an economic criteria based on the We also work on a stochastic dynamic programming (SDP) algorithm that aims at minimizing an economic criteria based on the total energy consumption of a range extender electric vehicle (REEV). This algorithm integrates information from the REEV’s navigation system in order to obtain some information about future expected vehicle speed. The model of the vehicle’s energetic system, which consists of a high-voltage (HV) battery, the main energy source, and an internal combustion engine (ICE), working as an auxiliary energy source), is written as a hybrid dynamical system and the associated optimization problem in the hybrid optimal control framework. The hybrid optimal control problem includes two important physical constraints on the ICE, namely, an activation delay and a decision lag. Three methods for the inclusion of such physical constraints are studied. After introducing the SDP algorithm formulation we comment on numerical results of the stochastic algorithm and its deterministic counterpart.

6.7. Optimal control of PDEs

**Participants:** Frédéric Bonnans, Francisco Silva [U. Roma], Térence Bayen [U. Montpellier II].

In the report [23] we consider an optimal control problem of a semi-linear elliptic equation, with bound constraints on the control. Our aim is to characterize local quadratic growth for the cost function \(J\) in the sense of strong solutions. This means that the function \(J\) grows quadratically over all feasible controls whose associated state is close enough to the nominal one, in the uniform topology. The study of strong solutions, classical in the Calculus of Variations, seems to be new in the context of PDE optimization. Our analysis, based on a decomposition result for the variation of the cost, combines Pontryagin’s principle and second order conditions. While these two ingredients are known, we use them in such a way that we do not need to assume that the Hessian of Lagrangian of the problem is a Legendre form, or that it is uniformly positive on an extended set of critical directions.

6.8. Global optimization of pipe networks by the interval analysis approach: the Belgium network case

**Participants:** Frédéric Bonnans, Grégoire Spiers, Jean-Léopold Vie.
In [26] we have shown that a classical test problem for the optimization of gas networks, namely the so-called Belgium gas network, could be solved by global optimization techniques. Until now only local algorithms had been used for solving this problem. Using techniques based on interval analysis and constraint propagation we actually recover (and therefore justify) the solution computed so far.
6. New Results

6.1. Convergence of adaptive finite element algorithms

Participants: Roland Becker, Shipeng Mao, David Trujillo.

The theoretical analysis of mesh-adaptive methods is a very active field of research. We have generalized our previous results concerning optimality of adaptive methods to nonconforming finite elements [49]. Our results include the error due to iterative solution of the system matrices by means of a simple stopping criterion related to the error estimator. The main difficulty was the treatment of the nonconformity which leads to a perturbation of the orthogonality relation at the heart of the proofs for conforming finite elements. We have been able to extend this result to the Stokes equations, considering different lowest-order nonconforming finite elements on triangular and quadrilateral meshes [16].

In [17] we have shown that the smallness assumption required in all former proofs of optimality of adaptive finite element methods can be overcome, at least in some situations.

Finally, we have shown optimality of a new goal-oriented method in [19].

Our theoretical studies, which are motivated by the aim to develop better adaptive algorithms, have been accompanied by software implementation with the Concha library, see Section 5.1. It hopefully opens the door to further theoretical and experimental studies.

6.2. Finite element methods for interface problems

Participants: Nelly Barrau, Roland Becker, Robert Luce, Erik Burman, Peter Hansbo.

The original formulation of NXFEM [62] is based on the doubling of elements. In some situations, as the case of a moving interface, it is computationally more convenient to have a method with local enrichment, as for the standard XFEM. In [41] we have developed such an approach based on NXFEM. We have developed an hierarchical formulation for a fictitious domain formulation in [13].

One of the technical difficulties is the simultaneous robustness of the method with respect to the size of the intersection of a mesh cell with the interface and with respect to the discontinuous diffusion parameters. This is the subject of the thesis of Nelly Barrau, supervised by Robert Luce and Eric Dubach (LMAP).

6.3. Discontinuous finite element methods

Participants: Roland Becker, Daniela Capatina, Julie Joie.

We have developed a new discontinuous Galerkin scheme for the Stokes equations and corresponding three-fields formulation. In this work, which is part of the Phd Thesis of Julie Joie, we introduce a modification of the stabilization term in the standard DG-IP method. This allows for a cheaper implementation and has a more robust behavior with respect to the stabilization parameter; we have shown convergence towards the solution of non-conforming finite element methods for linear, quadratic and cubic polynomial degrees. This scheme has been extended to the three-field formulation of the Stokes problem, which is a further step towards the polymer project of Section 4.2. Since it is well known that the non-conforming finite element approximations do not verify the discrete Korn inequality, an appropriate further stabilization term is introduced, see [42].

6.4. Stabilized finite element methods

Participants: Roland Becker, Erik Burman, Peter Hansbo.
We have developed a new stabilized finite element formulation based on implicit penalization of the singularities of higher-order derivatives of continuous finite element spaces in [12]. It has been applied to the convection-diffusion problem as well as to the incompressible Euler equations.

6.5. A posteriori error estimators based on $H'(\text{div})$-reconstructed fluxes

Participants: Roland Becker, Daniela Capatina, Robert Luce.

Mesh adaptivity is nowadays an essential tool in numerical simulations; in order to achieve it, reliable and efficient, easily computable a posteriori error estimators are needed. Such estimators obtained by reconstructing locally conservative fluxes in the Raviart-Thomas finite element space have been largely employed in the past years.

We have so far considered the convection-diffusion equation and proposed a unified framework for several finite element approximations (conforming, nonconforming and discontinuous Galerkin). The main advantage of our approach is to use, contrarily to the existing references, only the primal mesh for the flux reconstruction, which presents certain facilities from a computational point of view.

For this purpose, the construction of the $H'(\text{div})$-vector involved in the error estimator is inspired by the hypercircle method cf. [53] and is achieved on patches, which may overlap. A patch depends on the type of the employed finite elements and is defined as the support of a basis function.

Our first results were presented in [26]. We are working on the extension to higher-order approximations, to quadrilateral meshes and to other model problems.

6.6. A posteriori error analysis of sensitivities

Participants: Roland Becker, Daniela Capatina, Robert Luce, David Trujillo.

Most practical applications involve parameters $q = (q_i)_{1 \leq i \leq N}$ of different origins: physical (viscosity, heat conduction), modeling (computational domain, boundary conditions) and numerical (mesh, stabilization parameters, stopping criteria, values of a turbulence model). Numerical simulations can provide information related to the (first order) sensitivity of a quantity of physical interest $I(q)$ with respect to different parameters: $\frac{\partial I}{\partial q_i}$. Their computation can help to validate the physical model, to explain unexpected behaviour and also to guide efforts to improve both the physical and the computational models.

A posteriori error estimates for the functional itself, for fixed values of the parameters $q$, are well-known, cf. for example [47] where a goal-oriented error control is achieved by introducing an adjoint problem. Our goal is to provide a general framework for the a posteriori error estimation of sensitivities $\frac{\partial I}{\partial q_i} - \frac{\partial I}{\partial h_i}/\partial q_i$, which has not been given yet in the literature.

So far, we have applied the proposed method to the computation of the Nusselt number measuring the efficiency of a cooling process, described in the project Optimal. A cold liquid is injected in an annular domain through several inlets in order to cool a heated interior stator.

First numerical results, including adaptation with respect to the functional and to the sensitivity, have been carried out with the library Concha. They have been presented in [33], [30]. In Figure 10 one may see the computed temperature and velocity field, while the a posteriori error estimator for the sensitivity of the Nusselt number with respect to the inflow speed at the right-hand side inlet, as well as the adapted mesh, are given in Figure 11.

In the future, several important aspects related to the adaptive method are still to be investigated such as design of an appropriate adaptive algorithm, proof of its convergence and optimality etc.

6.7. Viscoelastic fluids modeling and numerical simulation

Participants: Roland Becker, Daniela Capatina, Didier Graebling, Julie Joie.

We have continued our activities with respect to numerical simulations of polymer flows.
Figure 10. Computed velocity and temperature fields.

Figure 11. Error estimator for the sensitivity and adapted mesh.
We have further validated [32] the code on both triangular and quadrilateral 2D meshes, by using a mixed non-conforming/DG method. In the case of Giesekus model, comparisons of velocity profile with semi-analytical solutions for Poiseuille flow [70] and with experimental data [73] for the 4:1 contraction are respectively shown in Figures 12 and 13).

The quadrilateral scheme needs an additional stabilization term, in order to ensure uniform consistency and stability for the underlying linear problem. It has been tested in particular on the benchmark case of flow around a 2D cylinder, for which our code converged for high values of the Weissenberg number (We $>$ 70).

We have computed the drag and compared it, see Figure 14, with numerical data found in the literature [58], [65], [76], [57] in the case of the Oldroyd-B model.

The system associated to this quasi-linear model has been solved on rather fine meshes (of about $10^6$ elements) thanks to a multigrid solver based on a Vanka type preconditioner. We have worked on the extension of the multigrid method to more complex and more realistic polymer models, involving nonlinear terms in the constitutive law such that Phan Thien-Tanner and Giesekus models.

The unsteady case has also been treated for different geometries.

The parallelization of the library is ongoing work as described in the section 5.3; first tests have been carried on in the case of viscoelastic liquids.

In order to better describe real polymer flows, the thermal aspects have to be taken into account. As a first step, the coupling of the flow with the energy equation in the Newtonian case was considered in [38], by using a conservative variable. Our next objective is to simulate anisothermal viscoelastic flows, including typical effects such that thermo-dependent viscosity and viscous heating.

With these tools, the long-term goal is to successively build up robust and efficient software in order to tackle design problems.

Moreover, we intend to consider other models of non-Newtonian fluids employed in other application domains such as biomedicine.

6.8. Positivity preserving schemes
Participants: Roland Becker, Daniela Capatina.
Figure 13. Comparison with experimental data for the 4:1 contraction.

<table>
<thead>
<tr>
<th>λ</th>
<th>CONCHA</th>
<th>Dou et al.</th>
<th>Hulsen et al.</th>
<th>Étienne et al.</th>
<th>Damanik et al.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>132,357</td>
<td>131,809</td>
<td>132,358</td>
<td>132,33</td>
<td>-</td>
</tr>
<tr>
<td>0.3</td>
<td>123,190</td>
<td>123,514</td>
<td>123,193</td>
<td>123,41</td>
<td>123,194</td>
</tr>
<tr>
<td>0.6</td>
<td>117,780</td>
<td>120,485</td>
<td>117,792</td>
<td>-</td>
<td>117,779</td>
</tr>
</tbody>
</table>

Figure 14. Comparison of drag coefficients with the literature.
The stability and robustness with respect to physical parameters of numerical schemes for polymer flows is a challenging question. Indeed, most algorithms encounter serious convergence problems for large Weissenberg numbers. In the recent years, this issue has been associated to the discrete positive definiteness of the so-called conformation tensor. It seems therefore essential for the numerical simulations to employ positivity preserving schemes.

In order to develop such schemes, we have adopted the approach proposed by Lee and Xu \[ 68 \] in the case of the quasi-linear Oldroyd-B model, based on the similarities between its constitutive equation and differential Riccati equations. We have applied it to a more general matrix equation; a typical example is the nonlinear Giesekus constitutive law. In agreement with our code (see Section 6.7), we have discretized it by a discontinuous Galerkin method combined with an upwinding of the transport terms, whereas the approach of Lee and Xu relies essentially on the characteristics method.

We have shown that a modification of Newton’s method yields a monotone and positive scheme, under certain hypothesis, and we have applied this study to both Oldroyd-B and Giesekus polymer models. This allowed us to better understand and explain the better behaviour of the latter for large Weissenberg numbers. These results have been presented in \[ 24 \],\[ 31 \].

We are working on several challenging questions such as the extension to other discretization methods, the improvement of the iterative method or the derivation of energy estimates for the coupled system.

6.9. Discretization of Euler’s equations

**Participants:** Roland Becker, Kossivi Gopki, Eric Schall, David Trujillo.

Over the past years, significant advances have been made in developing discontinuous Galerkin finite element methods (DGFEM) for applications in fluid flow and heat transfer. Certain features of the method have made it attractive as an alternative to other popular methods such as finite volume and more convenient finite element methods in thermal fluid engineering analyses. The DGFEM has been used successfully to solve hyperbolic systems of conservation laws. It makes use of the same local function space as the continuous method, but with relaxed continuity at inter-element boundaries. Since it uses discontinuous piecewise polynomial bases, the discretization is locally conservative and in the considered lowest-order case, the method preserves the maximum principle for scalar equations.

One of the challenges in Computational Fluid Dynamic (CFD) is to obtain as accurate as possible the solution of the problem under consideration at very low cost in terms of computational time. So our principal work is to find some relevant and robust strategies and technics of meshes adaptation in order to concentrate just the calculation where there are physical phenomena to capture. From Industrial point of view, the aim is to get the stationary solution as quick as possible with as much accuracy as possible. The main limitation of these results in CFD concern the underlying models: for example, nearly nothing seems to be known for (even linear) first-order systems or for realistic nonlinear equations. We therefore have developed different modern techniques, especially adaptive methods, to tackle this kind of problems in compressible CFD. The strategy is to iteratively improve the quality of the approximate solutions based on computed information (a posteriori error analysis). In this way, a sequence of locally refined meshes is constructed, which allows for better efficiency as compared to more classical approaches in the presence of different kind of singularities. The main goal is to improve the aerodynamical design process for complex configurations by significantly reducing the time from geometry to solution at engineering-required accuracy using high-order adaptive methods.

One of our strategies of refinement is based on the creation of hanging nodes commonly called non-conforming refinement. The figures 15 show superposition of two kinds of meshes. One is a non-conforming refined mesh (black color) and the other one is the initial grid (red color) on which the refinement has been performed. It shows the technic of cutting the cells where singularities occur in the scramjet inlet.

The mesh adaptation is designed using some criteria as a posteriori error estimates. We have designed criteria based on the calculation of the jump of physical quantities like density, pressure, entropy, temperature and mach number at the inter-element. This criteria seems to be a very good indicator for the mesh adaptation. Figure 16 is the comparison of isoline of the density in scramjet internal flow at mach 3 of the initial mesh,
the third and the sixth mesh after refinement. The indicator used is the density jump. It shows the impact and the accuracy of the solution obtained after the sixth iteration of the refinement.

![Superposition of non-conforming adapted black color grid and initial grid (red color) - (a) quadrangles and (b) triangles.](image)

*Figure 15. Superposition of non-conforming adapted black color grid and initial grid (red color) – (a) quadrangles and (b) triangles.*

The figure 17 shows the streamlines of the density in the scramjet inlet after the seventh iteration. This shows how the adaptation depicts almost clearly and accurately the shock waves and the expansion waves and their interactions in the domain.

![Cutlines along the symmetry axis of various meshes for the scramjet test case](image)

*Figure 16. Cutlines along the symmetry axis of various meshes for the scramjet test case*
Figure 17. Density streamlines on grid obtained after the seventh iteration of adaptive refinement procedure with density jump as indicator.

Figure 18 represent the density isolines of a flow past cylinder test case using the non-conforming mesh adaptation with quadrangular an triangular grids.

We have also settled another indication which is hierarchical. It measures the difference of $g_h$ with the physical quantity $g_{h/2}$ obtained by computation on a globally refined mesh $h/2$. This allows us the make comparison.
with the previous indicator. The case test considered for this comparison is an external flows past a cylinder airfoil at fixed free stream conditions: $M_\infty = 3$. The result is quite surprising the way one type of indicator can capture phenomenon that are not capture by the another one. In fact the hierarchical indicator seems to capture recirculation downstream to the obstacle which was not capture by the jump indicator (see figure 19).

![Streamlines coloured by the density on meshes generated with hierarchical indicator (a) and with jump indicator (b)](image)

Figure 19. Streamlines coloured by the density on meshes generated with hierarchical indicator (a) and with jump indicator (b)

We compare the computational time between a non-conforming mesh refinement and a globally mesh refined with nearly the same amount of cells. The meshes contain quadrangles or triangles. We can observe through the following tables that the adapted meshes wether triangular or quadrangular meshes allow to save 20 to 90 times the computational time than the normal globally refined mesh. (see tables 1 and 2)

<table>
<thead>
<tr>
<th>Scramjet test case at mach=3</th>
<th>Flow past cylinder test case at mach=3</th>
</tr>
</thead>
<tbody>
<tr>
<td>ScramQuad_4</td>
<td>Nodes: 17043, Cells: 15485, Segments: 34938, Compt. Time: 25.0236s</td>
</tr>
<tr>
<td>ScramQuad_Uniform</td>
<td>Nodes: 17182, Cells: 16640, Segments: 33824, Compt. Time: 865.077s</td>
</tr>
<tr>
<td>ScramTri_4</td>
<td>Nodes: 9951, Cells: 17005, Segments: 29138, Compt. Time: 22.3141s</td>
</tr>
<tr>
<td>CylQuad_5</td>
<td>Nodes: 11203, Cells: 10174, Segments: 23105, Compt. Time: 47.2187s</td>
</tr>
<tr>
<td>CylQuad_Uniform</td>
<td>Nodes: 10496, Cells: 10240, Segments: 20736, Compt. Time: 814.6168s</td>
</tr>
<tr>
<td>CylTri_6</td>
<td>Nodes: 6480, Cells: 10867, Segments: 19264, Compt. Time: 79.7836s</td>
</tr>
<tr>
<td>CylTri_Uniform</td>
<td>Nodes: 6032, Cells: 11776, Segments: 17808, Compt. Time: 4258.6618s</td>
</tr>
</tbody>
</table>

Table 1 Table 2

Figure 20. Comparison of computational times

In table 1, the gain in time is 35 times in quadrangular grid case and 90 times triangular ones and in table 2, the gain in time: 18 times in quadrangular grid case and 58 times triangular ones. So one can say that the adaptive mesh with the strategies and techniques we have settled are efficient and robust in capturing physical phenomenon at a very reasonable low cost.

In concluding, the procedure of refinement permit to save computational time and have good accuracy of the approximated solution computed. Our focus is to continue the improve our methods and strategies in order to meet the requirement of accuracy, robustness and efficiency. Many other works are in hand such as slope limiters for high-order Discontinuous Galerkin, low mach number computation with some remarkable approaches.
6. New Results

6.1. Analysis and control of fluids and of fluid-structure interactions

Participants: Thomas Chambrion, Antoine Henrot, Alexandre Munnier, Yu Ning Liu, Jean-François Scheid, Erica Schwindt, Mario Sigalotti, Takéo Takahashi, Marius Tucsnak, Jean-Claude Vivalda, Jérôme Lohéac.

The study of a fluid-structure system depends on the nature of the fluid considered and in particular on the Reynolds number. We have split the new results of this section according to the viscosity of the fluid. The first part is devoted to the case of a viscous fluid. This is the case that has received more attention from mathematicians in the recent years. In the second part, we have put the results concerning an inviscid fluid. This case is more classical in Fluid Mechanics and could be more interesting to understand self-propelled motions which is one of the main goal of our work. In the last part, we have given some numerical results.

6.1.1. Incompressible viscous fluids

- In [31], García and Takahashi present some abstract results giving a general connection between null-controllability and several inverse problems for a class of parabolic equations. They obtain some conditional stability estimates for the inverse problems consisting of determining the initial condition and the source term, from interior or boundary measurements. They apply this framework for Stokes system with interior and boundary observations, for a coupling of two Stokes system and a linear fluid-structure system.

- Nečasová, Takahashi and Tucsnak consider in [43] the three-dimensional motion of a self-propelled deformable structure into a viscous incompressible fluid. The deformation of the solid is given whereas its position is unknown. Such a system could model the propulsion of fish-like swimmers. The equations of motion of the fluid are the Navier-Stokes equations and the equations for the structure are deduced from Newton’s laws. The corresponding system is a free-boundary problem and the main result they obtain is the existence of weak solutions for this problem.

- In [29] we give a controllability result for a simplified 1D fluid-structure system.

- In [39] we give a detailed analysis of a phase field type model describing the motions of vesicles in a viscous incompressible fluid.

- In [40] we study a controllability problem for a simplified one dimensional model for the motion of a rigid body in a viscous fluid. One of the novelties brought in with respect to the existing literature consist in the fact that we use a single scalar control. Moreover, we introduce a new methodology, which can be used for other nonlinear parabolic systems, independently of the techniques previously used for the linearized problem. This methodology is based on an abstract argument for the null controllability of parabolic equations in the presence of source terms and it avoids tackling linearized problems with time dependent coefficients.

6.1.2. Ideal fluids

- In [42], the author studies the motion of an hyperelastic body immersed in a perfect fluid. The recourse to a strain energy density function in the modeling allows many different constitutive equations for the hyperelastic material to be considered. Numerical simulation are performed, aiming to study passive locomotion (i.e. locomotion at zero energy cost).

- In [27], we study the approximate controllability of 2D swimmer in an ideal fluid. The result includes an approximate tracking result of both the shape and the position of the swimmer.

6.2. Frequency tools for the analysis of PDE’s

Participants: Xavier Antoine, Bruno Pinçon, Karim Ramdani, Bertrand Thierry, Marius Tucsnak.
Our contribution in this direction mainly concerns the numerical approximation of scattering problems.

- In [21], we propose some strategies to solve numerically the difficult problem of multiple scattering by a large number of disks at high frequency. To achieve this, we combine a Fourier series decomposition with the EFIE integral equation. Numerical examples will be presented to show the efficiency of our method.

- In [20], we propose to simulate complex nonlinear physics problems related to the Schrödinger equations by using relaxation techniques coupled with absorbing boundary conditions or PMLs. This shows that these two methods are much more accurate than the usual complex scaling/absorbing potential approaches widely used in physics for domain truncation.

- In [19], complete high order absorbing boundary conditions are proposed, discretize and simulate for one- and two-dimensional nonlinear Schrödinger equations. In [38], we propose new accurate absorbing boundary condition for computing nonlinear eigenvalue problems related to the Schrödinger equation.

- In [57], we propose a review of how pseudo differential operators theory help in building analytical preconditioners and well-posed integral equations for acoustics scattering. In [26], we propose a new efficient and robust domain decomposition method for solving large scale three-dimensional acoustic scattering problems.

6.3. Observability, controllability and stabilization in the time domain

Participants: Fatiha Alabau, Xavier Antoine, Thomas Chambrion, Antoine Henrot, Karim Ramdani, Lionel Rosier, Mario Sigalotti, Takéo Takahashi, Marius Tucsnak, Jean-Claude Vivalda, Ghislain Haine, Roberto Guglielmi.

6.3.1. Observability

- The PhD of Ghislain Haine is devoted to the analysis of observers based techniques for solving inverse problems. In [34], we provide a convergence analysis of the iterative reconstruction algorithm proposed by Ramdani et al. in [81]. More precisely, we propose a complete numerical analysis for semi-discrete (in space) and fully discrete approximations of the iterative algorithm using finite elements in space and an implicit Euler method in time. In order to disseminate our reconstruction method in the community of Automatic and control engineering, we wrote an engineer’s oriented note [33] presenting the main ideas of our algorithm.

6.3.2. Control

- In [48], we develop a model that describes the impact of the amount of soot in the filter on the Diesel engine performance. This model is used to determine the optimal amount of soot on which the regeneration of the particulate filter shall start.

- In [49], we give sufficient conditions for the simultaneous approximate controllability of a bilinear Schrödinger equation driven by a single scalar control in the case where every energy level is non-degenerate and the control potential couples each pair of energy levels.

- In [25], we give sufficient conditions for the simultaneous approximate controllability of a bilinear Schrödinger equation driven by a single scalar control under a generic condition of coupling of all energy levels via a chain of non-degenerate transitions. The result applies for systems with degenerate energy levels or when the coupling operator does not couple directly each pair of energy levels.

- In [16], we prove exact controllability for symmetric coupled wave equations by a single control in the case of coupling and control regions which do not intersect. For this, we use and extend the two-level energy method introduced by Alabau-Boussouira (2001, 2003). Using transmutation, we derive null controllability results for coupled parabolic and Schrödinger equations. This is the first
positive quantitative result, in a multi-dimensional framework with control and coupling regions with empty intersection. Such questions have been considered using Carleman estimates but no positive quantitative results could be derived in the case of control and coupling regions which do not intersect.

- In [30] we propose a new method for the approximation of exact controls of a second order infinite dimensional system with bounded input operator. The algorithm combines Russell’s “stabilizability implies controllability” principle with the Galerkin method. The main new feature of this work consists of giving precise error estimates.

6.3.3. Stabilization

- In [44] we consider the wave equation with a time-varying delay term in the boundary condition in a bounded and smooth domain. We prove exponential stability of the solution, by introducing suitable energies and Lyapunov functionals. Such analysis is also extended to a nonlinear case.

- In [52] we present a course on stabilization of hyperbolic equations given at a CIME session on Control of PDE’s in Italy in July 2010, including well-known results, together with recent ones including nonlinear stabilization, memory-damping and stabilization of coupled systems by a reduced number of controls. In particular, we present the optimal-weight convexity method (Alabau-Boussouira 2005, 2010) in both the finite dimensional and infinite dimensional framework and give applications to semi-discretization of hyperbolic PDE’s.

- In [14], we consider stabilization of coupled systems of hyperbolic PDE’s with hybrid boundary conditions, by a reduced number of closed loop globally distributed controls. We establish polynomial stabilization for such systems under a new compatibility condition. We also derive decay rates for explicit initial data using interpolation theory.

- In [15], we consider stabilization of coupled systems of wave-type, with localized couplings and either localized internal closed loop controls or boundary control. We establish polynomial decay rates for coupling and damping regions which do not intersect in the one-dimensional case. We also derive results in the multi-dimensional case, under multiplier type conditions for both the coupling and damping regions. The novelty and difficulty is to consider localized couplings.

- In [13] we give a constructive proof of Gibson’s stability theorem, some extension and further positive and negative applications of this result.

- Very few lower energy estimates are available in the literature. The main one has been proved in the one-dimensional case for a locally distributed power-like damping for the wave equation in 1995 by Haraux. This approach does not generalize to multi-dimensional cases and for systems of equations. In [11], we prove strong energy and weak velocity lower estimates for the nonlinearly damped Timoshenko beams (coupled system), and for Petrowsky equations in two space dimensions.

- In [12], we show that if a linear system is observable through a locally distributed (resp. boundary) observation, then any dissipative nonlinear feedback locally distributed (resp. active only on a part of the boundary) stabilize the system and we give quasi-optimal energy decay rates, under the optimal condition of geometric optics of Bardos-Lebau-Rauch (1992). The approach is based on the optimal-weight convexity method (Alabau-Boussouira 2005, 2010). Our results generalize previous results by Haraux (1989) and Ammari and Tucsnak (2001) for linear feedbacks.

- In [17], we study the stabilization of Bresse system, which models vibrations of a beam through three coupled wave equations. We establish polynomial stabilization of the full system by a single feedback control.

- In [23], Badra (University of Pau) and Takahashi consider the stabilization of the system \( y' = Ay + Bu \) where \( A : \mathcal{D}(A) \rightarrow X \) be the generator of an analytic semigroup and \( B : U \rightarrow \mathcal{D}(A^*) \) a quasi-bounded operator. They consider controls \( u \) which are the linear combination of a finite family \( (v_1, ..., v_K) \). They show that if \( (A^*, B^*) \) satisfies a unique continuation property and if \( K \) is greater or equal to the maximum of the geometric multiplicities of the
the unstable modes of \( A \), then the system is generically stabilizable with respect to the family \((v_1, \ldots, v_K)\). With the same functional framework, they also prove the stabilizability of a class of nonlinear system when using feedback or dynamical controllers. They apply these results to stabilize the Navier–Stokes equations in 2D and in 3D by using boundary control with an optimal number of controllers.

- In [32] we tackle an unsolved difficulty in the control of vibrating systems, consisting in the fact that a small delay in the application of feedback control may destroy the stabilizing effect of the control. We consider a vibrating string that is fixed at one end and stabilized with a boundary feedback with delay at the other end and we show that certain delays (large, in general) in the boundary feedback preserve the exponential stability of the system.

- In [18] we consider \( N \) Euler-Bernoulli beams and \( N \) strings alternatively connected to one another and forming a chain beginning with a string. We study the strong and polynomial stabilities of this system on this network and the spectrum of the corresponding conservative system.

- In [45] we study the asymptotic behavior of the solution of the non-homogeneous elastic systems with voids and a thermal effect. Our main results concern strong and polynomial stabilities (since this system suffers of exponential stability).

- In [24] we are interested in an inverse problem for the wave equation with potential on a star-shaped network. We prove the Lipschitz stability of the inverse problem consisting in the determination of the potential on each string of the network with Neumann boundary measurements at all but one external vertices. Our main tool, proved in this article, is a global Carleman estimate.

- In [35] we consider switched systems on Banach and Hilbert spaces governed by strongly continuous one-parameter semigroups of linear evolution operators. We provide necessary and sufficient conditions for their global exponential stability, uniform with respect to the switching signal, in terms of the existence of a Lyapunov function common to all modes.

- In [47], we investigate sufficient conditions for the convergence to zero of the trajectories of linear switched systems. We apply our result to the synthesis of an observer for the three-cell converter.

6.3.4. Other problems

- In [37], we study a spectral problem related to a reaction-diffusion model where the preys and the predators do not live on the same area. We are interested in the optimal zone where the control should take place. First we prove existence of an optimal domain in a natural class. Then, it seems plausible that the optimal domain is localized in the intersection of the living areas of the two species. We prove this fact in one dimension for small size of domains.

- In [41], we explain why Donnelly’s proof of the gap conjecture is not correct.

- In [22], we study the set of points, in the plane, defined by \( \{(x, y) = (\lambda_1(\Omega), \lambda_2(\Omega)), |\Omega| = 1\} \), where \( (\lambda_1(\Omega), \lambda_2(\Omega)) \) are either the two first eigenvalues of the Dirichlet-Laplacian, or the two first non trivial eigenvalues of the Neumann-Laplacian. We consider the case of general open sets together with the case of convex open domains. We give some qualitative properties of these sets, show some pictures obtained through numerical computations and state several open problems.

- In [28], we look for the minimizers of the functional \( J_{\lambda}(\Omega) = \lambda |\Omega| - P(\Omega) \) among planar convex domains constrained to lie into a given ring. We prove that, according to the values of the parameter \( \lambda \), the solutions are either a disc or a polygon. In this last case, we describe completely the polygonal solutions by reducing the problem to a finite dimensional optimization problem. We recover classical inequalities for convex sets involving area, perimeter and inradius or circumradius and find a new one.
6. New Results

6.1. On the Existence of Strict Optimal Controls for Constrained, Controlled Markov Processes in Continuous-Time

Participant: François Dufour.

Closedness and convexity conditions are identified under which optimal controls in the class of strict controls exist for a large class of stochastic processes under infinite-horizon discounted, long-term average, first exit, finite-horizon and discretionary stopping criteria in the presence of hard and/or soft constraints. The results are more general than results obtained by Haussmann and Lepeltier for a controlled diffusion under a mixed optimal-stopping/finite-horizon/first-exit criterion. The approach taken in this work is to utilize equivalent linear programming formulations of the control problems which provides a unified LP formulation for the problems. The conditions of Haussmann and Lepeltier are shown to imply the sufficient conditions of this paper when the process is a controlled diffusion. Simpler conditions are also identified for Markov chains, simple Markov jump processes, diffusions with jumps, regime-switching diffusions and solutions to Levy stochastic differential equations.

These results have been obtained in collaboration with Richard Stockbridge, Department of Mathematical Sciences, University of Wisconsin Milwaukee, USA. It has been accepted for publication in Stochastics [14].

6.2. Approximation of Markov Decision Processes with General State Space

Participant: François Dufour.

In this work, we deal with a discrete-time finite horizon Markov decision process with locally compact Borel state and action spaces, and possibly unbounded cost function. Based on Lipschitz continuity of the elements of the control model, we propose a state and action discretization procedure for approximating the optimal value function and an optimal policy of the original control model. We provide explicit bounds on the approximation errors. Our results are illustrated by a numerical application to a fisheries management problem.

These results have been obtained in collaboration with Tomas Prieto-Rumeau, Department of Statistics and Operations Research, UNED, Madrid, Spain. It has been accepted for publication in Stochastics [13].

6.3. Asymmetry tests for Bifurcating Auto-Regressive Processes with missing data

Participants: Benoîte de Saporta, Anne Gégout-Petit.

Using the properties of the estimators studied in [20], we have constructed symmetry tests for bifurcating autoregressive processes (BAR) when some data are missing. BAR processes typically model cell division data. Each cell can be of one of two types odd or even. The goal of this work is to test asymmetry between odd and even cells in a single observed lineage. We have also derived asymmetry tests for the lineage itself, modeled by a two-type Galton-Watson process observed in a non standard scheme. We present applications on both simulated and real data.

This work is in collaboration with Laurence Marsalle of Lille 1 university. It will be soon submitted for publication.

6.4. Statistical study of asymmetry in cell lineage data

Participants: Benoîte de Saporta, Anne Gégout-Petit.
Simulation studies of the asymmetry in a single lineage data have shown a lack of power of the tests when the number of available generation is limited. This work proposes a rigorous methodology to study cell division data in the context of observation of several lineages. It generalizes [20]. We model the data by an asymmetric bifurcating autoregressive (BAR) process and take into account possibly missing data by modeling the genealogies with a two-type Galton Watson (GW) process. Our inference is based on several lineages, i.e. independent and identically distributed replicas of the coupled BAR and GW processes. We propose a least-squares estimator of the unknown parameters of the BAR process and an estimator of the parameters of the GW process, study their asymptotic properties and propose symmetry tests. Our results are applied on real data of Escherichia coli division.

This work is in collaboration with Laurence Marsalle of Lille 1 university. It will be soon submitted for publication.

6.5. Estimation of the jump rate of a PDMP

Participants: Romain Azaïs, François Dufour, Anne Gégout-Petit.

We estimate the jump rate of PDMP. We suppose the flow given by physics laws and we want to make some inference on $\lambda$, $\phi$ being deterministic, the problem can be rewritten as a problem of estimation of the rate $\lambda(z, t)$ with $z \in E$ with $E$ an open set of a separable metric space. We have an ergodicity assumption on the observed PDMP and the asymptotic is in the time of observation of the process.

We distinguish three cases:

1. $E$ is finite. In this case, we easily estimate each of the cumulated risk functions $\Lambda(z, t) = \exp(- \int_0^t \lambda(z, s) \, ds)$ corresponding to each of $z \in E$ by a Nelson Aalen estimator. The results is based on the decomposition in semi-martingale of the following counting process in an appropriate filtration:

$$\forall t \geq 0, \ N_n(z, t) = \sum_{i=0}^{n-1} 1\{S_{i+1} \leq t\} 1\{Z_i = z\}, \quad (1)$$

We obtain the estimator of the rate $\lambda(z, t)$ by smoothing of the estimator of $\Lambda$.

2. $E$ is an open set of a general separable metric space but the transition measure $Q$ does not depend on the time spent in the current regime. In this case, we suppose the rate $\lambda(z, t)$ Lipschitz and the process ergodic with a stationary law denoted by $\nu$. We first construct an estimation of the cumulated rate knowing that $z$ belongs to a set $A$ such that $\nu(A) > 0$ by:

$$\hat{L}_n(A, t) = \sum_{i=0}^{n-1} \frac{1}{Y_n(A, S_{i+1})} 1\{S_{i+1} \leq t\} 1\{Z_i \in A\}, \quad \text{with} \quad Y_n(A, t) = \sum_{i=0}^{n-1} 1\{S_{i+1} \geq t\} 1\{Z_i \in A\}. \quad (2)$$

We show the consistence of the estimator. Smoothing $\hat{L}_n(A, t)$ and using a fine partition of $E$ allow us to obtain an uniform result for the approximation of the rate $\lambda(z, t)$, in some sense in $t$ and $z$.

3. $E$ is an open set of a general separable metric space and the transition measure $Q$ depends on the time spent in the current regime. Here, we loose some conditional independence between the $S'_i$’s and the whole set of the locations of the jump $\{Z_1, ..., Z_n\}$. We have to make a detour for the estimation of the law of the time $S_{k+1}$ knowing the current $Z_k$ by the the law $S_{k+1}$ knowing $(Z_k, Z_{k+1})$. The method gives an estimation of the conditional density of $S_{k+1}$ given $Z_k$.

We have made simulation studies that give expected results. A R package for this estimation method is in progress.
This work is a part of the PhD Thesis of R. Azaïs founded by the ANR Fautocoes. R. Azaïs has presented a part of this work at “Rencontres des Jeunes Statisticiens” in 2011 September [28]. The work will be soon submitted to a international peer-reviewed journal for publication.

6.6. Detection of a damaged operating mode of optronic equipment using Hidden Markov Model

Participants: Camille Baysse, Anne Gégout-Petit, Jérôme Saracco.

As part of optimisation of the reliability, Thales Optronics now includes systems that examine the state of its equipment. This function is performed by HUMS (Health & Usage Monitoring System). The aim is to implement a program based on these observations that can determine the lifetime of this optronic equipment and optimize its maintenance.

Our study focuses on a simple example of HUMS. As part of our research, we are interested in a variable called "time-to cold" noted TMF, which reflects the state of the system. Using these informations about this variable, we seek to detect as soon as possible a damaged state and propose a maintenance before failure. This would allow the Thales Optronics company to improve its maintenance system and achieve many economies.

For this we use a hidden Markov model. The state of our system at time t is then modeled by a continuous time Markov chain X(t) with three states: stable, damaged and failure. However we do not observe directly this chain but indirectly through the TMF, a noisy function of this chain. Thanks to filtering equations, we obtained results on the probability that an equipment is in a damaged state at time t, knowing the history of the TMF until this moment. We have subsequently studied the method on simulated data, before applying these results on the analysis of our real data and we have checked that the results are consistent with the reality. This work will be used by Thales Optronique for the optimization of the maintenance.

This work is a part of the CIFRE PhD Thesis of Camille Baysse founded by Thales Optronic. It is the object of a technical report [48] and was presented in [29] in an internal Thales seminar.

6.7. Optimal quantization applied to Sliced Inverse Regression

Participants: Romain Azaïs, François Dufour, Anne Gégout-Petit, Jérôme Saracco.

We tackle the well known Slice Inverse Regression (SIR) method for a semiparametric regression model involving a quantitative variable $X$ and including a dimension reduction of $X$ via a parameter $\beta$. The response variable $Y$ is real. Our goal is to estimate $\beta$ and to predict the response variable conditionally to $X$. We adapt SIR method using optimal quantization [57] in the first time only for the independent variable $X$ for the estimation of $\beta$. In a second time, we quantize the variable $(\hat{\beta}_n, Y)$ in order to propose a discrete conditional law of $Y$ given $X = x$. We show the convergence of the estimator of $\beta$ and of the conditional law. Simulation studies show the numerical qualities of our estimates. This work is the object of a publication in Journal of Statistical Planning and Inference [15] and was presented in a national conference [23].

6.8. Multivariate Analysis for the detection of the effect of a treatment

Participant: Anne Gégout-Petit.

The aim of this work is to give some statistical rules to determine if a patient is meeting a given treatment (a BD here). The criterium commonly used to determine if a patient is meeting a BD treatment is based only on one physiological parameter: if this parameter increases, the patient is meeting. But now, many physiological parameters are measured in routine and physiologists "ont le sentiment" that a patient could have a global amelioration of his health state due to the treatment without an increase of the single used parameter.
Using the measures of six variables before and after the treatment, the expected value of this variable under the hypothesis of good health, we first propose indices of amelioration. Using standard multivariate analysis techniques, we first study the correlation between these indices. We use classification in order to constitute groups of patients whose present homogeneous treatment responses. The method used on a cohort of 100 subjects gives three groups: in the first one, the mean of the indices is near zero, the treatment has no effect. In the second one, subjects present significant amelioration regarding two of the indices but not the indices related to the commonly used one. The last group show an amelioration for all the indices.

We have proposed criteria to discriminate the three groups. These criteria where used on a validation cohort to constitute three groups. Their global characteristics were the same as in the original cohort and it comforts the validity of the method. This work will be used by physiologists to propose new criteria for the measure of the effect of a BD treatment. It is in collaboration with physiologists from Bordeaux and Nantes universities and is the object of a paper that will be soon submitted in a international peer-reviewed journal in the domain of pneumology.

6.9. Numerical computation of expectations of PDMP’s

Participants: Adrien Brandejsky, Benoîte de Saporta, François Dufour.

This work concerns the computation of expectations of functionals of piecewise deterministic Markov processes (PDMP’s). We propose a numerical scheme to approximate such expectations, analyze the convergence of our scheme and derive bounds for the convergence rate. More precisely, we are interested in the approximation of expectations of the form

$$E_x \left[ \int_0^{T_{N_0}} l(X_t) \, dt + \sum_{j=1}^{N} c(X_{T_j^-}) I_{(X_{T_j^-} \in \partial E)} \right]$$

where $(X_t)_{t \geq 0}$ is a PDMP, $(T_j)$ are its jump times and $l$ and $c$ are some non negative, real-valued, bounded functions. Such expectations are discussed by M.H.A. Davis in [50], chapter 3. They often appear as cost or reward functions in optimization problems. The first term is referred to as the running cost while the second may be called the boundary jump cost. Besides, they are quite general since M.H.A. Davis shows how a wide variety of apparently different functionals can be obtained from the above specific form. For example, this wide variety includes quantities such as a mean exit time and even, for any fixed $t \geq 0$, the distribution of $X_t$ (i.e. $E_x[I_F(X_t)]$ where $F$ is a measurable set).

There are surprisingly few works in the literature devoted to the actual computation of such expectations, using other means than direct Monte Carlo simulations. M.H.A Davis showed that these expectations satisfy integro-differential equations. However, the set of partial differential equations that is obtained is unusual. Roughly speaking, these differential equations are basically transport equations with a non-constant velocity and they are coupled by the boundary conditions and by some integral terms involving kernels that are derived from the properties of the underlying stochastic process. This approach is currently under study in this project by LATP. The main difficulty comes from the fact that the domains on which the equations have to be solved vary from one equation to another making their numerical resolution highly problem specific. Another similar approach has been recently investigated in [49], [54]. It is based on a discretization of the Chapman Kolmogorov equations satisfied by the distribution of the process $(X_t)_{t \geq 0}$. The authors propose an approximation of such expectations based on finite volume methods. Unfortunately, their method is only valid if there are no jumps at the boundary.

Our approach is completely different and does not rely on differential equations, but on the fact that such expectations can be computed by iterating an integral operator $G$. This operator only involves the embedded Markov chain $(Z_n, S_n)_{n \in \mathbb{N}}$ and conditional expectations. It is therefore natural to propose a computational method based on the quantization of this Markov chain, following the same idea as [8]. We also addressed two important aspects that had not been investigated in [8]. The first one consists in allowing $c$ and $l$ to be time depending functions, although still Lipschitz continuous, so that we may compute expectations of the form
\[ E_x \left[ \int_0^{T_N} l(X_t, t) dt + \sum_{j=1}^N c(X_{T_j^-}, T_j) I_{\{X_{T_j^-} \in \partial E\}} \right] . \]

This important generalization has huge applicative consequences. For instance, it allows discounted cost or reward functions such as \( l(x, t) = e^{-\delta t} l(x) \) and \( c(x, t) = e^{-\delta t} c(x) \) where \( \delta \) is some interest rate. To compute the above expectation, our strategy consists in considering, as it is suggested by M.H.A. Davis in [50], the time augmented process \( \tilde{X}_t = (X_t, t) \).

The second important generalization is to consider the deterministic time horizon problem. Indeed, it seems crucial, regarding to the applications, to be able to approximate

\[
E_x \left[ \int_0^t l(X_t, t) dt + \sum_{j \leq T_j} c(X_{T_j^-}, T_j) I_{\{X_{T_j^-} \in \partial E\}} I_{\{T_j \leq t\}} \right] = E_x \left[ \int_0^{+\infty} l(X_t, t) I_{\{t \leq T_f\}} dt + \sum_{j=1}^{+\infty} c(X_{T_j^-}, T_j) I_{\{X_{T_j^-} \in \partial E\}} I_{\{T_j \leq t\}} \right]
\]

for some fixed \( t_f > 0 \) regardless of how many jumps occur before this deterministic time. To compute this quantity, we start by choosing a time \( N \) such that \( P(T_N < t_f) \) be small so that the previous expectation boils down to

\[ E_x \left[ \int_0^{T_N} l(X_t, t) dt + \sum_{j=1}^N c(X_{T_j^-}, T_j) I_{\{X_{T_j^-} \in \partial E\}} I_{\{T_j \leq t_f\}} \right] . \]

We deal with the two indicator functions in two different ways. On the one hand, we prove that it is possible to relax the regularity condition on the running cost function so that our algorithm still converges in spite of the first indicator function. On the other hand, since the same reasoning cannot be applied to the indicator function within the boundary jump cost term, we bound it between two Lipschitz continuous functions. This provides bounds for the expectation of the deterministic time horizon functional.

An important advantage of our method is that it is flexible. Indeed, as pointed out in [47], a quantization based method is obstacle free which means, in our case, that it produces, once and for all, a discretization of the process independently of the functions \( l \) and \( c \) since the quantization grids merely depend on the dynamics of the process. They are only computed once, stored off-line and may therefore serve many purposes. Once they have been obtained, we are able to approximate very easily and quickly any of the expectations described earlier. This flexibility is definitely an important advantage of our scheme over standard methods such as Monte-Carlo simulations since, with such methods, we would have to run the whole algorithm for each expectation we want to compute.

The theoretical part of this work with rigorous proofs is under review for an international peer-reviewed journal [10]. F. Dufour presented this work in an invited session at an international conference [22].

### 6.10. Optimal stopping under partial observation

**Participants:** Adrien Brandejsky, Benoîte de Saporta, François Dufour.

In continuation of our work on optimal stopping for Piecewise deterministic Markov processes (PDMP’s) [8], we have started investigating the optimal stopping problem when the process is only partially observed. We supposed that the jump times of the process are observed, but the post jump locations are observed through a noise.
The first step is to rewrite the optimal stopping problem for the partially observed PDMP as a totally observed stopping problem for a new Markov chain, obtained by filtering the observation process. Then, one has to study precisely this filter, which is non standard due to the possible jumps of the process. The next step is to derive the dynamic programming equation adapted to our framework. Finally, we propose a numerical method based on quantization to approximate the value function and $\epsilon$-stopping times. Track is also kept of the error bounds all through our numerical procedure.

This work is still in progress and should be submitted to an international peer-reviewed journal shortly.

6.11. Efficient simulation of the availability of a feedwater control system

Participants: Benoîte de Saporta, François Dufour, Huilong Zhang.

In the reliability modeling of complex control systems, classical methodologies such as even-trees/fault-trees or Petri nets may not represent adequately the dynamic interactions existing between the physical processes (modeled by continuous variables) and the functional and dysfunctional behavior of its components (modeled by discrete variables). This paper proposes a framework for modeling and simulation of a water level control system in the steam generator (SG) in the secondary circuit of a nuclear power plant. A similar benchmark system was described by the U.S. Nuclear Regulatory Commission [46] to compare two approaches to dynamic reliability: DFM (Dynamic Flowgraph Methodology) and Markov/CCMT (Cell-to-Cell Mapping Technique). But the report released by the NRC is not sufficient to reconstruct a realistic model. We have developed a complete benchmark case. The behavioral model of SG is obtained from a linearized model published in 2000 by EDF [55]. Detailed description of the components, failure modes and control laws of the principal components is presented. For modeling the system, we use the piecewise deterministic Markov processes (PDP) framework [50] and for implementation we chose Simulink associated with Stateflow. PDP’s offer a very general modeling framework to deal with dynamic reliability problems; Simulink is a good tool to simulate non linear differential equations and their controller, while Stateflow implementation is appropriate for finite state machine descriptions of different components.

In our benchmark system, four physical processes are considered: feedwater flowrate, steam flow, narrow range water level and wide range water level. A PID controller is used to maintain the water level within limits of set-points. The system is composed of seven components: 1 passive system representing vapor transport system, 3 extraction pumps, 2 feeding turbopumps, and 1 waterflow regulation valve. The functional and dysfunctional behaviors and the failure rates of each component are based on operational experience.

We show that PDP modeling is suitable for dynamic reliability analysis, and that Simulink associated with Stateflow provides an interactive simulator (graphical block diagramming), which makes the simulator scalable. This work is submitted for presentation to an international conference in 2012.

6.12. Orthogonal Rotation in PCAMIX

Participants: Marie Chavent, Jéréôme Saracco.

The aim of this work is to propose an efficient algorithm for rotation in PCAMIX, a principal component method for a mixture of qualitative and quantitative variables. We give a new presentation of PCAMIX where the principal components and the squared loadings are obtained from a Singular Value Decomposition. The loadings of the quantitative variables and the principal coordinates of the categories of the qualitative variables are also obtained directly. In this context, we propose a computationaly efficient procedure for varimax rotation in PCAMIX and a direct solution for the optimal angle of rotation. A simulation study shows the good computational behavior of the proposed algorithm. An application on a real data set illustrates the interest of using rotation in MCA. All source codes are available in the R package “PCAmixdata”. This work is in revision for publication [45] and has been presented in [36].


Participants: Marie Chavent, Jéréôme Saracco.
In this work, we consider block-wise evolving data streams. When a semi-parametric regression model involving a common dimension reduction direction $\beta$ is assumed for each block, we propose an adaptive SIR (for sliced inverse regression) estimator of $\beta$. This estimator is faster than usual SIR applied to the union of all the blocks, both from computational complexity and running time points of view. We show the consistency of our estimator at the root-$n$ rate and its asymptotic normality. We also propose an extension of this method to multiple indices model. In simulation studies, we illustrate the good numerical behavior of our estimator. We also provide a graphical tool in order to detect if there exists a drift of the dimension reduction direction or some aberrant blocks of data. We illustrate our approach with various scenarios. We apply this approach on the following real data problem.

As an illustration, we consider a nonlinear inverse problem in remote sensing. The goal is to estimate the physical properties of surface materials on the planet Mars from hyperspectral data. The method is based on the estimation of the functional relationship between some physical parameters $Y$ and observed spectra $X$. For this purpose, a database of synthetic spectra is generated by a physical radiative transfer model. We propose to reduce the high dimension of spectra ($p = 352$ wavelengths) with a regularized version of SIR. The need to regularize SIR in very high dimensions is well-known. In practice, the database of synthetic spectra may be so large that it cannot be stored in a computer memory. Thus, a stream of smaller sub-databases is generated and we apply our “SIR datastream” approach to this context.

This work will be submitted for publication very soon and it has been presented in the international conference [31].

### 6.14. Classification of EEG data by evolutionary algorithm for the study of vigilance states

**Participants:** Marie Chavent, Laurent Vézard.

The objective of this work [42] is to predict the state of vigilance of an individual from the study of its brain activity (EEG signals). The variable to predict is binary (alertness "normal" or "relaxed"). EEG of 44 participants in both states (88 records) were collected with a helmet with 58 electrodes. After a pretreatment step and data validation, a test called "test slope" was chosen. The usual methods of supervised classification (k nearest neighbors, binary classification trees, random forests, and discriminant sparse PLS) were used to provide predictions of the state of participants. The test was then refined using a genetic algorithm, which has built a reliable model (average true classification rate by using CART equal to 86.68 ± 1.87%) and to select an electrode from the initial 58. This work is in collaboration with Pierrick Legrand (EPI Alea) and Frédérique Faïta (EA 487 "Sciences cognitives et facteur humain”

### 6.15. Comparison of Kernel Density Estimators on Environmental Data with Assumption on Number of Modes

**Participants:** Jérôme Saracco, Raphaël Coudret.

In this work, we consider valvometric data samples, that is measurements of distances between the two parts of the shell of bivalves. The movements of a few oysters are monitored in different places (like Arcachon Bay or Santander port) by a laboratory called Environnements et Paléoenvironnements Océaniques et Continentaux (EPOC). The aim of these experiments is to determine water quality via the observation of the bivalves behavior. Previous related works have been published on this subject. EPOC team studied this animals in a copper pollution context via investigations using kernels methods. In our study, we consider each day the density of the distance between the two parts of the shell. This density is reasonably assumed to have 2 modes, the first one corresponds to a close status of the shell and the second one to an open status. The study of the evolution of this density along the time provides informations on bivalves behavior. We provide theoretical results on our bandwidth choice with a bounded support kernel and we give a procedure to determine this bandwidth. We also provide asymptotic results for a density kernel estimator with a kernel that has it support on $\mathbb{R}$. We present an extensive simulation study in order to compare numerical performances of various density
estimators based or not on the two modes assumption. From the results obtained from the simulated data, we derive the suitable estimator for our real data application. This work will be submitted for publication very soon and it has been presented in the international conference [37].
6. New Results

6.1. Sampling methods for inverse scattering problems

6.1.1. Sampling methods with time dependent data
Participants: Houssem Haddar, Armin Lechleiter, Simon Marmorat.

We considered the extension of the so-called Factorization method to far-field data in the time domain. For a Dirichlet scattering object and incident wave fronts, the inverse problem under investigation consists in characterizing the shape of the scattering object from the behaviour of the scattered field far from the obstacle (far-field measurements). We derive a self-adjoint factorization of the time-domain far-field operator and show that the middle operator of this factorization possesses a weak type of coercivity. This allows to prove range inclusions between the far-field operator and the time-domain Herglotz operator.

We also extended the near-field version of the linear sampling method to causal time-dependent wave data for smooth, band-limited incident pulses, considering different boundary conditions as for instance Dirichlet, Neumann or Robin conditions [27].

6.1.2. Inverse problems for periodic penetrable media
Participants: Armin Lechleiter, Dinh Liem Nguyen.

Imaging periodic penetrable scattering objects is of interest for non-destructive testing of photonic devices. The problem is motivated by the decreasing size of periodic structures in photonic devices, together with an increasing demand in fast non-destructive testing. In this project, linked to the thesis project of Dinh Liem Nguyen, we considered the problem of imaging a periodic penetrable structure from measurements of scattered electromagnetic waves. As a continuation of earlier work, we considered an electromagnetic problem for transverse magnetic waves (earlier work treats transverse electric fields), and also the full Maxwell equations. In both cases, we treat the direct problem by a volumetric integral equation approach and construct a Factorization method.

6.1.3. Inverse problems for Stokes-Brinkmann flows
Participants: Armin Lechleiter, Tobias Rienmüller.

Geometric inverse problems for flows arise for instance when controlling pipelines and oil reservoirs. In this project, we considered the Stokes-Brinkmann equations that model, for instance, porous penetrable inclusions in a free background. The factorization method is able to characterize the inclusions from the relative Dirichlet-to-Neumann operator. Numerical examples show the feasibility of the method.

6.1.4. Inverse scattering from screens with impedance boundary conditions
Participants: Yosra Boukari, Houssem Haddar.

We are interested in solving the inverse problem of determining a screen (or a crack) from multi-static measurements of electromagnetic (or acoustic) scattered field at a given frequency. An impedance boundary condition is assumed to be verified at both faces of the screen. We extended the so-called factorization method to this setting. We also analyzed a data completion algorithm based on integral equation method for the Helmholtz equation. This algorithm is then coupled to the so-called RG-LSM algorithm to retrieve cracks inside a locally homogeneous background. This work is conducted in collaboration with F. Ben Hassen.

6.1.5. Transmission Eigenvalues and their application to the identification problem
Participants: Anne Cossonnière, Houssem Haddar, Giovanni Giorgi.
The so-called interior transmission problem plays an important role in the study of inverse scattering problems from (anisotropic) inhomogeneities. Solutions to this problem associated with singular sources can be used for instance to establish uniqueness for the imaging of anisotropic inclusions from multi-static data at a fixed frequency. It is also well known that the injectivity of the far field operator used in sampling methods is equivalent to the uniqueness of solutions to this problem. The frequencies for which this uniqueness fails are called transmission eigenvalues. We are currently developing approaches where these frequencies can be used in identifying (qualitative informations on) the medium properties. Our research on this topic is mainly done in the framework of the associate team ISIP http://www-direction.inria.fr/international/PHP/Networks/LiEA.php with the University of Delaware. A review article on the state of art concerning the transmission eigenvalue problem has been written in collaboration with F. Cakoni [24].

The main topic of the PhD thesis of A. Cossonnière is to extend some of the results obtained above (for the scalar problem) to the Maxwell’s problem. In this perspective, theoretical results related to solutions of the interior transmission problem for medium with cavities and existence of transmission eigenvalues have been obtained [14]. This work is then extended to the case of medium with perfectly conducting inclusions. Only the scalar case has been studied [35]. In collaboration with M. Fares and F. Collino from CERFACS we investigated the use of a surface integral equation approach to find the transmission eigenvalues for inclusions with piecewise constant index. The main difficulty behind this procedure is the compactness of the obtained integral operator in usual Sobolev spaces associated with the forward scattering problem. We solved this difficulty by introducing a preconditioning operator associated with a “coercive” transmission problem. The obtained procedure has been validated numerically in 2D and 3D cases. We also analyzed the transmission eigenvalue problem using this surface integral equation approach. This technique allowed us to generalize discretness results on the spectrum to cases where the contrast can change sign [2].

With G. Giorgi, we developed a method that give estimates on the material properties using the first transmission eigenvalue. This method is based on reformulating the interior transmission eigenvalue problem into an eigenvalue problem for the material coefficients. We validated our methodology for homogeneous and inhomogeneous inclusions and backgrounds. We also treated the case of a background with absorption and the case of scatterers with multiple connected components of different refractive indexes [26].

6.1.6. The factorization method for EIT with inhomogeneous background

Participants: Giovanni Migliorati, Houssem Haddar.

We developed a numerical inversion scheme based on the Factorization Method to solve the (continuous model of) Electrical Impedance Tomography problem with inhomogeneous background. The numerical scheme relies on the well chosen approximation by the finite element method of the solution to the dipole-like Neumann boundary-value problem. Two regularization techniques are tested, i.e. the Tikhonov regularization embedding Morozov principle, and the classical Picard Criterion. The numerical analysis of the method and the results obtained are presented in the INRIA report [28].

6.2. Iterative Methods for Non-linear Inverse Problems

6.2.1. Inverse medium problem for axisymmetric eddy current models

Participants: Houssem Haddar, Zixian Jiang, Armin Lechleiter.

We are interested in shape optimization methods for inclusion detection in an axisymmetric eddy current model. This problem is motivated by non-destructive testing methodologies for steam generators. We investigated the validity of the eddy current model for these kinds of problems and developed numerical methods for the solution of the direct problem in weighted Sobolev spaces. Then we computed the shape derivative of an inclusion which allows to use regularized iterative methods to solve the inverse problem [23]. We also develop asymptotic models to identify thin highly conducting deposits.

6.2.2. Hybrid methods for inverse scattering problems

Participants: Grégoire Allaire, Houssem Haddar, Dimitri Nicolas.
It is well admitted that optimization methods offer in general a good accuracy but are penalized by the cost of solving the direct problem and by requiring a large number of iterations due to the ill-posedness of the inverse problem. However, profiting from good initial guess provided by sampling methods these method would become viable. Among optimization methods, the Level Set method seems to be well suited for such coupling since it is based on capturing the support of the inclusion through an indicator function computed on a cartesian grid of probed media. Beyond the choice of an optimization method, our goal would be to develop coupling strategies that uses sampling methods not only as an initialization step but also as a method to optimize the choice of the incident (focusing) wave that serves in computing the increment step.

We investigated a coupling approach between the level set method and LSM where the initialization is done using a crude estimate provided by the linear sampling method. The obtained results validate the efficiency of this coupling in the case of simply and multiply connected obstacles that are well separated.

6.3. Shape and topology optimization

6.3.1. Incorporating manufacturing constraints in topology optimization

**Participant:** Grégoire Allaire.

With G. Michailidis and F. Jouve we study how to incorporate manufacturing constraints in topology optimization of structures using the level set method. The goal is to obtain a structure with optimal mechanical behaviour, which at the same time respects some predefined constraints imposed by the fabrication process. In this way, the final optimal shape is manufacturable and thus the method of shape and topology optimization turns to be industrially applicable.

The first constraints we have tackled are related to the limits of thickness a structure is forced to respect. We need to avoid optimal shape that contain very thin or thick members or even members that are very close between them. To achieve this, we have adopted two different approaches, a geometrical and a mechanical one. In the geometrical one, we have made extensive use of the notion of the signed-distance function to a domain. We have formulated a global constraint which guarantees that, at the end of the optimization process, the optimal structure respects the thickness limits. In the mechanical approach, we have tried to simulate the solidification process of a structure constructed via casting. We have set a time contraint, i.e. we have required that the structure cools earlier than some predefined time limit. We have started working on a more complicated thermal equation, a non-linear model with phase change, in order to describe more accurately the solidification process.

6.3.2. Optimization of composite materials draping

**Participant:** Grégoire Allaire.

With G. Delgado we work on the optimization of composite materials draping. These composite structures are constructed by lamination of a sequence of unidirectional reinforced layers or plies. Each ply is typically a thin sheet of carbon fibers impregnated with polymer matrix material. The optimization variables are the geometries of these layers and they are parameterized by a level set function. In a first instance, we treat the problem of mass minimization (with a constraint on the maximal compliance) for a cantilever type composite structure, laminated with four layers of a given orthotropic material at different angles. The elasticity analysis is performed with the software Freefem++, coupled with a C++ routine to solve, by a finite difference scheme, the evolution of the level sets.

6.3.3. A hybrid optimization method

**Participant:** Grégoire Allaire.
With Ch. Dapogny and P. Frey we develop a new method of geometric optimization for structures that relies on two alternative descriptions of shapes: on the one hand, they are exactly meshed so that mechanical evaluations by finite elements are accurate; on the other hand, we resort to a level-set characterization to describe their deformation along the shape gradient. The key ingredient is a meshing algorithm for building a mesh, suitable for numerical computations, out of a piecewise linear level-set function on an unstructured mesh. Therefore, our approach is at the same time a geometric optimization method (since shapes are exactly meshed) and a topology optimization method (since the topology of successive shapes can change thanks to the power of the level-set method). Our first results in 2-d have been announced. We continue to work on the 3-d case.

6.3.4. DeHomogenization

Participant: Olivier Pantz.

In most shape optimization problems, the optimal solution does not belong to the set of genuine shapes but is a composite structure. The homogenization method consists in relaxing the original problem thereby extending the set of admissible structures to composite shapes. From the numerical viewpoint, an important asset of the homogenization method with respect to traditional geometrical optimization is that the computed optimal shape is quite independent from the initial guess (even if only a partial relaxation is performed). Nevertheless, the optimal shape being a composite, a post-treatment is needed in order to produce an almost optimal non-composite (i.e. workable) shape. The classical approach consists in penalizing the intermediate densities of material, but the obtained result deeply depends on the underlying mesh used and the details level is not controllable. We proposed in [40] a new post-treatment method for the compliance minimization problem of an elastic structure. The main idea is to approximate the optimal composite shape with a locally periodic composite and to build a sequence of genuine shapes converging toward this composite structure. This method allows us to balance the level of details of the final shape and its optimality. Nevertheless, it was restricted to particular optimal shapes, depending on the topological structure of the lattice describing the arrangement of the holes of the composite. We lifted this restriction in order to extend our method to any optimal composite structure for the compliance minimization problem in [39]. Since, the method has been improved and a new article presenting the last results is in preparation. Moreover, we intend to extend this approach to other kinds of cost functions. A first attempt, based on a gradient method, has been made. Unfortunately, it was leading to local minima. Thus a new strategy has to be worked out. It will be mainly based on the same ideas than the one developed for the compliance minimization problem, but some difficulties are still to be overcome.

6.3.5. Level-Set Method

Participant: Olivier Pantz.

We have begun to work, with Gabriel Delagado, on a new level-set optimization method, based on a gradient method. The key idea consists in computing directly the derivative of the discretized cost functions. The main advantage is that it is usually more simple to implement than the standard approach (consisting in using a discretized version of the gradient of the cost function). Moreover, the results obtained are as good or even better than the one obtained in previous works. Nevertheless, this method has its drawbacks, since the cost function is only derivable almost everywhere (the zero level-set has to be transverse to the triangulation of the mesh). It follows that convergence toward the minimum by the gradient method is not granted. To overcome this problem, we intend to use a mix-formulation for the state function. An article is in preparation on this subject.

6.3.6. Robust Optimization

Participant: Olivier Pantz.

One of the main problem in shape optimization problems is due to the fact that the gradient is never computed exactly. When the current solution is far from a local optimum, this is not a problem: even a rough approximation of the gradient enable us to exhibit a descent direction. On the contrary, when close to a local optimal, a very precise computation of the gradient is needed. We intend, with G. Delgado, to use a-posteriori error estimates evaluate the errors made on the computation of the gradient and to ensure that at each step, a genuine descent direction is used in the gradient method.
6.3.7. **Optimization of a sodium fast reactor core**  
**Participants:** Grégoire Allaire, Olivier Pantz.

In collaboration with D. Schmidt, G. Allaire and E. Dombre, we apply the geometrical shape optimization method for the design of a SFR (Sodium Fast reactor) core in order to minimize a thermal counter-reaction known as the sodium void effect. In this kind of reactor, by increasing the temperature, the core may become liable to a strong increase of reactivity $\rho$, a key-parameter governing the chain-reaction at quasi-static states. We first use the 1 group energy diffusion model and give the generalization to the 2 groups energy equation. We then give some numerical results in the case of the 1 group energy equation. Note that the application of our method leads to some designs whose interfaces can be parametrized by very smooth curves which can stand very far from realistic designs. We don’t explain here the method that it would be possible to use for recovering an operational design but there exists several penalization methods that could be employed to this end. This work was partially sponsored by EDF. Our results will be published in the proceedings of the CEMRACS’11, during which part of the results have been obtained.

6.4. **Asymptotic models**

6.4.1. **Inverse scattering problem for coated obstacles**  
**Participants:** Nicolas Chaulet, Houssem Haddar.

In collaboration with L. Bourgeois, we considered the inverse scattering problem consisting in the identification of both an obstacle and its “equivalent impedance” from farfield measurements at a fixed frequency. The first specificity of this work is to consider the cases where this impedance is not a scalar function but a second order surface operator. The latter can be seen as a more general model for effective impedances and is for instance widely used for scattering from thin coatings. The second specificity of this work is to characterize the derivative of a least square cost functional with respect to this complex configuration. We provide in particular an extension of the notion of shape derivative to the cases where the impedance parameters cannot be considered as the traces of given functions. For instance, the obtained derivative does not vanish (in general) for tangential perturbations. The efficiency of considering this type of derivative is illustrated by some 2D numerical experiments based on a (classical) steepest descent method. The feasibility of retrieving both the obstacle and the impedance functionals is discussed in further numerical experiments [33].

6.4.2. **Interface conditions for thin dielectrics**  
**Participant:** Houssem Haddar.

Jointly with B. Delourme and P. Joly we established transmission conditions modelling thin interfaces that has (periodic) rapid variations along tangential coordinates. Motivated by non destructive testing experiments, we considered the case of cylindrical geometries and time harmonic waves. We already obtained a full asymptotic description of the solution in terms of the thickness in the scalar case using so called matched asymptotic expansions. This asymptotic expansion is then used to derive generalized interface conditions and establish error estimates for obtained approximate models [15]. The analysis of the approximate problem for Maxwell’s equations is the subject of a forthcoming publication.

6.4.3. **Homogenization**  
**Participant:** Grégoire Allaire.

With I. Pankratova and A. Piatnitski we considered the homogenization of a non-stationary convection-diffusion equation posed in a bounded periodic heterogeneous domain with homogeneous Dirichlet boundary conditions. Assuming that the convection term is large, we give the asymptotic profile of the solution and determine its rate of decay. In particular, it allows us to characterize the “hot spot”, i.e., the precise asymptotic location of the solution maximum which lies close to the domain boundary and is also the point of concentration. Due to the competition between convection and diffusion, the position of the “hot spot” is not always intuitive as exemplified in some numerical tests.
With Z. Habibi we studied the homogenization of heat transfer in periodic porous media where the fluid part is made of long thin parallel cylinders, the diameter of which is of the same order than the period. The heat is transported by conduction in the solid part of the domain and by conduction, convection and radiative transfer in the fluid part (the cylinders). A non-local boundary condition models the radiative heat transfer on the cylinder walls. To obtain the homogenized problem we first use a formal two-scale asymptotic expansion method. The resulting effective model is a convection-diffusion equation posed in a homogeneous domain with homogenized coefficients evaluated by solving so-called cell problems where radiative transfer is taken into account. In a second step we rigorously justify the homogenization process by using the notion of two-scale convergence. One feature of this work is that it combines homogenization with a 3D to 2D asymptotic analysis since the radiative transfer in the limit cell problem is purely two-dimensional. Eventually, we provide some 3D numerical results in order to show the convergence and the advantages of our homogenization method.

6.4.4. Modelling and simulation for underground nuclear waste storage.
Participants: Grégoire Allaire, Harsha Hutridurga.

In the framework of the GDR MOMAS (Groupement de Recherches du CNRS sur les MOdélisations MAthématiques et Simulations numériques liées aux problèmes de gestion des déchets nucléaires) I am working with R. Brizzi, H. Hutridurga, A. Mikelic and A. Piatnitski on upscaling of microscopic models by homogenization (i.e. finding macroscopic models and effective coefficients).

We studied the Taylor dispersion of a contaminant in a porous medium. The originality of the model is that it takes into account surface diffusion and convection on the pores boundaries. We rigorously obtained the homogenized equation and studied the behavior of the effective dispersion tensor when varying various parameters.

In collaboration with a team of chemists (around J.-F. Dufrêche from the GNR Paris), we have undertaken the rigorous homogenization of a system of PDEs describing the transport of a N-component electrolyte in a dilute Newtonian solvent through a rigid porous medium. The motion is governed by a small static electric field and a small hydrodynamic force, which allowed us to use O’Brien’s linearized equations as the starting model. Convergence of the homogenization procedure was established and the homogenized equations were discussed. Based on the rigorous study of the underlying equations, it was proved that the effective tensor satisfies Onsager properties, namely is symmetric positive definite. This result justified the approach of many authors who used Onsager theory as a starting point.

6.4.5. A new membrane/plate modeling
Participant: Olivier Pantz.

Using a formal asymptotic expansion, we have proved with K. Trabelsi, that non-isotropic thin-structure could behave (when the thickness is small) like a shell combining both membrane and bending effects. It is the first time to our knowledge that such a model is derived. An article on this project is in preparation.

6.4.6. A new Liouville type Rigidity Theorem
Participant: Olivier Pantz.

We have recently developed a new Liouville type Rigidity Theorem. Considering a cylindrical shaped solid, we prove that if the local area of the cross sections is preserved together with the length of the fibers, then the deformation is a combination of a planar deformation and a rigid motion. The results currently obtained are limited to regular deformations and we are currently working with B. Merlet to extend them. Nevertheless, we mainly focus on the case where the conditions imposed to the local area of the cross sections and the length of the fibers are only "almost" fulfilled. This will enable us to derive rigorously new non linear shell models combining both membranar and flexural effects that we have obtained using a formal approach.

6.4.7. Lattices
Participant: Olivier Pantz.
With A. Raoult and N. Meunier (Université Paris Descartes), we have computed the asymptotic limit of a square lattice with three-points interactions. An article currently under review has been submitted on this work.

6.4.8. Homogenization of axon Bundles

Participant: Olivier Pantz.

With E. Mandonnet (Lariboisière Hospital), we have developed a new modeling for bundles of axons using homogenization technique. Previous works only focus (even if not explicitly) in the low density case: That is when the axon density is small. The aim is to determine which kind of electrical stimulation could trigger a signal into the axon. Under the low density assumption, the external electric field is independent of the membrane potential of the axon. If not, both are strongly coupled. Moreover, we have performed numerical simulations to determine what is the best position of the electrodes to enable the activation of the axons. This work has lead to the publications of an article [20] and a technical report [29]. Finally, we have begun to investigate more realistic modelings of the ionic flux based on the works of FitzHugh-Nagumo with a student, Xinxin Cheng, who spent three months at the CMAP.

6.5. Diffusion MRI

6.5.1. Homogenized diffusion tensor and approximate analytical formulae for the long time ADC

Participants: Jing-Rebecca Li, Houssem Haddar.

We model the bulk magnetization in biological tissue due to a diffusion gradient at the voxel level by a two compartment Bloch-Torrey partial differential equation. The cell membranes are modelled as infinitely thin permeable interfaces. We show the simulated long time apparent diffusion tensor of the PGSE sequence is close to the effective diffusion tensor from homogenization theory for both isotropic and anisotropic diffusion. For nearly isotropic diffusion we give analytical approximate formulae for the long time apparent diffusion coefficient in two and three dimensions. The approximate formulae allow us to robustly estimate the change in the cellular volume fraction from ADC measurements before and after cell swelling if the cells are approximately uniform in size. We can also use the formulae to estimate the average cell size.

6.5.2. General ODE model of diffusion MRI signal attenuation

Participants: Jing-Rebecca Li, Hang Tuan Nguyen.

We model the magnetization in biological tissue due to a diffusion gradient by a two compartment Bloch-Torrey partial differential equation with infinitely thin permeable membranes. We formulate a ODE model for the magnetization and show the simpler ODE model is a good approximation to the Bloch-Torrey PDE model for a variety of gradient shapes. Using the ODE model we determine of the change in the cellular volume fraction from the signal attenuation obtained before and after cell swelling. This method requires only the ADC and Kurtosis of the two signal attenuations and the numerical solution of an ODE system.
DISCO Team

6. New Results

6.1. Algorithmic study of linear functional systems

Participants: Alban Quadrat [correspondent], Thomas Cluzeau [ENSIL, Univ. Limoges]

In [77], [76], [97], it is shown that every linear functional system (e.g., PD systems, differential time-delay systems, difference systems) is equivalent to a linear functional system defined by an upper block-triangular matrix of functional operators: each diagonal block is respectively formed by a generating set of the elements of the system satisfying a purely $i$-codimensional system. Hence, the system can be integrated in cascade by successively solving (inhomogeneous) $i$-codimensional linear functional systems to get a Monge parametrization of its solution space [129][89]. The results are based on an explicit construction of the grade/purity filtration of the module associated with the linear functional system. This new approach does not use involved spectral sequence arguments as is done in the literature of modern algebra [102], [103]. To our knowledge, the algorithm obtained in [77], [97] is the most efficient algorithm existing in the literature of non-commutative algebra. It was implemented in the PURITYFILTRATION package developed in Maple (see Section 5.7) and in the homalg package of GAP 4 (see Section 5.8).

Given a linear multidimensional system (e.g., ordinary/partial differential systems, differential time-delay systems, difference systems), Serre’s reduction aims at finding an equivalent linear multidimensional system which contains fewer equations and fewer unknowns. Finding Serre’s reduction of a linear multidimensional system can generally simplify the study of structural properties and of different numerical analysis issues, and it can sometimes help solving the linear multidimensional system in closed form. In [17], Serre’s reduction problem is studied for underdetermined linear systems of partial differential equations with either polynomial, formal power series or analytic coefficients and with holonomic adjoints in the sense of algebraic analysis [102], [103]. These linear partial differential systems are proved to be equivalent to a linear partial differential equation. In particular, an analytic linear ordinary differential system with at least one input is equivalent to a single ordinary differential equation. In the case of polynomial coefficients, we give an algorithm which computes the corresponding linear partial differential equation.

In [45], we give a complete constructive form of the classical Fitting’s lemma in module theory which studies the relation between equivalences of linear systems and isomorphisms of their associated finitely presented modules. The corresponding algorithms were implemented in the OREMORPHISMS package (see Section 5.5).

6.2. A new approach to classes of quasilinear PD systems

Participants: Alban Quadrat [correspondent], Thomas Cluzeau [ENSIL, Univ. Limoges], Daniel Robertz [Univ. Aachen]

Many partial differential systems appearing in mathematical physics, engineering sciences and mathematical biology are nonlinear. Unfortunately, algebraic analysis and $D$-module theory, which were successful for the algorithmic study of linear partial differential systems, cannot consider nonlinear PD systems. This project aims at developing a generalization of the algebraic analysis approach to certain classes of quasilinear partial differential systems appearing in mathematical physics and engineering sciences (e.g., Burgers’ equation, shallow water, Euler equations for an incompressible fluid, traffic flow, gas flow). In [44], we have shown how constructive methods of differential algebra and algebraic analysis could be combined to extend results obtained in [45] and [104] for linear PD systems and how they allowed us to compute conservation laws, internal symmetries and decompositions of the solution space of certain classes of quasilinear PD systems. The algorithms have been implemented in the JANETMORPHISMS package dedicated to this new approach (see Section 5.6).
6.3. Stabilization problems & Noncommutative geometry

**Participant:** Alban Quadrat [correspondent].

In [124], [123], [122], it was shown how the fractional representation approach to analysis and synthesis problems developed by Vidyasagar, Desoer, Callier, Francis, Zames..., could be recast into a modern algebraic analysis approach based on module theory (e.g., fractional ideals, algebraic lattices) and the theory of Banach algebras. This new approach successfully solved open questions in the literature. Basing ourselves on this new approach, we explain in [126] why the non-commutative geometry developed by Alain Connes is a natural framework for the study of stabilizing problems of infinite-dimensional systems. Using the 1-dimensional quantized calculus developed in non-commutative geometry and results obtained in [124], [123], [122], we show that every stabilizable system and their stabilizing controllers naturally admit geometric structures such as connections, curvatures, Chern classes... These results are the first steps toward the use of the natural geometry of the stabilizable systems and their stabilizing controllers in the study of the important $H_{\infty}$ and $H_2$-problems.

6.4. Stabilization of time-delay systems

**Participants:** Alban Quadrat [correspondent], Arnaud Quadrat [SAGEM, MASSY].

In [127], we study the stabilization problem of a linear system formed by a simple integrator and a time-delay. We show that the stabilizing controllers of such a system can be rewritten as the closed-loop system defined by the stabilizing controllers of the simple integrator and a distributed delay. This result is used to study tracking problems appearing in the study of inertially stabilized platforms for optical imaging systems.

6.5. Singular boundary problems

**Participants:** Georg Regensburger [correspondent], Anja Korporal [RICAM, Linz], Bruno Buchberger [RISC, Linz], Markus Rosenkranz [Univ. Kent].

In [55], we present results on algorithmic methods for singular boundary problems for linear ordinary differential equations. Moreover, the implementation of integro-differential operators and the corresponding algorithms for boundary problems in the computer algebra system Maple is discussed. The operations implemented for regular boundary problems include computing Green’s operators as well as composing and factoring boundary problems. For singular boundary problems, compatibility conditions and generalized Green’s operators can be computed. In [94], we give a survey and new results on our algebraic and symbolic approach to boundary problem developed over the last years. The construction of integro-differential operators and polynomials over an integro-differential algebra is described in detail along with a generic implementation of the corresponding canonical forms and algorithms.

6.6. Model of reaction networks

**Participants:** Georg Regensburger [correspondent], Stefan Müller [RICAM, Linz].

In a joint work, Stefan Müller and G. Regensburger propose a notion of generalized mass action systems that could serve as a more realistic model for reaction networks in intracellular environments; classical mass action systems capture chemical reaction networks in homogeneous and dilute solutions, see e.g. [111] and [114]. We show that several results of Chemical Reaction Network Theory carry over to the case of generalized mass action kinetics. Our main result essentially states that, if the sign vectors of the stoichiometric and the kinetic-order subspace coincide, there exists a unique positive complex balancing equilibrium in every stoichiometric compatibility class.

6.7. Ruin probabilities

**Participants:** Georg Regensburger [correspondent], Hansjörg Albrecher [Univ. Lausanne], Corina Constantinescu [Univ. Lausanne], Zbigniew Palmowski [Univ. Wrocław], Markus Rosenkranz [Univ. Kent].
In a cooperation with Hansjörg Albrecher, Corina Constantinescu (both University of Lausanne), Zbigniew Palmowski (University of Wroclaw), and Markus Rosenkranz (University of Kent), we developed a symbolic technique to obtain asymptotic expressions for ruin probabilities and discounted penalty functions in renewal insurance risk models when the premium income depends on the present surplus of the insurance portfolio. The analysis is based on boundary problems for linear ordinary differential equations with variable coefficients and the corresponding Green’s operators (integral operators), generalizing the approach from [99]. We obtain also closed-form solutions for more specific functions in the compound Poisson risk model.

6.8. Systems of linear ordinary integro-differential equations

Participants: Alban Quadrat, Georg Regensburger [correspondent].

A. Quadrat and G. Regensburger (in the frame of his grant) are working on a new approach for studying algebraic and algorithmic properties of systems of linear ordinary integro-differential equations with boundary conditions. In a recent series of papers, in particular, [100], [101], V. V. Bavula obtained numerous algebraic results for modules over the ring of integro-differential operators with polynomial coefficients using generalized Weyl algebras. We are interested in how far some of his approach can be made algorithmic and generalized to boundary problems. First results in this direction were presented at the Journées Nationales de Calcul Formel (JNCF 2011).

6.9. Stabilization of MIMO fractional systems with delays

Participants: Catherine Bonnet, Le Ha Vy Nguyen, Alban Quadrat.

In order to yield the set of all stabilizing controllers of a large class of MIMO fractional time-delay systems, we may look for coprime factorizations of the transfer function and their corresponding Bézout factors. As primary results, in considering $H_\infty$ stability, left coprime factorizations and left Bézout factors have been determined analytically from the transfer function. Then a particular stabilizing controller has been derived. We also proved the existence of right coprime factorizations.

6.10. Stability analysis of (fractional) delay systems of neutral type

Participants: Catherine Bonnet, André Fioravanti [UNICAMP], Le Ha Vy Nguyen.

The $H_\infty$-stability analysis of (fractional) delay systems of neutral type has been studied in [14]. It was shown that the chains of poles are asymptotic to vertical axes which position depend on the roots of a certain polynomial (easily determined from the given transfer function). In [14] the case of roots of multiplicity one was completely treated whereas the case of roots of multiplicity greater than one was considered only in the particular case of neutral systems whose transfer function is a product of transfer functions of neutral systems with one delay. This year, we have studied more deeply the case of roots of multiplicity greater than one for both, standard and fractional delay systems.

6.11. Matrix Norm Approach for Control of Linear Time-Delay Systems

Participants: Catherine Bonnet, André Fioravanti [UNICAMP], José Claudio Geromel [UNICAMP], Silviu Niculescu.

In [46], we have treated the time-delay linear systems control design in the framework of complete and partial information. We were able to find linear controllers that increase the first stability window imposing at the same time that the delay-free system is stable using some properties about the norms of the state-space matrices. Our method treated the design problem by numeric routines based on Linear Matrix Inequalities (LMI) arisen from classical linear time invariant system theory coupled together with a unidimensional search. Both the state and output feedback design, were solved.

6.12. Interval Observers

Participants: Frédéric Mazenc, Silviu Niculescu, Olivier Bernard [UNICAMP].
The technique, based on the notion of interval observer, is a recent state estimation technique, which offers the advantage of providing information on the current state of a system at any instant of time. The first interval observers where relying on the assumption that the system was cooperative or, roughly speaking, “almost” cooperative.

In the contribution [25], we have proved that, for any time-invariant exponentially stable linear system with additive disturbances, time-varying exponentially stable interval observers can be constructed. The technique of construction relies on the Jordan canonical form that any real matrix admits and on time-varying changes of coordinates for elementary Jordan blocks which lead to cooperative linear systems. We applied our to the case of linear systems with input and output that are detectable.

The paper [31] focused on the analysis and the design of families of interval observers for linear systems with a point-wise delay. First, we proved that classical interval observers for systems without delays are not robust with respect to the presence of delays, no matter how small delays are. Next, we have shown that, in general, for linear systems with delay, the classical interval observers endowed with a point-wise delay are unstable. A new type of design of interval observers enabling to circumvent these obstacles is proposed. It provides with framers that incorporate distributed delay terms. The proposed interval observers are assessed through a non-linear biotechnological model.

6.13. Partial Differential Equations

Participants: Frédéric Mazenc, Christophe Prieur [GIPSA Lab, CNRS].

In [33] and [65], for families of partial differential equations (PDEs) with particular stabilizing boundary conditions, we have constructed strict Lyapunov functions. The PDEs under consideration were parabolic and, in addition to the diffusion term, might contain a nonlinear source term plus a convection term. The boundary conditions were the classical Dirichlet conditions, or the Neumann boundary conditions or a periodic one. The constructions relied on the knowledge of weak Lyapunov functions for the nonlinear source term. The strict Lyapunov functions were used to prove asymptotic stability in the framework of an appropriate topology. Moreover, when an uncertainty is considered, our construction of a strict Lyapunov function made it possible to establish some robustness properties of Input-to-State Stability (ISS) type.

In [64], we have considered a family of time-varying hyperbolic systems of balance laws. The partial differential equations of this family can be stabilized by selecting suitable boundary conditions. For the stabilized systems, the classical technique of construction of Lyapunov functions provides with a function which is a weak Lyapunov function in some cases, but is not in others. We transform this function through a strictification approach which gives a time-varying strict Lyapunov function which allows us to establish asymptotic stability in the general case and a robustness property with respect to additive disturbances of Input-to-State Stability (ISS) type.


Participants: Frédéric Mazenc, Siviu Niculescu.

In the work [32], we propose a new approach for the stabilization of nonlinear time-varying forward-complete systems with delay in the input. This approach is a new reduction model approach, which relies on operators of a new type. It presents three advantages. First, the corresponding control laws do not include distributed terms. Second, it yields closed-loop systems with positive solutions that can be easily derived. Finally, the stabilized systems possess some robustness properties (for instance of ISS type) that can be estimated.

6.15. Backstepping with delay

Participants: Frédéric Mazenc, Siviu Niculescu, Mounir Bekaïk.
In the contributions [30] and [63], we revisited the problem of constructing globally asymptotically stabilizing control laws for nonlinear systems in feedback form with a known pointwise delay in the input. The result we obtained covers a family of systems wider than those studied in the literature and endows with control laws with a single delay, in contrast to those given in previous works which include two distinct pointwise delays or distributed delays. The strategy of design is based on the construction of an appropriate Lyapunov-Krasovskii functional. An illustrative example ends the paper.

6.16. Certifying good performance for adaptive systems

**Participants:** Frédéric Mazenc, Michael Malisoff [Louisiana State University].

The usual adaptive control problem is to design a controller that forces all trajectories of the system to track a prescribed trajectory, while keeping the estimator of the unknown constant parameter vector bounded. We studied the important and more difficult adaptive tracking and estimation problem of simultaneously (1) forcing the trajectories of the system to track a given trajectory and (2) identifying the parameter vector. This problem was known to be solvable when the regressor satisfies a persistency of excitation condition, but the known results did not provide a strict Lyapunov function for the augmented error dynamics and so could not certify good performance such as ISS with respect to uncertainty added to the controller.

Our main result from [118] covers adaptively controlled first-order nonlinear systems that satisfy the persistency of excitation condition and are affine in the unknown parameter vector. Our contribution consists in particular in constructing global strict Lyapunov functions for the augmented error dynamics. In [28], we extended [118] to adaptive tracking for nonlinear systems in feedback form with multiple inputs and unknown high-frequency gains multiplying the controllers. The control gains must be identified as part of the control objective. High-frequency gains are important for electric motors, flight dynamics, and robot manipulators. We used a persistency of excitation condition that again ensured tracking and parameter identification and led to the explicit construction of a global strict Lyapunov function for the closed loop augmented error dynamics. The strict Lyapunov function was key to proving integral ISS with respect to time varying uncertainties added to the unknown parameters.

6.17. Improved Model Predictive Control design

**Participants:** Sorin Olaru, Morten Hovd [NTNU, Trondheim, Norway].

New results [38] have been obtained toward the computation of feasible sets for linear model predictive control techniques, based on set relations and not on the conventional orthogonal projection. Further, the problem of computing suitable inner approximations of the feasible sets was considered. Such approximations are characterized by simpler polytopic representations, and preserve essential properties as convexity, positive invariance, inclusion of the set of expected initial states.

6.18. Particle Swarm Optimization for reduced order Hinfini controllers with constraints handling

**Participants:** Guillaume Sandou, Gilles Duc [Supelec (E3S) Control Department], Mohamed Yagoubi [Ecole des Mines de Nante].

Efficient dedicated methods have been developed for Hinfinity controller synthesis. However, such methods require translating the design objectives using weighting filters, whose tuning is not easy; in addition they lead to high order controllers which have to be reduced. A particle swarm optimization method is used to solve both problems successively: after having optimized the filters according to the design specifications using a full order controller but no reformulation of the constraints, a reduced-order one is computed using the obtained filters. Experimental tests for a pendulum in the cart exhibit much than satisfactory results [95], [52]. In addition, the design of Hinfinity static output feedback has been done and tested using the Compleib library and exhibiting similar results to those obtained with the HIFOO solver [86].
6.19. Ant colony for symbolic regression

**Participants:** Guillaume Sandou, Bianca Minodora Heiman.

The identification of systems is a key feature to get representative models and so to design efficient control laws. Numerous methods exist to identify the parameters of nonlinear systems when the global structure of the model is given. The problem appears to be much more difficult when this structure is unknown (symbolic regression). We introduce ant colony optimization (ACO) in solving the problem of non linear systems identification in the case of an unknown structure of the mode [53]. Numerical results prove the viability of the approach.

6.20. Robust optimization for energy management

**Participants:** Guillaume Sandou, Henri Borsenberger, Philippe Dessante [Supelec (E3S) Energy Department].

Many studies have considered the solution of Unit Commitment problems for the management of energy networks. In this field, earlier work addressed the problem in deterministic cases and in cases dealing with demand uncertainties. In this paper [15], the authors develop a method to deal with uncertainties related to the cost function. Indeed, such uncertainties often occur in energy networks (waste incinerator with an a priori unknown waste amounts, cogeneration plant with uncertainty of the sold electricity price...). The corresponding optimization problems are large scale stochastic non-linear mixed integer problems. The developed solution method is a recourse based programming one. The main idea is to consider that amounts of energy to produce can be slightly adapted in real time, whereas the on/off statuses of units have to be decided very early in the management procedure. Results show that the proposed approach remains compatible with existing Unit Commitment programming methods and presents an obvious interest with reasonable computing loads.

6.21. Modeling and control of Acute Myeloid Leukemia

**Participants:** José Luis Avila Alonso, Annabelle Ballesta [BANG project-team], Frédéric Bonnans [COMMANDS project-team], Catherine Bonnet, Jean Clairambault [BANG project-team], Xavier Dupuis [COMMANDS project-team], Pierre Hirsch [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Jean-Pierre Marie [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Faten Merhi [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Silviu Niculescu, Hitay Özbay [Bilkent University, Ankara, Turkey], Ruoping Tang [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris].

In order to better take into account physiological phenomena as well as better understand the effect of the new anti-FLT3 therapy for AML, we have modified the model M. Adimy and F. Crauste in two ways [88]:
- we have introduced a modeling of quick self-renewal of cells in each stage of maturation.
- we have modeled each phase of the proliferating compartment (that is $G_1$, $S$, $G_2$ and $M$) separately. For the time being, only the $M$-phase is supposed to have a fixed time duration as it is well-known that the short time necessary to perform mitosis is hardly submitted to any variation.

In parallel to this modeling task, Faten Merhi and Annabelle Ballesta have performed experiments in order to identify parameters of the model. These experiments (which will continue in 2012) tend to show that we will converge to a patient-dependant model.

6.22. Chronic myelogenous leukemia

**Participants:** Frédéric Mazenc, Siviu Niculescu, Peter Kim [Univ. of Sydney].
The paper [26] focuses on the stability analysis of a delay-differential system encountered in modeling immune dynamics during imatinib treatment for chronic myelogenous leukemia (CML). A simple algorithm is proposed for the analysis of delay effects on the stability. Such an algorithm takes advantage of the particular structure of the dynamical interconnections of the model. The analysis shows that the model yields three fixed points, two of which are always unstable and one of which is sometimes stable. The stable fixed point corresponds to an equilibrium solution in which the leukemia population is kept below the cytogenetic remission level. This result implies that, during imatinib treatment, the resulting anti-leukemia immune response can serve to control the leukemia population. However, the rate of approach to the stable fixed point is very slow, indicating that the immune response is largely ineffective at driving the leukemia population towards the stable fixed point. To extend the stability analysis with respect to the delay parameter, we conduct a global nonlinear analysis to demonstrate the existence of unbounded solutions. We provide sufficient conditions based on initial cell concentrations that guarantee unbounded solutions and comment on how these conditions can serve to predict whether imatinib treatment will result in a sustained remission based on a patient’s initial leukemia load and initial anti-leukemia T cell concentration.

6.23. Control of a model of human heart

Participants: Frédéric Mazenc, Michael Malisoff [Louisiana State University], Marcio de Queiroz [Louisiana State University].

We worked on the control of human heart rate during exercise [27], which is a problem that has implications for the development of protocols for athletics, assessing physical fitness, weight management, and the prevention of heart failure. We provided new stabilization techniques, based on the notion of tracking, for a recently-proposed nonlinear model for human heart rate response that describes the central and peripheral local responses during and after treadmill exercise. The control we proposed respect the sign constraint imposed by the model and we proposed observers to cope with the case (important from a practical point of view) where some of the variables are not measured.

6.24. Marine Robotic Surveys

Participants: Frédéric Mazenc, Michael Malisoff [Louisiana State University], Fumin Zhang [Georgia Tech].

The works [24], [61] was inspired by the recent Deepwater Horizon oil spill disaster. The goal was to develop and implement robotic surveying methods to evaluate the immediate and longer term environmental impacts of the oil spills. It was joint with Michael Malisoff from the LSU and a Georgia Tech robotics team led by Fumin Zhang. Robotic surveying methods provide a low cost and convenient way to collect data in marsh areas that are difficult to access by human based methods. We designed strict Lyapunov functions that made it possible to use ISS to quantify the robustness of collision avoiding curve tracking controllers under controller uncertainty. The controllers are designed to keep the robot a fixed distance from, but moving parallel to, a two dimensional curve. Four challenges in applying ISS to curve tracking are (a) the need to restrict the magnitudes of the uncertainty to keep the state in the state space and build a strict Lyapunov function, (b) the likelihood of time delays in the controllers in real time applications, (c) possible parameter uncertainty such as unknown control gains, and (d) generalizations to three dimensional curve tracking. We overcame challenge (a) by finding maximum bounds on the perturbations that maintain forward invariance of a nested family of hexagons that fill the state space and transforming a nonstrict Lyapunov function into a strict Lyapunov function on the full state space. To address challenge (b), we used a Lyapunov-Krasovskii approach from [119] to convert the strict Lyapunov function into a Lyapunov-Krasovskii functional. This led to an upper bound on the admissible controller delay that can be introduced into the controller while still maintaining ISS.
6. New Results

6.1. Generalizing dual-feasible functions to the multi-dimensional case

Participant: F. Clautiaux.

Dual-feasible functions have been used in the past to compute lower bounds and valid inequalities for different combinatorial optimization and integer programming problems. Until now, all the dual-feasible functions proposed in the literature were 1-dimensional functions. We extended the principles of dual-feasible functions to the m-dimensional case by introducing the concept of vector packing dual-feasible function. We explored the theoretical properties of these functions in depth, and we proposed general schemes for generating some instances of these functions. Additionally, we proposed and analyzed different new families of vector packing dual-feasible functions. All the proposed approaches were tested extensively using benchmark instances of the 2-dimensional vector packing problem. Our computational results show that these functions can approximate very efficiently the best known lower bounds for this problem and improve significantly the convergence of branch-and-bound algorithms.

6.2. Aggregation algorithms for network flow mathematical models

Participant: F. Clautiaux.

We designed a general framework for solving very large network flow mathematical models that use a pseudo-polynomial number of variables. It is based on an initial aggregation of the vertices of the model and its iterative refinement using different optimization techniques. This led to large improvements for a special case of vehicle routing problem. Several theoretical questions regarding convergence, worst-case analysis and approximation algorithms are raised by our work and are now under study.

6.3. New Price Setting models in the Energy Field

Participant: L. Brotcorne.

The electricity supply industry is facing in many countries a restructuring process towards deregulation and competition. In that context classical marginal cost based approaches based on estimation of cost production function and demand functions are not well-suited anymore. Indeed, the energy prices have to be defined not only to retrieve the production costs but also in order to take into account the consumer behavior. Consumers make their choice of service, or of energy provider in order to minimize their disutility values. Failing to recognize that may lead to tremendous lack on revenues. In order to capture this hierarchical decision process where a leader (the energy provider) takes explicitly into account the reaction of a follower (the consumers) in his decision process (see for example [1]) the energy pricing problems addressed in this subject have been modeled as bilevel programs.

6.4. Bi-level formulation for a Long-Distance Freight Transportation Problem

Participants: M. Diaby, L. Brotcorne, E.-G. Talbi.

A company wants to convey different types of products from origin \(i\) to points of destination \(j\). It can deliver the goods itself or hire a transport company, and subcontract part of the application. The transport company must offer attractive prices while aiming to maximize its profit. The aim of this problem is to determine rates that allow the carrier to maximize its revenues and remain affordable for the customer. The problem is modeled as a bilevel program at the first level, the carrier (leader) wants to maximize its revenues; at the second level, the client (follower) aims to minimize its expenses. A primal dual heuristic has been proposed to solve the problem.
6.5. Advances in Parallel Metaheuristics on GPU  

**Participants:** Thé Van Luong, Nouredine Melab, El-Ghazali Talbi.

Nowadays, GPU computing has recently been revealed effective to deal with time-intensive problems. This new emerging technology is believed to be extremely useful to speed up many complex algorithms. One of the major issues for metaheuristics is to rethink existing parallel models and programming paradigms to allow their deployment on GPU accelerators. Generally speaking, the major issues we have to deal with are: the distribution of data processing between CPU and GPU, the thread synchronization, the optimization of data transfer between the different memories, the memory capacity constraints, etc. The contribution of our work is to deal with such issues for the redesign of parallel models of metaheuristics to allow solving of large scale optimization problems on GPU architectures. Our objective is to rethink the existing parallel models and to enable their deployment on GPUs.

Thereby, the new results involve a new generic guideline for building efficient parallel metaheuristics on GPU (e.g. tabu search, iterated local search, island model for evolutionary algorithms, pareto local search or multi-start algorithms). Our challenge is to come out with the GPU-based design of the whole hierarchy of parallel models. In this purpose, very efficient approaches are proposed for CPU-GPU data transfer optimization, thread control, mapping of solutions to GPU threads or memory management. These approaches have been exhaustively experimented using eleven optimization problems and six GPU configurations. Compared to a CPU-based execution, experiments report up to 80-fold acceleration for large combinatorial problems and up to 2000-fold speed-up for a continuous problem. The different works related to our work have been accepted in a dozen of publications, including the IEEE Transactions on Computers journal.

6.6. Parallel Evolutionary Algorithms for Energy-Aware Scheduling  

**Participants:** Y. Kessaci, M. Mezmaz, N. Melab, E.-G. Talbi, D. Tuyttens.

In the last decades, energy becomes an increasingly important issue in computing and embedded systems. In computing systems, minimizing energy consumption can significantly reduce the amount of energy bills. The demand for computing systems steadily increases and the cost of energy continues to rise. In embedded systems, reducing the use of energy allows to extend the autonomy of these systems. In addition, the reduction of energy decreases greenhouse gas emissions. Therefore, many researches are carried out to develop new methods in order to consume less energy. In this work, we propose an overview of the main methods used to reduce the energy consumption in computing and embedded systems.

As a use case and to give an example of a method, this work describes our new parallel bi-objective hybrid genetic algorithm that takes into account the completion time and the energy consumption. In terms of energy consumption, the obtained results show that our approach outperforms previous scheduling methods by a significant margin. In terms of completion time, the obtained schedules are also shorter than those of other algorithms.

6.7. A Parallel Bi-objective Hybrid Metaheuristic for Energy-Aware Scheduling for Cloud Computing Systems  


In this work, we investigate the problem of scheduling precedence-constrained parallel applications on heterogeneous computing systems (HCSs) like cloud computing infrastructures. This kind of applications was studied and used in many research works. Most of these works propose algorithms to minimize the completion time (makespan) without paying much attention to energy consumption. We propose a new parallel bi-objective hybrid genetic algorithm that takes into account, not only makespan, but also energy consumption. We particularly focus on the island parallel model and the multi-start parallel model. Our new method is based on dynamic voltage scaling (DVS) to minimize energy consumption. In terms of energy consumption, the obtained results show that our approach outperforms previous scheduling methods by a significant margin. In terms of completion time, the obtained schedules are also shorter than those of other algorithms. Furthermore, our study demonstrates the potential of DVS.
6.8. A Pareto-based GA for Scheduling HPC Applications on Distributed Cloud Infrastructures

Participants: N. Melab, E.-G. Talbi.

Reducing energy consumption is an increasingly important issue in cloud computing, more specifically when dealing with High Performance Computing (HPC). Minimizing energy consumption can significantly reduce the amount of energy bills and then increase the provider’s profit. In addition, the reduction of energy decreases greenhouse gas emissions. Therefore, many researches are carried out to develop new methods in order to consume less energy. In this work, we present a multi-objective genetic algorithm (MO-GA) that optimizes the energy consumption, $CO_2$ emissions and the generated profit of a geographically distributed cloud computing infrastructure. We also propose a greedy heuristic that aims to maximize the number of scheduled applications in order to compare it with the MO-GA. The two approaches have been experimented using realistic workload traces from Feitelson’s PW A Parallel Workload Archive. The results show that MO-GA outperforms the greedy heuristic by a significant margin in terms of energy consumption and $CO_2$ emissions. In addition, MO-GA is also proved to be slightly better in terms of profit while scheduling more applications.

6.9. An Hybrid Multiobjective Method to solve Biclustering of Microarray Data

Participants: L. Jourdan, K. Seridia, E.-G. Talbi.

In this work, we propose an a multi-objective genetic algorithm (NSGA-II) with a heuristic to solve the biclustering problem of micro-array data. Due to the huge size of the datasets, we propose a new representation based on a string of integers and the associated operators. Experimental results on real data set show that our approach can find significant biclusters of high quality.

6.10. Flexibility and dynamic optimization

Participants: L. Jourdan, M. Khouadjia, E.-G. Talbi.

In this work we propose a new method that explicitly searches for routes that are flexible enough to allow easy adaptation after a new order has arrived. For that we propose a measure of flexibility, and show that better solutions can be obtained when the "planning horizon" for all subproblems is modified to take the flexibility of solutions into account.

6.11. Indicator-based Multi-objective Local Search

Participant: A. Liefooghe.

In the last few years, a significant number of multi-objective metaheuristics have been proposed in the litterature in order to address real-world problems. Local search methods play a major role in many of these metaheuristic procedures. We adapt a recent and popular indicator-based selection method in order to define a population-based multi-objective local search. The proposed algorithm is designed in order to be easily adaptable, parameter independent and to have a high convergence rate. The capacity of our algorithm to reach these goals is evaluated on a large bunch of experiments. Three combinatorial optimization problems are investigated: a flow-shop scheduling problem, a ring star problem and a nurse scheduling problem. The experiments show that our algorithm can be applied with success to different types of multi-objective optimization problems and that it outperforms some classical metaheuristics. Furthermore, the parameter sensitivity analysis enables us to provide some useful guidelines about how to set the main parameters.

6.12. Connectedness and Local Search for Bi-objective Knapsack Problems

Participant: A. Liefooghe.
In [33], we report an experimental study on a given structural property of connectedness of the set of Pareto-optimal solutions for two variants of the bi-objective knapsack problem. A local search algorithm that explores this property is then proposed and its performance is compared against exact algorithms in terms of running time and number of optimal solutions found. The experimental results indicate that this simple local search algorithm is able to find a representative set of Pareto-optimal solutions in most of the cases, and in much less time than exact approaches.

6.13. Set-based Multiobjective Fitness Landscapes
Participants: S. Verel, A. Liefooghe, C. Dhaenens.

Fitness landscape analysis aims to understand the geometry of a given optimization problem in order to design more efficient search algorithms. However, there is a very little knowledge on the landscape of multiobjective problems. In [43], we consider multiobjective optimization as a set problem. Then, we give a general definition of set-based multiobjective fitness landscapes. An experimental set-based fitness landscape analysis is conducted on the multiobjective NK-landscapes with objective correlation. The aim is to adapt and to enhance the comprehensive design of set-based multi-objective search approaches, motivated by an a priori analysis of the corresponding set problem properties. Our experimental study shows that tools from single-objective fitness landscapes can directly be extended for analyzing set-based multiobjective search approaches. The relevant features of multimodality and ruggedness has been highlighted for this particular class of problems.

Participants: S. Verel, A. Liefooghe, L. Jourdan, C. Dhaenens.

The structure of the search space explains the behavior of multiobjective search algorithms, and helps to design well-performing approaches. In this work, we analyze the properties of multiobjective combinatorial search spaces, and we pay a particular attention to the correlation between the objective functions. To do so, we extend the multiobjective NK-landscapes in order to take the objective correlation into account. We study the co-influence of the problem dimension, the degree of non-linearity, the number of objectives, and the objective correlation on the structure of the Pareto optimal set, in terms of cardinality and number of supported solutions [45], as well as on the number of Pareto local optima [46]. This work concludes with guidelines for the design of multiobjective local search algorithms, based on the main fitness landscape features. All our results show that no expectation on the performance of multiobjective local search algorithms can be drawn without taking the problem properties into account very precisely. Indeed, it has now become clear that the number of objectives is one of the key issue to explain a problem complexity, but we also pointed out that the objective correlation is at least as important. Multiobjective fitness landscape analysis plays a central role to explain the performance of local search algorithms, and to design more efficient methods, that suit better the problem features.

6.15. On the Neutrality of Combinatorial Optimization Problem to Design an Efficient Neutrality-based Local Search
Participants: M.-E. Marmion, C. Dhaenens, L. Jourdan, A. Liefooghe, S. Verel.

In the context of the permutation flowshop scheduling problem, a deep landscape analysis focused on the neutrality property has been driven [hal-00550356]. This analysis characterizes the neutral networks of the local optima in order to make propositions about the way to exploit it in algorithms. Hence, NILS has been designed in order to exploit the neutrality of local optima [hal-00563459]. As soon as a local optimum is found, the search is allowed to move to equivalent neighbors. Moreover, NILS gets only one parameter that controls the number of solutions allowed to be visited with the same fitness value. NILS has been tested on flowshop and has shown promising results.
6.16. Guiding the Search over Neutral Networks

**Participants:** M.-E. Marmion, C. Dhaenens, L. Jourdan, A. Liefooghe, S. Verel.

In a context of neutrality, VEGAS was designed to escape from neutral networks based on the evolvability of solutions, and on a multi-armed bandit by selecting the more promising solution from the neutral network [hal-00579990]. Its main feature is to consider the whole evaluated solutions of a neutral network rather than the last accepted solution as classical methods. VEGAS was tested on NKq-landscapes (problems built to present neutral properties) and results show the importance of considering the whole identified solutions from the neutral network and of guiding the search explicitly.

6.17. DAMS: Distributed Adaptive Metaheuristic Selection

**Participants:** B. Derbel, S. Verel.

In this work, we design a new Distributed Adaptive Metaheuristic Selection (DAMS) scheme. DAMS is dedicated to adaptive optimization in distributed environments. Given a set of metaheuristics, the goal of DAMS is to coordinate their local execution on distributed nodes in order to optimize the global performance of the distributed system. DAMS is based on three-layer architecture allowing nodes to decide distributively what local information to communicate, and what metaheuristic to apply while the optimization process is in progress. Within this context, we specialize DAMS by describing a particular instantiation called Select Best and Mutate (SBM). Its is a simple, yet efficient, adaptive distributed algorithm using an exploitation component allowing nodes to select the metaheuristic with the best locally observed performance, and an exploration component allowing nodes to detect the metaheuristic with the actual best performance. SBM features are analyzed from both a parallel and an adaptive point of view, and its efficiency is demonstrated through experimentations and comparisons with other adaptive strategies (sequential and distributed).

6.18. A Method to Combine Combinatorial Optimization and Statistics to Mine High-Throughput Genotyping Data

**Participants:** J. Hamon, C. Dhaenens, J. Jacques.

In collaboration with Gènes Diffusion, we are interested in high-throughput genotyping data in order to select a subset of genes explaining a trait of interest. We suggest to study these high-throughput data combining combinatorial optimization and statistical methods. A first method based on an ILS (Iterated Local Search) and using a statistical criterion to calculate the fitness was suggested and compared with classical statistical approaches.


**Participants:** J. Jacques, L. Jourdan, C. Dhaenens.

This work focuses on helping clinical trials investigators to screen more patients. First, we performed an analysis of clinical trial business and medical data available in French hospitals. We carried on several expert interviews. Then we developed a model to this problem as an association rules mining problem. After a statistical study of rule interestingness measures, we proposed an improvement of this model as a multi-objective combinatorial optimization problem.

6.20. Reducing Thread Divergence in GPU-based B&B Applied to the Flow-shop Problem

**Participants:** I. Chakroun, A. Bendjoudi, N. Melab.
Branch-and-Bound (B&B) algorithms are attractive methods for solving to optimality combinatorial optimization problems. Nevertheless, they are time-intensive when dealing with large problem instances. Therefore, several parallel B&B strategies based on large computer clusters and grids have been proposed in the literature. However, to the best of our knowledge no contribution has been proposed for designing B&B algorithms on GPUs (Graphic Processing Units). Because of their tremendous computing power and remarkable cost efficiency, GPUs have been recently revealed as a powerful way to achieve high performance on long-running scientific applications. In this research work, we propose to revisit the design and implementation of B&B algorithms on GPU. We focus on the parallel evaluation of the bounds since preliminary experiments performed on the Flow-Shop scheduling problem (FSP) have shown that the bounding operation consumes over 98% of the execution time of the B&B algorithm. To deal with thread divergence reduction issue caused by the bounding operation a code refactoring approach have been proposed.

6.21. Fitness Landscapes: Local Optima Network

Participant: S. Vérel.

A new methodology to study the structure of the configuration spaces of hard combinatorial problems. It consists in building the network that has as nodes the locally optimal configurations and as edges the weighted oriented transitions between their basins of attraction. We apply the approach to the detection of communities in the optima networks produced by two different classes of instances of a hard combinatorial optimization problem: the quadratic assignment problem (QAP). We provide evidence indicating that the two problem instance classes give rise to very different configuration spaces. For the so-called real-like class, the networks possess a clear modular structure, while the optima networks belonging to the class of random uniform instances are less well partitionable into clusters.
3. New Results

3.1. 3D matrix-free $P^1$-exact conservative interpolation

Participants: F. Alauzet [correspondant]

The interpolation stage is a crucial step for time-dependent mesh adaptive simulations. Indeed, if the accuracy of the solution is spoiled during this stage then it is lost for ever. In the past, we have demonstrated the superiority of conservative interpolation for 2D compressible flow simulations. Here, we develop the 3D extension of the method on tetrahedral meshes. The same conclusion arose. The main difficulty was to design a fast, accurate and robust mesh intersection algorithm.

3.2. A changing-topology ALE numerical scheme

Participants: F. Alauzet [correspondant] and G. Olivier

The main difficulty arising in numerical simulations with moving geometries is to handle the displacement of the domain boundaries, i.e., the moving bodies. Only vertices displacement is not sufficient to achieve complex movement such as shear. We proved that the use of edge swapping allows us to achieve such complex displacement. We therefore developed an ALE formulation of this topological mesh modification to preserve the solver accuracy and convergence order [31].

3.3. Advanced boundary layer meshing

Participants: F. Alauzet [correspondant], J. Castelneau, L. Marechal, D. Marcum and A. Loseille

We design a new method to generate structured boundary layer meshes which are mandatory to accurately compute compressible flows a high Reynolds number (several millions). It couple the specification of the optimal boundary layer from the geometry boundary and moving mesh techniques to extrude the boundary layer in an already existing mesh. The main advantage of this approach is its robustness, i.e., at each step of the algorithm we have always a valid mesh.

3.4. Applications du maillage à la cryptographie

Participants: T. Grosges [correspondant], D. Barchiesi, Michael François


3.5. Applications du maillage à l’électromagnétisme et modélisation multi-physiques

Participants: D. Barchiesi [correspondant], T. Grosges, Houman Borouchaki, Laurence Giraud-Moreau, Sameh Kessentini, Anis Chaari

3.6. Automatic decomposition of discretized surfaces for parallel processing

Participants: S.H. Lo, H. Borouchaki [correspondant], P. Laug

Free-form surfaces and industrial surface forms could nowadays be conveniently generated efficiently by laser-based digitizing techniques or from a CAD graphics system. Automatic algorithms are imperative to decompose complex triangulated objects consisting of thousands to millions of nodal points into simpler surface parts for parameterization and parallel processing. Non-manifold complicated discretized objects will be handled and simple closed or open surface parts (manifolds) are retrieved by means of pure topological considerations. Each of the simple surface part, which is topologically equivalent to an open surface, a sphere or a torus will be decomposed by entirely topological operations into n equal pieces based on a specified geometrical criterion such as surface area, Gaussian curvatures or magnitudes of dihedral angles, etc. Cut planes could be conveniently defined normal to the axes of inertia of the object to be decomposed. The cut plane which produces surface parts which best respect the given criterion will be chosen, and the two bisected surfaces will be made as equal as possible by means of some general balancing mechanisms. This procedure could be repeated as often as necessary until sub-surface parts satisfying the given criterion are obtained. Each piece of the decomposed surfaces having an intact topological boundary can then be processed independently in parallel, and all the pieces could be put back together to recover the original object or partially to represent a certain portion of the object [29].

3.7. Construction de maillages de degré 2 – Triangle et tétraèdre P2

Participants: P.L. George [correspondant], H. Borouchaki, P. Laug

There is a need for finite elements of degree 2 or more to solve various P.D.E. problems. This study discusses a method to construct such meshes in the case of triangular element (in the plane or for a surface) or tetrahedral element (in the volume case), restricting at degree 2. This first part considers the planar case and, to begin with, returns to Bézier curves and Bézier triangles of degree 2. In the case of triangles, the relation with Lagrange P2 finite element is shown. Validity conditions are discussed and some invalid elements are shown while proposing a method to correct them. A construction method is then proposed [34].

3.8. Continuous and Discrete Adjoints to the Euler Equations for Fluids

Participants: F. Alauzet [correspondant] and O. Pironneau

Adjoints are used in optimization to speed-up computations, simplify optimality conditions or compute sensitivities. Because time is reversed in adjoint equations with first order time derivatives, boundary conditions and transmission conditions through shocks can be difficult to understand. In this work, we analyzed the adjoint equations that arise in the context of compressible flows governed by the Euler equations of fluid dynamics. We showed that the continuous adjoints and the discrete adjoints computed by automatic differentiation agree numerically; in particular the adjoint is found to be continuous at the shocks and usually discontinuous at contact discontinuities by both [9].

3.9. Goal-oriented anisotropic mesh adaptation for unsteady problems

Participants: F. Alauzet [correspondant], A. Belme and A. Dervieux

We have extended our previous work on goal-oriented mesh adaptation to time dependent simulations. This requires to set up a global fixed point algorithm in which state and adjoint variables problems are solved. The adjoint problem is solved backward in time [24].

3.10. Hessian recovery techniques on isotropic and anisotropic meshes

Participants: M. Picasso, F. Alauzet [correspondant], H. Borouchaki and P.-L. George
Mesh adaptation required the use of second order derivatives, \textit{i.e.}, the Hessian. For second order numerical scheme, the provided numerical solution is only piecewise linear. Consequently, numerical methods are considered to recover second order derivatives, the famous recovery techniques. In this work, we have investigated several methods. Numerical results on 2D and 3D isotropic and anisotropic meshes indicate that the quality of the results is strongly linked to the mesh topology and that no convergence can be insured in general \cite{22}.

3.11. High Quality Geometric Meshing of CAD Surfaces

\textbf{Participants} : P. Laug [correspondant], H. Borouchaki

We propose a general scheme of an indirect approach for generating isotropic and anisotropic geometric meshes of a surface constituted by a conformal assembly of parametric patches, based on the concept of metric. The different steps of the scheme are considered and, in particular, the definition of the geometric metric at each point of the surface (internal to a patch, belonging to an interface or boundary curve, or extremity of such a curve) as well as its corresponding induced metric in parametric domains.

Isotropic or anisotropic geometric metrics can locally produce significant size variations (internal to a patch or across interface curves) and can even be discontinuous across the interface curves. The larger the rate of the mesh size variation, the worse is the shape quality of the resulting mesh. To control this size variation, various methodologies based on metric reduction have been proposed in the case of a continuous isotropic metric. We introduce a novel iterative mesh gradation approach for discontinuous metrics. The approach uses a particular metric reduction procedure in order to ensure the convergence of the gradation process. In particular, we show that in the worst case the anisotropic discontinuous geometric metric map is reduced to an isotropic continuous geometric metric map for which the gradation is controlled \cite{27}.

3.12. Linéarisation et maillage des surfaces paramétrées

\textbf{Participants} : P. Laug, H. Borouchaki [correspondant], E. Renaut

Nous proposons une méthodologie pour simplifier la paramétrisation des surfaces composées de carreaux paramétrés issues généralement des environnements CAO (conception assistée par ordinateur). Cette nouvelle paramétrisation est définie via la construction d’une triangulation adaptative appelée support pour chaque carreau. L’adaptation est gouvernée par le contrôle de l’écart entre la paramétrisation initiale et celle issue de la triangulation support. Ce support est utilisé pour générer les maillages de ces surfaces selon une approche indirecte, dans laquelle le maillage est généré via les domaines des paramètres. Le support permet de déconnecter le maillleur du système CAO, et en ce sens constitue un outil universel pour le maillage de telles surfaces \cite{19}.

3.13. Mesh adaptation for very high-order numerical scheme

\textbf{Participants} : F. Alauzet, A. Loseille [correspondant] and E. Mbinky

In the past, we have demonstrated that multi-scale anisotropic mesh adaptation is a powerful tool to accurately simulate compressible flow problem and to obtain faster convergence to continuous solutions. But, this was limited to second order numerical scheme. Nowadays, numerous teams are working on the development of very high-order numerical scheme (e.g. of third or greater order): Discontinuous Galerkin, Residual Distribution scheme, Spectral method, ...

This work extend interpolation error estimates to higher order numerical solution representation. We have examined the case of third-order accuracy. The first step is to reduce the tri-linear form given by the third order error term into a quadratic form based on the third order derivative. From this local error model, the optimal mesh is exhibited thanks to the continuous mesh framework.

3.14. Méthodes de remaillages adaptatifs pour le formage incrémental de tôles minces

\textbf{Participants} : L. Moreau [correspondant], A. Cherouat, H. Borouchaki
Développement d’une méthode de remaillage adaptatif surfacique 3D permettant de raffiner de déraffiner le maillage localement autour de l’outil sphérique au cours des simulations numériques de formage incrémental.

3.15. Méthodes de remaillages adaptatifs surfacique dans le cadre des simulations numériques d’emboutissage de structure minces

Participants : L. Moreau [correspondant], A. Cherouat, H. Borouchaki
Développement de méthodes de remaillage adaptatif surfacique 3D, développement de méthodes d’interpolation et transfert des champs, interface avec le code EF Abaqus et application sur des exemples concrets de mise en forme de structure métalliques et composites.

3.16. Multi-scale anisotropic mesh adaptation for unsteady problems

Participants : F. Alauzet [correspondant] and G. Olivier
We focused on the extension of the multi-scale anisotropic mesh adaptation to unsteady flows. It leads to the development of a global fixed point mesh adaptation algorithm and space-time error estimates. Moreover, the mesh adaptation methodology has been extended to the case of moving meshes simulations [25].

3.17. Parallel CAD surface meshing

Participants : P. Laug, H. Borouchaki [correspondant]
A wide range of surfaces can be defined by means of composite parametric surfaces as is the case for most CAD modelers. There are, essentially, two approaches to meshing parametric surfaces: direct and indirect. Popular direct methods include the octree-based method, the advancing-front-based method and the paving-based method working directly in the tridimensional space. The indirect approach consists in meshing the parametric domain and mapping the resulting mesh onto the surface. Using the latter approach, we have proposed a general meshing scheme which consists in discretizing each interface curve and meshing each parametric domain according to the above boundary discretizations. Complex surfaces such as a car engine or a complete aircraft are composed of thousands of patches, and meshing these surfaces using the above sequential scheme can be inefficient. We propose a parallel version of the general meshing scheme while naturally balancing the load to each processor.

3.18. Quelques avancées dans les algorithmes de maillages tétraédriques

Participants : H. Borouchaki, P.-L. George [correspondant], L. Marechal
Le comportement en complexité des algorithmes de triangulation sur les "gros" maillage nous amène à utiliser les algorithmes de renumérotation de type Hilbert qui minimise les défauts de cache. Cette technique est également utilisée comme aide à l’optimisation des "gros" maillages. L’algorithme de renumérotation est multi-cœurs.
Des triangulations de plusieurs dizaines de millions de sommets sont construites en utilisant un "simple" ordinateur. La vitesse d’insertion trole le million de tétraèdre à la seconde.

3.19. Reconstruction de surface 3D à partir d’images numériques 2D

Participants : L. Moreau [correspondant], H. Borouchaki, A. Cherouat
Mise au point d’une méthode d’acquisition de la morphologie du buste féminin, développement d’algorithmes de reconstruction de surface 3D d’un buste féminin à partir de photos numériques dans l’objectif de modéliser le comportement du sein en modes statiques et dynamiques.

3.20. Validité des éléments finis usuels

Participants : H. Borouchaki, P.-L. George [correspondant], P. Laug, L. Marechal
On étudie les conditions assurant la validité géométrique des éléments finis usuels de degré 1 et 2. La formulation éléments finis mne conduisant pas toujours à une conclusion simple, on formule les éléments finis sous leur forme de Bézier. Ceci conduit à exhiber des conditions suffisantes (parfois nécessaires et suffisantes) de validité des éléments, c’est-à-dire de la positivité de leur jacobien. Pour les éléments de degré 2, on donne l’interprétation géométrique de ces conditions. Les éléments étudiés sont le triangle à 3 nœuds, le triangle à 6 nœuds, le quadrilatère à 4 nœuds et les quadrilatère à 8 et 9 nœuds, le tétraèdre à 4 nœuds et le tétraèdre à 10 nœuds puis les hexaèdres à 8, 27 et 20 nœuds.

Le cas du simplexe de dimension quelconque et de degré quelconque est traité.

Quelques remarques sur les maillages de surface de degré 2 indiquent quelques pistes à suivre ou à éviter lors de la construction de tels éléments.

### 3.21. Visualisation et modification des maillages courbes d’ordre élevé

**Participants**: J. Castelneau, A. Loseille [correspondant], L. Maréchal

Dans le cadre du projet ILab, des nouveaux algorithmes de visualisation et de modifications interactives des maillages courbes et hybrides ont été développés. En effet, une des principales difficultés dans la génération de maillages courbes reste la visualisation. Il est également nécessaire de disposer d’algorithmes de corrections interactifs car les maillages de surfaces initiaux (de degré 2) sont pour la plupart faux.
GECO Team

5. New Results

5.1. New results: geometric control

A first set of new results concerns sub-Riemannian geometry.

- In [3] we continued the study of almost-Riemannian structures, which are rank-varying sub-Riemannian structures locally generated by a number of vector fields equal to the dimension of the ambient manifold. In particular, two-dimensional almost-Riemannian structures are generalized Riemannian structures on surfaces for which local orthonormal frames are Lie bracket generating pair of vector fields that can become collinear. We considered the Carnot–Carathéodory distance canonically associated with an almost-Riemannian structure and studied the problem of Lipschitz equivalence between two such distances on a given compact oriented surface. We analyzed the generic case, allowing in particular for the presence of tangency points, i.e., points where two generators of the distribution and their Lie bracket are linearly dependent. The main result of the paper provides a characterization of the Lipschitz equivalence class of an almost-Riemannian distance in terms of a labeled graph associated with it.

- In [1] we studied nilpotent 2-step, corank 2 sub-Riemannian metrics. Such metrics naturally appear as nilpotent approximations of general sub-Riemannian ones. We exhibited optimal syntheses for these problems. It turns out that in general the cut time is not equal to the first conjugate time but has a simple explicit expression. As a byproduct of this study we proved some smoothness properties of the spherical Hausdorff measure in the case of a generic 6-dimensional, 2-step corank 2 sub-Riemannian metric.

- In [12] we started from the remark that in Carnot–Carathéodory spaces the class of 1-rectifiable sets does not contain smooth non-horizontal curves. We were looking for a new definition of rectifiable sets including non-horizontal curves. We introduced, for any metric space, a new class of curves, called continuously metric differentiable of degree \( k \), which are Hölder but not Lipschitz continuous when \( k > 1 \). Replacing Lipschitz curves by this kind of curves we defined \((H^k,1)\)-rectifiable sets and showed a density result generalizing the corresponding one in Euclidean geometry. This theorem has been obtained as a consequence of computations of Hausdorff measures along curves, for which we gave an integral formula. In particular, we showed that both spherical and standard Hausdorff measures along curves coincide with a class of dimensioned lengths and are related with an interpolation complexity, for which estimates have already been obtained in Carnot–Carathéodory spaces.

A class of problems for which tracking and motion planning is crucial, is given by the control of unmanned aerial vehicles (UAV). In order to develop improved planning tasks that take into account payload requirements, optimal costs and obstacles avoidance (or no flight zones), it is important to develop reliable and flexible simulators. One such simulator for a UAV ground control station is proposed in [9]. The research focuses on the connection between the UAV trajectories and its sensors. Our proposal includes a module-based description of the architecture of the simulator and is based on a nonlinear model of a fixed wing aircraft.

5.2. New results: quantum control

New results have been obtained for the control of the bilinear Schrödinger equation, with two different approaches.
• In [2] we proved an approximate controllability result by finite-dimensional methods, considering
the Galerkin approximations. The approach improves the technique that we developed in [40].
The result requires less restrictive non-resonance hypotheses on the spectrum of the uncontrolled
Schrödinger operator than those already known. The control operator is not required to be bounded
and we are able to extend the controllability result to the density matrices. The proof is based on
fine controllability properties of the finite-dimensional Galerkin approximations and allows to get
estimates for the $L^1$ norm of the control. The general controllability result is applied to the problem
of controlling the rotation of a bipolar rigid molecule confined on a plane by means of two orthogonal
external fields.

• In [4] we presented a constructive method to control the bilinear Schrödinger equation via
two controls. The method is based on adiabatic theory and works if the spectrum of the Hamiltonian admits conical eigenvalue intersections. We provided sharp estimates of the relation between the error and the controllability time. We also showed that for a Hamiltonian of the kind
$-\Delta + V_0(x) + u_1 V_1(x) + u_2 V_2(x)$ on a domain of $\mathbb{R}^n$ the eigenvalue intersections are conical
generically with respect to $V_0, V_1, V_2$.

5.3. New results: neurophysiology

We gave new contributions to the developing theory of human locomotion modeled through optimal control
problems. In this paradigm, the trajectories are assumed to be solutions of an optimal control problem whose
cost has to be determined.

• The purpose of [6] has been to analyze the class of optimal control problems defined in this way.
We proved strong convergence of their solutions, on the one hand for perturbations of the initial and
final points (stability), and on the other hand for perturbations of the cost (robustness).

• In [5] we discussed the modeling of both the dynamical system and the cost to be minimized, and
we analyzed the corresponding optimal synthesis. The main results describe the asymptotic behavior
of the optimal trajectories as the target point goes to infinity.

In [10] we studied the model of geometry of vision due to Petitot, Citti and Sarti [81]. One of the main
features of this model is that the primary visual cortex V1 lifts an image from $\mathbb{R}^2$ to the bundle of directions of
the plane. Neurons are grouped into orientation columns, each of them corresponding to a point of this bundle.
In this model a corrupted image is reconstructed by minimizing the energy necessary for the activation of
the orientation columns corresponding to regions in which the image is corrupted. The minimization process
intrinsically defines an hypoelliptic heat equation on the bundle of directions of the plane. In the original
model, directions are considered both with and without orientation giving rise respectively to a problem on the
group of rototranslations of the plane $SE(2)$ or on the projective tangent bundle of the plane. We provided a
mathematical proof of several important facts for this model. We first proved that the model is mathematically
consistent only if directions are considered without orientation. We then proved that the convolution of a
$L^2(\mathbb{R}^2, \mathbb{R})$ function (e.g. an image) with a 2D Gaussian is generically a Morse function. This fact is important
since the lift of Morse functions to the projective tangent bundle of the plane is defined on a smooth manifold.
We then provided the explicit expression of the hypoelliptic heat kernel on the projective tangent bundle of the
plane in terms of Mathieu functions. Finally, we presented the main ideas of an algorithm which allows
in real non-academic images. The algorithm is massively parallelizable and
needs no information on where the image is corrupted.

5.4. New results: switched systems

New results on switched systems have been obtained in three directions:

• **Discrete-time systems.** In [14] we dealt with the stability properties of linear discrete-time switched
systems with polytopic sets of modes. The most classical and viable way of studying the uniform
asymptotic stability of such a system is to check for the existence of a quadratic Lyapunov function. It
is known from the literature that letting the Lyapunov function depend on the time-varying switching parameter improves the chance that a quadratic Lyapunov function exists. The contribution of [14] is twofold. We first proved that under a non-degeneracy assumption the dependence on the switching function can be actually assumed to be linear with no prejudice on the effectiveness of the method. Moreover, we showed that no gain is obtained even if we allow the Lyapunov function to depend on the time. Second, we introduced the notion of eventual accessible sets and we showed that, in the degenerate case, it leads to a relaxation of the LMI conditions to check stability of switched linear systems. As a consequence, equivalence between different notions of quadratic stability can still be established under an additional assumption but, in general, allowing the Lyapunov function to depend on time leads to less conservative LMI conditions, as we explicitly showed through an example. We also discussed the case where the variation of the switching parameter is bounded by a prescribed constant between two subsequent times.

- **Continuous-time systems subject to persistent-excitation.** In [11] we studied linear control systems for which the controlled part can be switched off by a signal subject to a persistent excitation condition. We were interested in the stabilization problem of this system by a linear state feedback and we positively answered a question asked in [41], proving the following: Assume that the class of persistently exciting signals is restricted to those which are $M$-Lipschitzian, where $M > 0$ is a positive constant. Then, given any $C > 0$, there exists a linear state feedback depending on the class of signals under consideration (but not an individual signal) so that the rate of exponential decay of the time-varying system associated with any signal is greater than $C$.

- **Infinite-dimensional continuous-time systems.** In [13] we partially extended the analysis of finite-dimensional systems subject to persistently exciting signals to the case of systems driven by PDEs. More precisely, we studied the asymptotic stability of a dissipative evolution in a Hilbert space subject to intermittent damping. We observed that, even if the intermittence satisfies a persistent excitation condition, if the Hilbert space is infinite-dimensional then the system needs not being asymptotically stable (not even in the weak sense). Exponential stability is recovered under a generalized observability inequality, allowing for time-domains that are not intervals. Weak asymptotic stability is obtained under a similarly generalized unique continuation principle. Strong asymptotic stability is proved for intermittences that do not necessarily satisfy some persistent excitation condition, evaluating their total contribution to the decay of the trajectories of the damped system. Our results are discussed using the example of the wave equation and the linear Schrödinger equation.
6. New Results

6.1. Multiplicative cascades in real/synthetic oceanographic signals: application to the evaluation of ocean dynamics

**Participants:** Hussein Yahia, Oriol Pont, Joel Sudre, Véronique Garçon, Claire Pottier [CNES], Antonio Turiel, Christine Provost.

This work is performed during the final year of the HIRESUBCOLOR contract with CNES/NASA. From a fundamental point of view, significant advances have been worked out in the application of complex systems methods for the derivation of new methods for computing ocean dynamics at high spatial resolution (high resolution Sea Surface Temperature, pixel size: 4 x 4 kms). No temporal information is used. Instead, the norms and orientations of low resolution vector fields derived from altimetry and scatterometers are propagated along the scales of turbulent signals. This year, specific study on the propagation of the norm of the vector field has been conducted, resulting in the complete mapping (norm and orientation) of ocean dynamics at the high resolution of Sea Surface Temperature data. Validation is performed by comparison with the output of the ROMS 3D simulation model, with excellent results, and buoy validation is under way. H. Yahia has been invited to the AGU (American Geophysical Union) Fall meeting in San Francisco to make a presentation of HIRESUBCOLOR results (December 5-9 2011) and also to the EGU meeting to be held in Vienna in 2012. In 2011 the complete method for the determination of ocean dynamics has been finalized and it includes:

- the determination of both norm and orientation of the vector fields,
- the propagation along the scale of both geostrophic and ageostrophic dynamics.

These methods are generic, and can be applied to the determination of high resolution information for ocean/atmosphere interaction. For that matter, the Oceanflux proposal has been submitted and accepted, starting November 1, 2011, and a new proposal called MULTICARO will be submitted to CNES-OSTST. Related publications: [28], [21].

6.2. Endocardial potential analysis for cardiac arrhythmias

**Participants:** Oriol Pont, Hussein Yahia, Harish Kumar Goddabanahalli, Michel Haissaguerre, Nicolas Derval, Mélèze Hocini.

Cardiac diseases are the main cause of morbidity and mortality in western countries. Both the pathogenic areas and the evolution of the condition are complex to detect and estimate in the case of arrhythmias, specially in atrial and ventricular fibrillation. In these cases, the dynamics of the cardiac potential behaves chaotically in a highly complex way that challenges its description. Under this context, we are working in the direction of the characterization of the heartbeat dynamics in a model-agnostic way. We have performed analysis of heartbeat dynamics through singularity spectrum, empirical analysis of LPEs for the deart-beat Dynamics, and a novel innovative method to distinguish arrhythmic heartbeat with rhythmic heartbeat has been studied.

Intracardial potential is measured by means of electrode catheters used during the radiofrequency ablation procedure for cases of atrial fibrillation. We have found that in the electric potential signal, a simple fast-changing three-state orientation signal can be sifted from its complex but slow-changing modulation. The fast dynamics experimentally fits a Markovian process which can be described in a simple and compact way and its parameters are robustly estimated. This shows a clear change of signature even with small statistics that can be used to detect transitions in the arrhythmia and identify different regimes in them.

Related publications: [18], [23], [19], [24].
6.3. Phonetic segmentation

Participants: Vahid Khanagha, Joshua Winebarger, Khalid Daoudi, Oriol Pont, Hussein Yahia, Régine André-Obrecht.

Previously we had developed a novel phonetic segmentation method based on Microcanonical Multiscale Formalism (MMF). The algorithm was based on precise computation of Local Predictability Exponents (LPEs) at each point, and then using their integration over time axis (ACC) as a quantitative representative of changes in behavior of distribution of these exponents between neighboring phonemes. The piecewise linear estimation of ACC had provided very good segmentation precision. By performing error analysis of the original algorithm, we proposed a 2-step technique which better exploits LPEs to improve the segmentation accuracy. In the first step, we detect the boundaries of the original signal and of a low-pass filtered version, and we consider the union of all detected boundaries as candidates. In the second step, we use a hypothesis test over the local LPE distribution of the original signal to select the final boundaries. In summary following steps have been taken:

- Detailed error analysis of the original method, which resulted in the realization of the fact that a high-pass filtering can help to detect some of the missed boundaries.
- Development of the hypothesis test method, using the Log Likelihood Ratio Test for final decision over a list of candidates.
- Evaluation of the overall 2-step algorithm on the whole train part of the TIMIT database, to compare with the original method.
- Evaluation on test part of TIMIT database to compare with the state of the art methods.

Related publications: [13], [14].

We continued and improved the adaptation of speaker segmentation methods to develop new (nonlinear) techniques for phonetic segmentation. We succeeded in proposing simple and efficient new algorithms that outperform existing ones. Even with new approaches, our nonlinear approach was still competitive.

Related publications: [22], [27].

6.4. Optimal wavelets, unpredictable points manifold and the emergence of complexity

Participants: Oriol Pont, Hussein Yahia, Suman Maji.

We have found new theoretical developments that link the optimal wavelet description with the information transfer that characterizes the singularity exponents in complex signals. This fact is particularly relevant when there exists a microcanonical cascade as an effective dynamics for the underlying complex system. The implication of this is that under a multiscale hierarchy, the unpredictable set of a signal can be described in terms of its optimal wavelet coefficients and the multiplicative cascade relations between them, easing the information inference or reconstructability between resolution levels.

A new method for the detection of the unpredictable points manifold has been developed, enhancing previous implementations. The algorithm exploits the basic signal symmetries that can be easily verified and has the advantage of not assuming any underlying model. That work is the result of the collaboration between our team and A. Turiel’s team at Institute of Marine Sciences of Barcelona.

Additionally, we have developed a new algorithm that allows for the first time a very robust detection of the optimal wavelet in 2D signals. The main advantage of this new algorithm is that it optimizes the wavelet shape in a totally unconstrained way, therefore not restricting to specific wavelet families.

Related Publications: [7], [20].

6.5. Discriminative learning for automatic speaker recognition

Participants: Khalid Daoudi, Reda Jourani, Régine André-Obrecht.
Most of the speaker recognition systems rely on generative learning of Gaussian Mixture Models (GMM). During the last decade, discriminative approaches have been an interesting and valuable alternative to address directly the classification problem. For instance, Support Vector Machines (SVM) combined with GMM supervectors are among state-of-the-art approaches in speaker recognition. Recently a new discriminative approach for multiway classification has been proposed, the Large Margin Gaussian mixture models (LM-GMM). These latter methods have the same advantage as SVM in term of the convexity of the optimization problem to solve. However they differ from SVM because they draw nonlinear class boundaries directly in the input space, and thus no kernel trick is required. We continued our work on investigating simplified versions of LM-GMM for speaker recognition that can handle large scale databases. We developed a new and efficient learning algorithm and evaluated it on NIST-SRE data. The results show that this new algorithm not only outperforms both the original LM-GMM and the traditional GMM, but also outperforms state-of-the-art discriminative methods such as GMM-supervectors SVM.

Related Publications: [9], [10], [11], [12], [8].

6.6. A multiscale approach to phase reconstruction for Adaptive Optics

Participants: Suman Kumar Maji, Hussein Yahia, Oriol Pont, Thierry Fusco, Vincent Michau, Joel Sudre.

Atmospheric turbulence in Earth’s atmosphere upper layers plays a fundamental role in limiting the resolution of ground based instruments. These turbulent layers perturbate to a great extent incoming light from outer space. One of the best known solutions to overcome this hurdle is Adaptive Optics (AO). It provides real-time compensation by deforming a mirror through a servo-loop, according to phase measurements provided by a wavefront sensor (WFS). We propose and experiment with a new model for phase reconstruction from an acquired subimage of the perturbated phase: instead of reconstructing the phase gradient using conventional methods of AO, we propagate along the scales phase information, from the low resolution of the WFS to higher resolution, using specific wavelet projections that mimic inference along the scales associated to cascading properties of fully developed turbulence.

Related Publication: [17].

6.7. Reconstruction of Speech signal from its Unpredictable Points Manifold

Participants: Vahid Khanagha, Khalid Daoudi, Oriol Pont, Hussein Yahia.

Local Predictability Exponents (LPEs) can be used to classify a given signal’s samples according to their predictability. In particular, the Unpredictable Points Manifold, the subset of less predictable points can be formed as the ensemble of points having the least value of singularity exponents. We call these exponents the Local Predictability Exponents since they are are computed according to a procedure based on the evaluation of the degree of reconstruction at a given point. We demonstrate in the case of Speech signal that LPEs are key quantities related to predictability in the framework of reconstructible systems: it is possible to reconstruct the whole Speech signal by applying a reconstruction kernel to the UPM. This provides a strong indication of the importance of the UPM, already demonstrated for other types of complex signals. Experiments show that a UPM containing a small number of the points provides very good perceptual reconstruction quality. In summary following steps have been taken:

- Using the LPEs to form the UPM for Speech signal and coping with the implementation issues in particular case of Speech signal.
- Successful reconstruction of Speech signal from the UPM. The performance was measured using objective measures of reconstruction quality.
- Detailed study of geometrical implications of the points in UPM, and proposition of a new multiscale measure, to be used in estimation procedure of exponents, which is more appropriate for speech analysis. In fact, the same quality of reconstruction is achieved with a quite smaller UPM.
- Development of a very simple compression algorithm (8-bit differential nonuniform quantizer) which overperforms the traditional DPCM coding method.
6.8. New upwelling indices from complex system methods

**Participants:** Hussein Yahia, Ayoub Tamim, Khalid Minaoui, Driss Aboutajdine, Véronique Garçon, Joel Sudre.

We started the Volubilis project on the study of upwelling in the Moroccan coast by satellite imaging. A Ph.D. student, Ayoub Tamim, was recruited in January 2011. We organized the kick-off meeting of the project in Rabat from June 6 to June 10 with the presence of all partners (INRIA, LEGOS, FSR and CRTS). A. Tamim implemented a new version of the software CRTS is using to compute an upwelling index. He then joined GEOSTAT (from September 5 to December 15, 2011) where he first constituted a small validation database of SST and Chlorophyll images. He then implemented some algorithms such as Morand index and multiscale entropy. The goal being to define in a first step a coarse segmentation of upwelling regions.

6.9. A detailed analysis of multisensor fusion of moderate resolution imaging spectroradiometer

**Participants:** Harish Kumar Goddabanahalli, Dharmendra Singh, Hussein Yahia.

Related publication: [16].
I4S Team

6. New Results

6.1. identification of linear systems

6.1.1. Modular identification and damage detection for large structures

Participants: Michael Döhler, Laurent Mevel.

In Operational Modal Analysis (OMA) of large structures it is often needed to process sensor data from multiple non-simultaneously recorded measurement setups, especially in the case of large structures. In this work a new efficient variant of the PreGER algorithm is presented that avoids the numerical explosion of the calculation by using a modular approach, where the data from the measurement setups is processed setup by setup and not at the same time [16].

6.1.2. Fast multi order subspace identification algorithm

Participants: Michael Döhler, Laurent Mevel.

Stochastic subspace identification methods are an efficient tool for system identification of mechanical systems in Operational Modal Analysis (OMA), where modal parameters are estimated from measured vibrational data of a structure. System identification is usually done for many successive model orders, as the true system order is unknown and identification in results at different model orders need to be compared to distinguish true structural modes from spurious modes in so-called stabilization diagrams. An algorithm to estimate the system matrices at multiple model orders has been derived [20].

6.1.3. Evaluation of confidence intervals and computation of sensitivities for subspace methods

Participants: Michael Döhler, Xuan Lam, Laurent Mevel.

In Operational Modal Analysis, the modal parameters (natural frequencies, damping ratios and mode shapes) obtained from Stochastic Subspace Identification (SSI) of a structure, are afflicted with statistical uncertainty. A variant of this approach has been derived for the Eigenvalue-Realization-Algorithm (ERA) [25]. Another version has been proposed for the merging subspace algorithm [17], [17]. This approach has been validated on large scale examples[14].

6.2. damage detection for mechanical structures

6.2.1. Damage detection and localisation

Participants: Michael Döhler, Laurent Mevel.

Statistical methods using output-only data have been shown to offer a robust solution to the damage detection task. These techniques have also been combined with sensitivities extracted from finite element models to offer information on the location of damage accounting for uncertainties in the finite element sensitivities. In some applications, however, the formulation of the finite element model makes implementation impractical and this motivates the search for model-free damage localization alternatives. One option is to use experimentally extracted sensitivities but their computation requires a set of constants (usually absorbed in the normalization of the eigenvectors) that are not available in output only identification. The noted limitation can be circumvented by adding a known perturbation to the mass distribution and repeating the output only identification, a procedure that can be practical in some cases. Linking a null-space based subspace damage index with experimentally extracted sensitivities allows us to infer on the position of damage without formulating a
finite element model and without the need for input measurements. The performance of the algorithm is illustrated on simulated data [19]. Damage detection has also been applied to a large scale example of an european project [23], [15].

6.2.2. Robust subspace damage detection

Participants: Michael Döhler, Laurent Mevel.

Subspace methods enjoy some popularity, especially in mechanical engineering, where large model orders have to be considered. In the context of detecting changes in the structural properties and the modal parameters linked to them, some subspace based fault detection residual has been recently proposed and applied successfully. However, most works assume that the unmeasured ambient excitation level during measurements of the structure in the reference and possibly damaged condition stays constant, which is not possible in any application. This work addresses the problem of robustness of such fault detection methods. A subspace-based fault detection test is derived that is robust to excitation change but also to numerical instabilities that could arise easily in the computations [21].

6.3. Instability monitoring of aeronautical structures

6.3.1. Subspace identification for hinged-blades helicopters

Participants: Ahmed Jhinaoui, Laurent Mevel.

In this work, an extension of the output-only subspace identification, to the class of linear periodically time-varying (LPTV) systems, is proposed. The goal is to identify a useful information about the system’s stability using the Floquet theory which gives a necessary and sufficient condition for stability analysis [24].

6.3.2. Optimal input design for identification and detection

Participants: Alireza Esna Ashari, Laurent Mevel.

Output only techniques rely on the presence on unknown turbulence, which may or may not be enough to excite the system. A new approach for applying artificial input to the system for maximizing detection and identifiability has been developed. This work considers the problem of auxiliary input design for subspace-based fault detection methods. In several real applications, particularly in the damage detection of mechanical structures and vibrating systems, environment noise is the only input to the system. In some applications, white noise produces low quality output data for the subspace-based fault detection method. In those methods, a residual is calculated to detect the fault based on the output information. However, some modes of the system may not influence the outputs and the residual appropriately if the input is not exciting enough for those modes. In this work, rotated inputs method is implemented to excite the system modes. In addition to produce a residual more sensitive to the weak modes, it is possible to detect system order changes due to the fault using the rotated inputs. Simulation results demonstrate the efficiency of injecting these auxiliary inputs to improve the subspace-based fault detection methodology [22]. This work is funded by FP7-NMP Large Scale Integrated Project IRIS.
5. New Results

5.1. Asymptotic preserving schemes

Participant: Nicolas Crouseilles.

In [18], we extend the micro-macro decomposition based numerical schemes developed previously to the collisional Vlasov-Poisson model in the diffusion and high-field asymptotics. In doing so, we first write the Vlasov-Poisson model as a system that couples the macroscopic (equilibrium) part with the remainder part. A suitable discretization of this micro-macro model enables to derive an asymptotic preserving scheme in the diffusion and high-field asymptotics. In addition, two main improvements are presented: On the one hand a self-consistent electric field is introduced, which induces a specific discretization in the velocity direction, and represents a wide range of applications in plasma physics. On the other hand, as suggested in a previous reference, we introduce a suitable reformulation of the micro-macro scheme which leads to an asymptotic preserving property with the following property: It degenerates into an implicit scheme for the diffusion limit model when \( \varepsilon \to 0 \), which makes it free from the usual diffusion constraint \( \Delta t = O(\Delta x^2) \) in all regimes. Numerical examples are used to demonstrate the efficiency and the applicability of the schemes for both regimes.

In [45], a Two-Scale Macro-Micro decomposition of the Vlasov equation with a strong magnetic field is derived. This consists in writing the solution of this equation as a sum of two oscillating functions with circumscribed oscillations. The first of these functions has a shape which is close to the shape of the Two-Scale limit of the solution and the second one is a correction built to offset this imposed shape.

5.2. Resolution of the quasi-neutrality equation

Participant: Nicolas Crouseilles.

In reference [39], different parallel algorithms are proposed for the numerical resolution of the quasi-neutrality equation in the GYSELA code. A set of benchmarks on a parallel machine has permitted to evaluate the performance of the different versions of the quasi-neutrality solver. In particular, in [40], these improvements are combined with memory optimization which enable a scalability of the GYSELA code up to 64k cores.

In [20], a new discretization scheme of the gyrokinetic quasi-neutrality equation is proposed. It is based on Isogeometric Analysis; the IGA which relies on NURBS functions, seems to accommodate arbitrary coordinates and the use of complicated computation domains. Moreover, arbitrary high order degree of basis functions can be used. Here, this approach is successfully tested on elliptic problems like the quasi-neutrality equation.

5.3. High order schemes for Vlasov-Poisson system

Participant: Nicolas Crouseilles.

In [44], we derive the order conditions for fourth order time splitting schemes in the case of the 1D Vlasov-Poisson system. Computations to obtain such conditions are motivated by the specific Poisson structure of the Vlasov-Poisson system: this structure is similar to Runge-Kutta-Nyström systems. The obtained conditions are proved to be the same as RKN conditions derived for ODE up to the fourth order. Numerical results are performed and show the benefit of using high order splitting schemes in that context.
In [19], we present a discontinuous Galerkin scheme for the numerical approximation of the one-dimensional periodic Vlasov-Poisson equation. The scheme is based on a Galerkin-characteristics method in which the distribution function is projected onto a space of discontinuous functions. We present comparisons with a semi-Lagrangian method to emphasize the good behavior of this scheme when applied to Vlasov-Poisson test cases.

The CEMRACS is an annual summer research session promoted by the SMAI. The 15th edition of 2010 has been organized by N. Crouseilles, H. Guillard, B. Nkonga and E. Sonnendrücker around "Numerical modeling of fusion plasmas". The volume [38] gathers artext resulting from research projects initiated during the CEMRACS 2010.

5.4. Second order averaging for the nonlinear Schrödinger equation with strong anisotropic potential

Participants: Florian Méhats, François Castella.

In [10], we consider the three dimensional Gross-Pitaevskii equation (GPE) describing a Bose-Einstein Condensate (BEC) which is highly confined in vertical $z$ direction. The confining potential induces high oscillations in time. If the confinement in the $z$ direction is a harmonic trap – an approximation which is widely used in physical experiments – the very special structure of the spectrum of the confinement operator implies that the oscillations are periodic in time. Based on this observation, it can be proved that the GPE can be averaged out with an error of order of $\epsilon$, which is the typical period of the oscillations. In this article, we construct a more accurate averaged model, which approximates the GPE up to errors of order $O(\epsilon^2)$. Then, expansions of this model over the eigenfunctions (modes) of the confining operator $H_z$ in the $z$-direction are given in view of numerical applications. Efficient numerical methods are constructed to solve the GPE with cylindrical symmetry in 3D and the approximation model with radial symmetry in 2D, and numerical results are presented for various kinds of initial data.

5.5. A problem of moment realizability in quantum statistical physics

Participant: Florian Méhats.

This work [34] is a generalization of the results previously obtained by F. Méhats and O. Pinaud, in J. Stat. Phys. (2010), in a one-dimensional setting: we revisit the problem of the minimization of the quantum free energy (entropy + energy) under local constraints (moments) and prove the existence of minimizers in various configurations. While the above quoted article addressed the 1D case on bounded domains, we treat in the present paper the multi-dimensional case as well as unbounded domains and non-linear interactions as Hartree/Hartree-Fock. Moreover, whereas this article dealt with the first moment only, namely the charge density, we extend the results to the second moment, the current density.

5.6. Orbital stability of spherical galactic models

Participant: Florian Méhats.

In [33] we consider the three dimensional gravitational Vlasov Poisson system which is a canonical model in astrophysics to describe the dynamics of galactic clusters. A well known conjecture is the stability of spherical models which are nonincreasing radially symmetric steady states solutions. This conjecture was proved at the linear level by several authors in the continuation of the breakthrough work by Antonov in 1961. In a previous work, we derived the stability of anisotropic models under spherically symmetric perturbations using fundamental monotonicity properties of the Hamiltonian under suitable generalized symmetric rearrangements first observed in the physics literature. In this work, we show how this approach combined with a new generalized Antonov type coercivity property implies the orbital stability of spherical models under general perturbations.
5.7. The Schrödinger Poisson system on the sphere

Participant: Florian Méhats.

In [31], we study the Schrödinger-Poisson system on the unit sphere \( S^2 \) of \( \mathbb{R}^3 \), modeling the quantum transport of charged particles confined on a sphere by an external potential. Our first results concern the Cauchy problem for this system. We prove that this problem is regularly well-posed on every \( H^s(S^2) \) with \( s > 0 \), and not uniformly well-posed on \( L^2(S^2) \). The proof of well-posedness relies on multilinear Strichartz estimates, the proof of ill-posedness relies on the construction of a counterexample which concentrates exponentially on a closed geodesic. In a second part of the paper, we prove that this model can be obtained as the limit of the three dimensional Schrödinger-Poisson system, singularly perturbed by an external potential that confines the particles in the vicinity of the sphere.

5.8. A boundary matching micro-macro decomposition for kinetic equations

Participant: Florian Méhats.

In [32], we introduce a new micro-macro decomposition of collisional kinetic equations which naturally incorporates the exact space boundary conditions. The idea is to write the distribution function \( f \) in all its domain as the sum of a Maxwellian adapted to the boundary (which is not the usual Maxwellian associated with \( f \)) and a reminder kinetic part. This Maxwellian is defined such that its ‘incoming’ velocity moments coincide with the ‘incoming’ velocity moments of the distribution function. Important consequences of this strategy are the following. i) No artificial boundary condition is needed in the micro/macro models and the exact boundary condition on \( f \) is naturally transposed to the macro part of the model. ii) It provides a new class of the so-called ‘Asymptotic preserving’ (AP) numerical schemes: such schemes are consistent with the original kinetic equation for all fixed positive value of the Knudsen number \( \epsilon \), and if \( \epsilon \to 0 \) with fixed numerical parameters then these schemes degenerate into consistent numerical schemes for the various corresponding asymptotic fluid or diffusive models. Here, the strategy provides AP schemes not only inside the physical domain but also in the space boundary layers. We provide a numerical test in the case of a diffusion limit of the one-group transport equation, and show that our AP scheme recovers the boundary layer and a good approximation of the theoretical boundary value, which is usually computed from the so-called Chandrasekhar function.

5.9. 1D quintic nonlinear equation with white noise dispersion

Participant: Arnaud Debussche.

Under certain scaling the nonlinear Schrödinger equation with random dispersion converges to the nonlinear Schrödinger equation with white noise dispersion. The aim of these works is to prove that this latter equation is globally well posed in \( L^2 \) or \( H^1 \). In [28], we improve the Strichartz estimates obtained previously for the Schrödinger equation with white noise dispersion in one dimension. This allows us to prove global well posedness when a quintic critical nonlinearity is added to the equation. We finally show that the white noise dispersion is the limit of smooth random dispersion.

5.10. Weak approximation of stochastic partial differential equations: the nonlinear case

Participant: Arnaud Debussche.

In [22] we study the error of the Euler scheme applied to a stochastic partial differential equation. We prove that as it is often the case, the weak order of convergence is twice the strong order. A key ingredient in our proof is Malliavin calculus which enables us to get rid of the irregular terms of the error. We apply our method to the case a semilinear stochastic heat equation driven by a space-time white noise.

5.11. Ergodic BSDEs under weak dissipative assumptions

Participant: Arnaud Debussche.
In [27] we study ergodic backward stochastic differential equations (EBSDEs) dropping the strong dissipativity assumption needed previously. In other words we do not need to require the uniform exponential decay of the difference of two solutions of the underlying forward equation, which, on the contrary, is assumed to be non degenerate. We show existence of solutions by use of coupling estimates for a non-degenerate forward stochastic differential equations with bounded measurable non-linearity. Moreover we prove uniqueness of “Markovian” solutions exploiting the recurrence of the same class of forward equations. Applications are then given to the optimal ergodic control of stochastic partial differential equations and to the associated ergodic Hamilton-Jacobi-Bellman equations.

5.12. Asymptotic first exit times of the Chafee-Infante equation with small heavy tailed noise
Participant: Arnaud Debussche.
Motivated by paleoclimatological issues, we determine in [26] asymptotic first exit times for the Chafee-Infante equation forced by heavy-tailed Levy diffusions from reduced domains of attraction in the limit of small intensity. We show that in contrast to the case of Gaussian diffusion the expected first exit times are polynomial in terms of the intensity.

5.13. Stochastic Cahn-Hilliard equation with double singular nonlinearities and two reflections
Participant: Arnaud Debussche.
In [25] we consider a stochastic partial differential equation with two logarithmic nonlinearities, two reflections at 1 and -1, and a constraint of conservation of the space average. The equation, driven by the derivative in space of a space-time white noise, contains a bi-Laplacian in the drift. The lack of a maximum principle for the bi-Laplacian generates difficulties for the classical penalization method, which uses a crucial monotonicity property. Being inspired by the works of Debussche, Goudenège, and Zambotti, we obtain existence and uniqueness of a solution for initial conditions in the interval \((-1, 1)\). Finally, we prove that the unique invariant measure is ergodic, and we give a result of exponential mixing.

5.14. Diffusion limit for a stochastic kinetic problem
Participants: Arnaud Debussche, Erwan Faou.
In [29] we consider numerical approximations of stochastic differential equations by the Euler method. In the case where the SDE is elliptic or hypo-elliptic, we show a weak backward error analysis result in the sense that the generator associated with the numerical solution coincides with the solution of a modified Kolmogorov equation up to high order terms with respect to the stepsize. This implies that every invariant measure of the numerical scheme is close to a modified invariant measure obtained by asymptotic expansion. Moreover, we prove that, up to negligible terms, the dynamic associated with the Euler scheme is exponentially mixing.

5.15. Convergence of stochastic gene networks to hybrid piecewise deterministic processes
Participant: Arnaud Debussche.
In [47] we consider numerical approximations of stochastic differential equations by the Euler method. In the case where the SDE is elliptic or hypo-elliptic, we show a weak backward error analysis result in the sense that the generator associated with the numerical solution coincides with the solution of a modified Kolmogorov equation up to high order terms with respect to the step-size. This implies that every invariant measure of the numerical scheme is close to a modified invariant measure obtained by asymptotic expansion. Moreover, we prove that, up to negligible terms, the dynamic associated with the Euler scheme is exponentially mixing.
5.16. Exponential mixing of the 3D stochastic Navier-Stokes equations driven by mildly degenerate noise

**Participant:** Arnaud Debussche.

In [11], we prove the strong Feller property and exponential mixing for 3D stochastic Navier-Stokes equation driven by mildly degenerate noises (i.e., all but finitely many Fourier modes are forced) via Kolmogorov equation approach.

5.17. Ergodicity results for the stochastic Navier-Stokes equations: an introduction

**Participant:** Arnaud Debussche.

In this survey article [46], we review recent progresses in the study of ergodicity for the stochastic Navier-Stokes equations. The first part introduces general concepts, the second deals with the 2D case and the 3D case is treated in the third part.

5.18. Local Martingale and Pathwise Solutions for an Abstract Fluids Model

**Participant:** Arnaud Debussche.

In the first article [23], we establish the existence and uniqueness of both local martingale and local pathwise solutions of an abstract nonlinear stochastic evolution system. The primary application of this abstract framework is to infer the local existence of strong, pathwise solutions to the 3D primitive equations of the oceans and atmosphere forced by a nonlinear multiplicative white noise. In the second article [24], global existence is obtained.

5.19. Geometric numerical integration and Schrödinger equations

**Participant:** Erwan Faou.

The goal of geometric numerical integration is the simulation of evolution equations by preserving their geometric properties over long times. This question is of particular importance in the case of Hamiltonian partial differential equations typically arising in many application fields such as quantum mechanics or wave propagations phenomena. This implies many important dynamical features such as energy preservation and conservation of adiabatic invariants over long times. In this setting, a natural question is to know how and to which extent the reproduction of such long time qualitative behavior is ensured by numerical schemes.

Starting from numerical examples, these notes [37] try to provide a detailed analysis in the case of the Schrödinger equation in a simple setting (periodic boundary conditions, polynomial nonlinearities) approximated by symplectic splitting methods. This text analyzes the possible stability and instability phenomena induced by space and time discretization, and provides rigorous mathematical explanations for them.

5.20. On the influence of the geometry on skin effect in electromagnetism

**Participant:** Erwan Faou.

In [14], we consider the equations of electromagnetism set on a domain made of a dielectric and a conductor subdomain in a regime where the conductivity is large. Assuming smoothness for the dielectric-conductor interface, relying on recent works we prove that the solution of the Maxwell equations admits a multiscale asymptotic expansion with profile terms rapidly decaying inside the conductor. This skin effect is measured by introducing a skin depth function that turns out to depend on the mean curvature of the boundary of the conductor. We then confirm these asymptotic results by numerical experiments in various axisymmetric configurations. We also investigate numerically the case of a nonsmooth interface, namely a cylindrical conductor.
5.21. **Reconciling alternate methods for the determination of charge distributions: A probabilistic approach to high-dimensional least-squares approximations**

**Participant:** Erwan Faou.

In [17], we propose extensions and improvements of the statistical analysis of distributed multipoles (SADM) algorithm put forth by CHIPOT in 1998 for the derivation of distributed atomic multipoles from the quantum-mechanical electrostatic potential. The method is mathematically extended to general least-squares problems and provides an alternative approximation method in cases where the original least-squares problem is computationally not tractable, either because of its ill-posedness or its high-dimensionality. The solution is approximated employing a Monte Carlo method that takes the average of a random variable defined as the solutions of random small least-squares problems drawn as subsystems of the original problem. The conditions that ensure convergence and consistency of the method are discussed, along with an analysis of the computational cost in specific instances.

5.22. **Hamiltonian interpolation of splitting approximations for nonlinear PDEs**

**Participant:** Erwan Faou.

In [30], we consider a wide class of semi linear Hamiltonian partial differential equations and their approximation by time splitting methods. We assume that the nonlinearity is polynomial, and that the numerical trajectory remains at least uniformly integrable with respect to an eigenbasis of the linear operator (typically the Fourier basis). We show the existence of a modified interpolated Hamiltonian equation whose exact solution coincides with the discrete flow at each time step over a long time. While for standard splitting or implicit-explicit schemes, this long time depends on a cut-off condition in the high frequencies (CFL condition), we show that it can be made exponentially large with respect to the step size for a class of modified splitting schemes.

5.23. **Energy cascades for NLS on the torus**

**Participant:** Erwan Faou.

In the work [16], we consider the nonlinear Schrödinger equation with cubic (focusing or defocusing) nonlinearity on the multidimensional torus. For special small initial data containing only five modes, we exhibit a countable set of time layers in which arbitrarily large modes are created. The proof relies on a reduction to multiphase weakly nonlinear geometric optics, and on the study of a particular two-dimensional discrete dynamical system.

5.24. **A Nekhoroshev type theorem for the nonlinear Schrödinger equation on the d-dimensional torus**

**Participant:** Erwan Faou.

In [49] we prove a Nekhoroshev type theorem for the nonlinear Schrödinger equation

\[ iu_t = -\Delta u + V u + \partial \pi g(u, \pi), \quad x \in T^d, \]

where \( V \) is a typical smooth Fourier multiplier and \( g \) is analytic in both variables. More precisely we prove that if the initial datum is analytic in a strip of width \( \rho > 0 \) whose norm on this strip is equal to \( \epsilon \) then, if \( \epsilon \) is small enough, the solution of the nonlinear Schrödinger equation above remains analytic in a strip of width \( \rho/2 \), with norm bounded on this strip by \( C \epsilon \) over a very long time interval of order \( \epsilon^{-\alpha} \ln \epsilon^\beta \), where \( 0 < \beta < 1 \) is arbitrary and \( C > 0 \) and \( \alpha > 0 \) are positive constants depending on \( \beta \) and \( \rho \).
5.25. Sobolev stability of plane wave solutions to the cubic nonlinear Schrödinger equation on a torus

**Participant:** Erwan Faou.

In [48], it is shown that plane wave solutions to the cubic nonlinear Schrödinger equation on a torus behave orbitally stable under generic perturbations of the initial data that are small in a high-order Sobolev norm, over long times that extend to arbitrary negative powers of the smallness parameter. The perturbation stays small in the same Sobolev norm over such long times. The proof uses a Hamiltonian reduction and transformation and, alternatively, Birkhoff normal forms or modulated Fourier expansions in time.

5.26. Approximate travelling wave solutions to the 2D Euler equation on the torus

**Participants:** Erwan Faou, Nicolas Crouseilles.

In [43], we consider the two-dimensional Euler equation with periodic boundary conditions. We construct approximate solutions of this equation made of localized travelling profiles with compact support propagating over a stationary state depending on only one variable. The direction or propagation is orthogonal to this variable, and the support is concentrated around flat points of the stationary state. Under regularity assumptions, we prove that the approximation error can be made exponentially small with respect to the width of the support of the travelling wave. We illustrate this result by numerical simulations.

5.27. Markov chains competing for transitions: applications to large-scale distributed systems

**Participant:** François Castella.

We consider in [12] the behavior of a stochastic system composed of several identically distributed, but non-independent, discrete-time absorbing Markov chains competing at each instant for a transition. The competition consists in determining at each instant, using a given probability distribution, the only Markov chain allowed to make a transition. We analyze the first time at which one of the Markov chains reaches its absorbing state. We obtain its distribution and its expectation and we propose an algorithm to compute these quantities. We also exhibit the asymptotic behavior of the system when the number of Markov chains goes to infinity. Actually, this problem comes from the analysis of large-scale distributed systems and we show how our results apply to this domain.

5.28. Analysis of a large number of Markov chains competing for transitions

**Participant:** François Castella.

This text [41] generalizes the previous one [12] in the following sense.

In the situation on the previous article, we analyze the first time at which one of the Markov chains reaches its absorbing state. When the number of Markov chains goes to infinity, we analyze the asymptotic behavior of the system for an arbitrary probability mass function governing the competition. We give conditions for the existence of the asymptotic distribution and we show how these results apply to cluster-based distributed storage when the competition is handled using a geometric distribution.

5.29. Splitting methods with complex coefficients for some classes of evolution equations

**Participant:** Philippe Chartier.
We are concerned in [13] with the numerical solution obtained by splitting methods of certain parabolic partial differential equations. Splitting schemes of order higher than two with real coefficients necessarily involve negative coefficients. In a previous paper, Castella et al. demonstrated the possibility to overcome this second-order barrier by considering splitting methods with complex-valued coefficients and built up methods of orders 3 to 14. In this paper, we reconsider the technique employed therein and show that it is inherently bound to order 14 and largely sub-optimal with respect to error constants. As an alternative, we solve directly the algebraic equations arising from the order conditions and construct several methods of orders 4, 6, 8 and 16 that are the most accurate ones available at present time.

5.30. Higher-order averaging, formal series and numerical integration

Participant: Philippe Chartier.

The paper [42] considers non-autonomous oscillatory systems of ordinary differential equations with \( d = 1 \) non-resonant constant frequencies. Formal series like those used nowadays to analyze the properties of numerical integrators are employed to construct higher-order averaged systems and the required changes of variables. With the new approach, the averaged system and the change of variables consist of vector-valued functions that may be written down immediately and scalar coefficients that are universal in the sense that they do not depend on the specific system being averaged and may therefore be computed once and for all. The new method may be applied to obtain a variety of averaged systems. In particular we study the quasi-stroboscopic averaged system characterized by the property that the true oscillatory solution and the averaged solution coincide at the initial time. We show that quasi-stroboscopic averaging is a geometric procedure because it is independent of the particular choice of co-ordinates used to write the given system. As a consequence, quasi-stroboscopic averaging of a canonical Hamiltonian (resp. of a divergence-free) system results in a canonical (resp. in a divergence-free) averaged system. We also study the averaging of a family of near-integrable systems where our approach may be used to construct explicitly formal first integrals for both the given system and its quasi-stroboscopic averaged version. As an application we construct three first integrals of a system that arises as a nonlinear perturbation of coupled harmonic oscillators with one slow frequency and four resonant fast frequencies.

The stroboscopic averaging method (SAM) is a technique for the integration of highly oscillatory differential systems \( \dot{y} = f(y, t) \) with a single high frequency. The method may be seen as a purely numerical way of implementing the analytical technique of stroboscopic averaging which constructs an averaged differential system \( \dot{Y} = F(Y) \) whose solutions \( Y \) interpolate the sought highly oscillatory solutions \( y \). SAM integrates numerically the averaged system without using the analytic expression of \( F \); all information on \( F \) required by the algorithm is gathered on the fly by numerically integrating the originally given system in small time windows. SAM may be easily implemented in combination with standard software and may be applied with variable step sizes. Furthermore it may also be used successfully to integrate oscillatory DAEs. The paper [15] provides an analytic and experimental study of SAM and two related techniques: the LISP algorithms of Kirchgraber and multirevolution methods.
5. New Results

5.1. Numerical probability

5.1.1. Simulation of stochastic processes

Participants: Aurélien Alfonsi, Benjamin Jourdain, Arturo Kohatsu-Higa [(University of Osaka)], Abdelkodoudsse Ahdida, Stefanp De Marco.

5.1.1.1. Pathwise convergence of the Euler scheme

A. Alfonsi, B. Jourdain and A. Kohatsu-Higa are studying the convergence for the Wasserstein distance of the Euler scheme towards the limit diffusion. This would be a way to analyze the weak pathwise convergence of this discretization scheme. They have obtained some promising results in the one-dimensional case.

5.1.2. Multi-dimensional models and correlation issues

A. Alfonsi and his PhD Student A. Ahdida have submitted a paper on the simulation of Wishart processes. They have introduced a new family of stochastic differential equations that are defined in the space of correlation matrices and provided high order discretization schemes for such processes [51]. They are currently trying to use this type of matrix valued processes to model the dependence between assets. In particular, they would like to calibrate Index options data. This work is still in progress. The thesis of A. Ahdida has been defended on December 1st [12].

The PhD student Lokmane Abbas Turki has worked on numerical methods for American option pricing based on Malliavin calculus and parallel implementation. He has submitted a paper (co-authored with B. Lapeyre) [50]. He is now working on the dependence of option prices with respect to correlations between stocks in multi-dimensional models.

A. Alfonsi and S. De Marco (postdoc CERMICS) have studied how some option prices (such as spread options) are modified when the correlation between stocks is increased.

5.1.3. Stochastic volatility models

Exotic options and stochastic volatility models is the subject of Sidi Mohamed OULD ALY’s thesis, defended on June 16th, 2011. Sidi-Mohamed has results on the effective computation of option prices in a stochastic volatility model, in the context of variance swap modelling. He has worked out a new model, in the spirit of Bergomi’s approach. This model has remarkable features in terms of tractability and calibration. S.M. Ould Aly has developed numerical methods and an original variance reduction method for models with a log normal volatility. He also has results on the monotony of option prices with respect to the correlation between the stock price and the volatility in the Heston model. Sidi Mohamed has submitted three papers.

5.2. American options

Participants: Benjamin Jourdain, Maxence Jeunesse, Damien Lamberton, Ayech Bouselmi.

5.2.1. American put option with discrete dividends.

B. Jourdain and M. Jeunesse are interested in the regularity of the optimal exercise boundary for the American Put option when the underlying asset pays discrete dividends at known times during the lifetime of the option. The dividend amounts are deterministic functions of the asset prices just before the dividend dates. B. Jourdain and M. Jeunesse have proved continuity of the exercise boundary and smooth contact for the value function under general assumptions on the dividend functions.
5.2.2. **American options in exponential Lévy models**

D. Lamberton and his PhD student A. Bouselmi are working on American options within multi-dimensional exponential Lévy models. They also have preliminary results on the asymptotic behaviour of the exercise boundary of the American put near maturity in the one dimensional case when the limit is strictly smaller than the strike price.

5.3. **Financial issues modelling**

**Participants:** Aurélien Alfonsi, Céline Labart, Jérôme Lelong, J. Acevedo.

5.3.1. **Credit risk modelling**

A. Alfonsi, C. Labart and J. lelong are studying loss models called “stochastic local intensity models” that have been proposed in the literature. First, they are interested in proving mathematically that these models are well posed (it exists and has a unique solution). Second, they aim to provide numerical tools to sample such dynamics.

5.3.2. **Limit order markets**

A. Alfonsi has an active collaboration with A. Schied (Mannheim University) on limit order book models. A. Schied has visited the CERMICS two weeks in February 2011 and one week on December. They are currently studying some type of non-Markovian resilience for the limit-order book for which they are able to get the optimal execution strategy in a closed form. Moreover A. Alfonsi and his PhD student J. Acevedo study impact models for which the limit order book shape evolves along the time.

5.3.3. **Control of systemic risk**

**Participants:** Agnès Sulem, J.Ph. Chancelier, Andreea Minca.

We are interested in contagion modeling and systemic risk in financial networks. We aim to contribute in particular to the domain of control of such systems in order to reduce the systemic risk. We model the propagation of distress in financial systems as an epidemic on a random graph in which the nodes represent financial institutions and edges the exposure between them. Cascade dynamics may be reduced to the evolution of a multi-dimensional markov chain that corresponds to a sequential discovery of exposures and determines the size of contagion. We study the optimal intervention strategy by a lender of last resort who wants to minimize the size of contagion under budget constraints.

5.4. **Stochastic control of jump processes and stochastic differential games**

**Participants:** Agnès Sulem [in collaboration with B. Øksendal (Oslo University) and T. Zhang (Manchester University)], John Joseph Absalom Hosking.

5.4.1. **Stochastic control under model uncertainty**

In [58], we study optimal stochastic control problems with jumps under model uncertainty. We rewrite such problems as (zero-sum) stochastic differential games of forward-backward stochastic differential equations. We prove general stochastic maximum principles for such games, both in the zero-sum case (finding conditions for saddle points) and for the non-zero sum games (finding conditions for Nash equilibria). We then apply these results to study optimal portfolio and consumption problems under model uncertainty. We combine the optimality conditions given by the stochastic maximum principles with Malliavin calculus to obtain a set of equations which determine the optimal strategies.

In [45], we consider some robust optimal portfolio problems for markets modeled by (possibly non-Markovian) Itô–Lévy processes. Mathematically the situation can be described as a stochastic differential game, where one of the players (the agent) is trying to find the portfolio which maximizes the utility of her terminal wealth, while the other player (“the market”) is controlling some of the unknown parameters of the market (e.g. the underlying probability measure, representing a model uncertainty problem) and is trying to minimize this maximal utility of the agent. This leads to a worst case scenario control problem for the agent.
In the Markovian case such problems can be studied using the Hamilton-Jacobi-Bellman-Isaacs (HJBI) equation, but these methods do not work in the non-Markovian case. We approach the problem by transforming it to a stochastic differential game for backward stochastic differential equations (BSDE game). Using comparison theorems for BSDEs with jumps we arrive at criteria for the solution of such games, in the form of a kind of non-Markovian analogue of the HJBI equation. The results are illustrated by examples.

5.4.2. Singular stochastic control

In [59], A. Sulem and B. Øksendal study partial information, possibly non-Markovian, singular stochastic control of Itô–Lévy processes and obtain general maximum principles. The results are used to find connections between singular stochastic control, reflected BSDEs and optimal stopping in the partial information case. As an application we give an explicit solution to a class of optimal stopping problems with finite horizon and partial information. Singular control of SPDEs.

In [57], A. Sulem, B. Øksendal and T. Zhang study general singular control problems for random fields given by a stochastic partial differential equation (SPDE). They show that under some conditions the optimal singular control can be identified with the solution of a coupled system of SPDE and a kind of reflected backward SPDE (RBSPDE). They also establish existence and uniqueness of solutions of RBSPDEs.

5.4.3. Optimal control with delay

In [44], we study optimal control problems for (time-)delayed stochastic differential equations with jumps. We establish sufficient and necessary stochastic stochastic maximum principles for an optimal control of such systems. The associated adjoint processes are shown to satisfy a (time-) advanced backward stochastic differential equation (ABSDE). Several results on existence and uniqueness of such ABSDEs are shown. The results are illustrated by an application to optimal consumption from a cash flow with delay.

In [48], we will prove a sufficient necessary stochastic maximum principles for the optimal control of SPDEs with delay and study associated time-advanced backward stochastic partial differential equations.

5.4.4. Stochastic differential games

In [55], J. Hosking has constructed a stochastic maximum principle (SMP) which provides necessary conditions for the existence of Nash equilibria in a certain form of N-agent stochastic differential game (SDG) of a mean-field type. The information structure considered for the SDG is of a possible asymmetric and partial type. To prove our SMP we use a spike-variation approach with adjoint representation techniques, analogous to that of S. Peng in the optimal stochastic control context. In our proof we apply adjoint representation procedures at three points. The first-order adjoint processes are defined as solutions to certain mean-field backward stochastic differential equations (BSDEs), and second-order adjoint processes of a fist type are defined as solutions to certain BSDEs. Second order adjoint processes of a second type are defined as solutions of backward stochastic equations of a type that we introduce in this paper, and which we term conditional mean-field BSDEs. From the resulting representations, we show that the terms relating to these second-order adjoint processes of the second type are of an order such that they do not appear in our final SMP equations.

5.5. Risk measures, BSDEs with jumps and nonlinear expectations

Participants: Agnès Sulem, Marie-Claire Quenez, Z. Chen [Shandong University].

In the Brownian case, links between dynamic risk measures and Backward Stochastic Differential Equations (BSDEs) have been established. A. Sulem and M.-C. Quenez are exploring these links in the case of stochastic processes with jumps. They have extended some comparison theorems for BSDEs with jumps, and provided a representation theorem of convex dynamic risk measures induced by BSDEs with jumps. They study optimal stopping problems for (non necessarily) convex dynamic risk measures induced by BSDEs with jumps and establish their connections with Reflected BSDEs with jumps. They also study the case of model ambiguity and its relation with mixed control/optimal stopping problems.
There are two classes of nonlinear expectations, one is the Choquet expectation given by Choquet (1955), the other is the Peng’s \( g \)-expectation given by Peng (1997) via backward differential equations (BSDE). Recently, Peng raised the following question: can a \( g \)-expectation be represented by a Choquet expectation? In [26], A. Sulem and Z. Chen provide a necessary and sufficient condition on \( g \)-expectations under which Peng’s \( g \)-expectation can be represented by a Choquet expectation for some random variables (Markov processes). It is well known that Choquet expectation and \( g \)-expectation (also BSDE) have been used extensively in the pricing of options in finance and insurance. Our result also addresses the following open question: given a BSDE (\( g \)-expectation), is there a Choquet expectation operator such that both BSDE pricing and Choquet pricing coincide for all European options? Furthermore, the famous Feynman-Kac formula shows that the solutions of a class of (linear) partial differential equations (PDE) can be represented by (linear) mathematical expectations. As an application of our result, we obtain a necessary and sufficient condition under which the solutions of a class of nonlinear PDE can be represented by nonlinear Choquet expectations [26].

5.6. Malliavin calculus and applications

5.6.1. Lower bounds for the density of functionals on the Wiener space

In collaboration with B. Fernandez and A. Meda from the University of Mexico, V. Bally gave a lower bound for general Itô processes to remain in a tube around a given curve. This is done under some ellipticity assumption in [21]. Now, with L. Caramellino (University Tor Vergata, Rome) he investigates the case of a diffusion processes which satisfies the Hörmander condition.

5.6.2. Malliavin Calculus for Poisson Point Processes and applications

V. Bally and E. Clément (Université Paris-Est Marne la Vallée) study the density of the law of the solution of a stochastic equation with jumps, which has discontinuous coefficients [18], [19]. Moreover, with N. Fournier (university of Creteil), V. Bally obtained results on the smoothness of the law of a bidimensional Bolzman equation [22].

5.6.3. Riesz transform and regularity of the law of a random variable

The idea of using the Riesz transform in order to study the regularity of the law of a random variable appears in former works of P. Malliavin and A. Thalmaier. In collaboration with L. Caramellino (University Tor Vergata, Rome) we gave regularity results using this tool [17].
6. New Results

6.1. Théorie spectrale max-plus et géométrie métrique/Max-plus spectral theory and metric geometry

6.1.1. Introduction

Participants: Marianne Akian, Stéphane Gaubert, Cormac Walsh.

Étant donné un noyau \( a : S \times S \rightarrow \mathbb{R} \cup \{-\infty\} \), on peut lui associer le problème spectral max-plus

\[
\sup_{y \in S} a(x, y) + u(y) = \lambda + u(x), \quad \forall x \in S,
\]

(3)
dans lequel on cherche le vecteur propre \( u : S \rightarrow \mathbb{R} \cup \{-\infty\} \) et la valeur propre correspondante \( \lambda \in \mathbb{R} \cup \{-\infty\} \). Comme nous l’avons rappelé dans les § 3.2 et 3.3 , le problème spectral ( 9 ) intervient en contrôle ergodique: l’ensemble \( S \) est l’espace des états, et l’application \( a(x, y) \) fournit le gain associé à la transition \( x \to y \). Le cas où \( S \) est fini est classique, l’on a alors un résultat précis de répresentation de l’espace propre, à l’aide d’un certain graphe, dit graphe critique. Des résultats existent également lorsque \( S \) est compact et que le noyau vérifie certaines propriétés de régularité.

Dans [ 79 ], nous avons considéré le cas où \( S \) est non compact. Lorsque \( \lambda = 0 \), l’espace propre est analogue à l’espace des fonctions harmoniques défini en théorie (classique ou probabiliste) du potentiel. En introduisant l’analogue max-plus de la frontière de Martin, nous avons obtenu un analogue de la formule de représentation de Poisson des fonctions harmoniques : toute solution \( u \) de ( 9 ) peut être représentée sous la forme :

\[
u = \sup_{w \in \mathcal{M}_m} w + \mu_u(w)
\]

(4)
où \( \mathcal{M}_m \subset (\mathbb{R} \cup \{-\infty\})^S \) est l’analogue max-plus de la frontière de Martin minimale (l’ensemble des fonctions harmoniques extrémales normalisées), et où \( \mu_u \) joue le rôle de la mesure spectrale. Nous avons montré aussi que les éléments de l’espace de Martin minimal peuvent être caractérisés comme les limites de “quasi-géodésiques”. La frontière de Martin max-plus généralise dans une certaine mesure la frontière d’un espace métrique construite à partir des horo-fonctions (fonctions de Busemann généralisées), ou horofrontière. Ces résultats inspirent les travaux des sections suivantes, qui portent sur des cas remarquables d’espaces métriques (§ 6.1.4 ) ou sur des applications en théorie des jeux (§ 6.1.2 ).

English version

Let the kernel \( a : S \times S \rightarrow \mathbb{R} \cup \{-\infty\} \) be given. One may associate the max-plus spectral equation ( 9 ), where the eigenvector \( u : S \rightarrow \mathbb{R} \cup \{-\infty\} \) and the eigenvalue \( \lambda \in \mathbb{R} \cup \{-\infty\} \) are unknown. As we recalled in § 3.2 and refinotone, this spectral problem arises in ergodic optimal control: the set \( S \) is the state space, and the map \( a(x, y) \) is the transition reward. The case when \( S \) is finite is classical, a precise spectral theorem is known, with a characterisation of the eigenspace in terms of a critical graph. Some results have been shown when \( S \) is compact, assuming that the kernel \( a \) satisfies some regularity properties.
In [79], we considered the case where $S$ is non-compact. When $\lambda = 0$, the eigenspace is analogous to the set of harmonic functions defined in classical or probabilistic potential theory. By introducing a max-plus analogue of the classical Martin boundary, we obtained an analogue of the Poisson representation of harmonic functions, showing that any solution $u$ of (9) may be represented as in (10) where $\mathcal{M}_m \subset (\mathbb{R} \cup \{-\infty\})^S$ is a max-plus analogue of the minimal Martin boundary (the set of normalised extremal harmonic functions), and $\mu_u$ plays the role of the spectral measure. We also showed that the elements of the minimal Martin boundary can be characterised as limits of certain “almost-geodesics”. The max-plus Martin boundary generalises to some extent the boundary of metric spaces defined in terms of horofunctions (generalised Busemann functions), or horoboundary. These results have inspired the work of the next sections, which deal either with remarkable examples of metric spaces (§ 6.1.4) or applications to zero-sum games (§ 6.1.2).

6.1.2. Une caractérisation maximin du taux de fuite d’applications nonexpansives/A maximin characterization of the escape rate of nonexpansive mappings

Participants: Stéphane Gaubert, Guillaume Vigeral.

Le problème de l’existence du gain moyen par unité de temps pour des jeux répétés à somme nulle conduit à étudier la limite $f^k(v)/k$ lorsque l’horizon $k$ tend vers l’infini, où $f$ (l’opérateur de programmation dynamique) est une application contractante au sens large sur un espace de Banach, voir § 3.3. La limite peut ne pas exister, mais un résultat de Kohlberg et Neyman montre qu’il existe toujours une forme linéaire $\phi$ de norme 1 telle que la limite de $\phi(f^k(x)/k)$ existe lorsque $k \to \infty$ et coïncide avec le “taux de fuite” $\lim_{k \to \infty} \|f^k(x)/k\|$. Dans [29], nous avons généralisé ce résultat au cas où $f$ est une application nonexpansive (contractante au sens large) sur un espace métrique vérifiant une forme affaiblie de l’hypothèse de courbure nonpositive au sens de Busemann. La forme linéaire $\phi$ est alors remplacée par une horofonction, et l’on obtient une caractérisation de type “maximin” du taux de fuite, qui étend la formule de Collatz-Wielandt en théorie de Perron-Frobenius (le vecteur propre apparaissant dans cette formule correspond à une famille d’horoboules invariantes par la dynamique). Ceci est motivé par les problèmes de contrôle ou de jeux quadratiques, dans lesquels l’espace métrique est le cône des matrices définies positives muni de sa métrique Riemannienne invariante ou de la métrique de Thompson.

English version

The problem of the existence of the mean payoff per time unit for repeated games leads to studying the existence of the limit of $f^k(v)/k$, where $f$ is a nonexpansive map (the dynamic programming operator) acting on a Banach space, see § 3.3 for more background. The limit may not exist, but a result of Kohlberg et Neyman shows that there is always a norm one linear form $\phi$ such that the limit of $\phi(f^k(x)/k)$ exists and coincides with the limit of $\|f^k(x)/k\|$ as $k$ tends to infinity (the escape rate). In [29], we extend this result to the case of a nonexpansive map defined on a metric space satisfying a mild form of Busemann nonpositive curvature condition. Then, the linear form $\phi$ is replaced by an horofunction, and we obtain a maximin type characterization of the escape rate, which extends the Collatz-Wielandt formula in Perron-Frobenius theory. This is motivated by the study of quadratic optimal control and game problems, in which the metric space is the cone of positive semi-definite matrices equipped with the Riemannian invariant metric or with Thompson metric.

6.1.3. Isométries de la géométrie de Hilbert/Isometries of the Hilbert geometry

Participants: Cormac Walsh, Bas Lemmens [Kent University, UK].

L’un des intérêts de l’horo-frontière est de renseigner sur le groupe des isométries d’un espace métrique. En effet, ce groupe agit naturellement sur l’horo-frontière, et cette action peut parfois être mieux comprise que l’action du groupe sur l’espace d’origine.

Nous étudions le groupe des isométries pour la métrique de Hilbert. De La Harpe [188] a donné plusieurs conjectures relatives à ce groupe. Nous conjecturons que le groupe des isométries est exactement le groupe des transformations linéaires projectives à moins que le domaine ne soit une coupe d’un cône symétrique non-Lorentzien. Nous avons démontré cette conjecture lorsque le domaine est un polytope [32].
Dans le cas général, on prouve, en utilisant les horo-fonctions, que si il existe une bijection entre deux cônes homogène de degré $-1$, antitone, et d’inverse antitone, ces deux cônes sont symétriques. Nous essayons maintenant de montrer que toute isométrie de Hilbert sur un domaine convexe est la version projective d’un automorphisme linéaire du cône sur le domaine, ou d’une bijection du cône, homogène de degré $-1$, qui est antitone et d’inverse antitone. Ce résultat permettrait de compléter la preuve de la conjecture proposée plus haut.

**English version**

One use for the horofunction boundary is to study the group of isometries of a metric space. This is because this group has a well defined action on the horoboundary and it is likely that in many cases this action will be easier to understand than the action on the space itself.

We have been investigating the isometries of the Hilbert geometry. De La Harpe [188] has previously made several conjectures about the isometry group of this space. We conjecture that the isometry group is exactly the group of projective linear transformations unless the domain on which the geometry is defined is a cross section of a non-Lorentzian symmetric cone. We have previously proved that this conjecture is true in the case of a polytope domain [32].

In the general case, we can now prove, using horofunctions, that if a bijection between cones is homogeneous of degree $-1$, order inverting, and has an order inverting inverse, then both cones are symmetric. We are working on showing that every Hilbert isometry on a convex domain arises by considering projectively either a linear automorphism on the cone over the domain, or a homogeneous $-1$, order inverting bijection on this cone with order inverting inverse. Establishing this result would complete our proof of the above conjecture.

**6.1.4. Espace de Teichmüller/Teichmüller space**

**Participant:** Cormac Walsh.

L’espace de Teichmüller d’une surface est un espace métrique composé des structures conformes de cette surface. On peut le voir comme l’ensemble des classes d’équivalence des métriques riemanniennes de cette surface, où deux métriques sont équivalentes si il existe une application conforme homotope à l’identité qui envoie l’une des métriques sur l’autre.


Néanmoins, la métrique la plus utilisée sur l’espace de Teichmüller est la métrique de Teichmüller. L’horo-frontière de cet espace métrique n’est autre que la frontière déjà introduite dans la littérature sous le nom de frontière de Gardiner–Masur. Nous étudions cette frontière, en particulier nous donnons explicitement ses points de Busemann.

Par la suite, nous avons l’intention d’utiliser cette propriété afin d’étudier les sous-groupes du groupe modulaire, qui est le groupe des isométries de la métrique de Teichmüller.

**English version**

An interesting metric space is the Teichmüller space of a surface. This is the space of conformal structures on the surface. One may think of it as the space of equivalence classes of Riemannian metrics on the surface, where two such metrics are regarded as being equivalent if there is a conformal map on the surface taking one to the other that is homotopic to the identity.

There are several natural metrics on Teichmüller space. Previously, we have worked with Thurston’s stretch metric and have shown [70] that the horofunction boundary with this metric is just the usual Thurston boundary.
However, the most commonly used metric on Teichmüller space is Teichmüller’s metric. The horofunction boundary of this metric space turns out to be the same as a previously defined boundary, called the Gardiner–Masur boundary. We have been investigating this boundary, in particular we have managed to work out explicitly its Busemann points.

In future work, we intend to apply this knowledge to study subgroups of the mapping class group, which is the isometry group of the Teichmüller metric.

6.2. Algèbre linéaire max-plus et convexité abstraite/Max-plus linear algebra and abstract convex analysis

6.2.1. Convexité max-plus ou tropicale/Max-plus or tropical convexity

Participants: Xavier Allamigeon, Stéphane Gaubert, Eric Goubault [CEA], Ricardo Katz [Conicet, Argentine].

On étudie les analogues max-plus ou tropicaux des ensembles convexes. Ceux-ci sont utiles en particulier pour représenter de manière effective les ensembles d’états accessibles de systèmes à événements discrets [9], ils sont aussi apparus récemment en géométrie tropicale, dans toute une série de travaux à la suite de Sturmfels et Develin [114]. Les polyèdres max-plus peuvent aussi être vus comme des limites de déformations de polyèdres classiques, sur lesquels ils donnent un éclairage de nature combinatoire. Toutes ces motivations ont inspiré la recherche d’analogues des résultats fondamentaux d’analyse convexe classique: séparation, projection, points extrémaux, à la suite en particulier de [8].


On en déduit un analogue tropical de la méthode de la double description [57] (méthode très utilisée sur les polyèdres classiques, et dûe à Motzkin et al. [158]). Cet algorithme permet de calculer les sommets d’un polyèdre défini de façon externe (intersection de demi-espaces ou d’hyperplans tropicaux). Grâce au critère combinatoire précédent, l’algorithme améliore de plusieurs ordres de grandeur les techniques connues jusqu’alors. Ceci est confirmé par de nombreuses expérimentations. Ce travail est motivé par des applications à l’analyse statique [82] et aux systèmes à événements discrets [116], dans lesquelles la manipulation de tels polyèdres est le goulot d’étranglement.

Dans un travail de X. Allamigeon, S. Gaubert, et R. Katz [57], on étend le théorème de McMullen au cas tropical: ce dernier caractérise le nombre maximal de points extrémaux d’un polyèdre, en fonction du nombre d’inégalités qui le définissent et de sa dimension. Nous montrons que la même borne est valide dans le cas tropical (à une modification triviale près). Cependant, le calcul de la borne optimale est encore ouvert dans ce cas.

Dans un travail de S. Gaubert et R. Katz [26], on étudie la représentation d’un polyèdre tropical comme intersection de demi-espaces, ou si l’on préfère, comme conjonction d’inégalités affines. Nous donnons notamment un contre-exemple, montrant les inconvénients de la représentation en termes de demi-espaces minimaux proposée précédemment dans la littérature tropicale.

We study the max-plus or tropical analogues of convex sets. These have been used in particular to represent effectively the accessible sets of certain discrete event systems [9]. They also appeared in tropical geometry, following the work of Sturmfels and Develin [114]. Max-plus polyhedra can be thought of as limits of deformations of classical polyhedra, on which they give a combinatorial insight. These motivations have inspired the investigation of analogues of basic results of classical convex analysis: separation, projection, representation by extreme points, following [8].

In a work of X. Allamigeon, S. Gaubert, and E. Goubault [57], we introduce a combinatorial criterion for the characterization of the vertices of tropically convex polyhedra. It is expressed in terms of directed hypergraphs and their strongly connected components. This criterion can be verified in almost linear time in the size of the hypergraph.

This allows to develop a tropical analogue of the double description method [57] (this method is widely used for classical convex polyhedra, and is due to Motzkin et al. [158]). This algorithm is able to determine all the vertices of a polyhedron defined externally (intersection of tropical half-spaces of hyperplanes). Thanks to the combinatorial criterion mentioned above, the algorithm improves the existing methods by several orders of magnitude. This is confirmed by several experiments. This is motivated by applications to static analysis [82] and discrete event systems [116], in which computing such polyhedra turns out to be the bottleneck.

In a work of X. Allamigeon, S. Gaubert, and R. Katz [19], we extend the McMullen upper bound theorem to the tropical case. This theorem characterizes the maximal number of extreme points of a polyhedron, as a function of the number of inequalities defining it, and of the dimension. We show that the same bound is valid in the tropical case (up to a trivial modification). However, computing the optimal bound is an open problem in this case.

In a work of S. Gaubert and R. Katz [26], we study the representation of a tropical polyhedron as an intersection of half-spaces. We give in particular a counter example, showing some inconvenients of the representation in terms of minimal half-spaces proposed previously in the tropical literature.

It is well-known that a tropical polyhedron can be represented as the convex hull of a minimal set of points and rays, provided by its vertices and extreme rays [125]. In an ongoing work of X. Allamigeon and R. Katz, partly done during the visit of R. Katz at INRIA (July 2011), the dual problem of characterizing the minimal representations by half-spaces is studied. We show that a tropical polyhedron admits essentially a unique minimal external representation by half-spaces, provided that their apices belong to the polyhedron.

We prove that the apices of these half-spaces correspond to certain vertices of the tropical complex introduced by Develin and Sturmfels [114]. We also establish a combinatorial criterion allowing to eliminate redundant half-spaces using directed hypergraphs.

### 6.2.2. Convexes max-plus et jeux avec paiements ergodiques/Max-plus convex sets and mean payoff games

**Participants:** Marianne Akian, Xavier Allamigeon, Stéphane Gaubert, Alexander Guterman [Moscow State University], Ricardo Katz [Conicet, Argentine], Sergei Sergeev.

Dans un travail d’Akian, Gaubert et Guterman [16], on montre un résultat d’équivalence entre les jeux ergodiques à somme nulle et les systèmes d’inégalités max-plus linéaires: décider la non-vacuité d’un polyèdre tropical est équivalent à vérifier si un jeu déterministe à somme nulle a un paiement moyen par unité de temps positif ou nul. Plus généralement, la même question pour un jeu stochastique à somme nulle est équivalente à vérifier si un convexe tropical (non-polyédral, i.e., défini par un système infini d’inégalités) est vide. Ces résultats sont démontrés à l’aide de techniques de théorie de Perron-Frobenius non-linéaire. Ils sont ensuite appliqués à l’étude de l’indépendance linéaire dans le semi-anneau tropical.

Le résultat de [16] a eu plusieurs retombées.
D’une part, dans un travail d’Allamigeon, Gaubert, et Katz [20], on établit un analogue tropical du théorème de Farkas: on montre que décider si une inégalité max-plus linéaire est une conséquence logique d’une famille de telles inégalités est également équivalent à un problème de jeu ergodique. Le travail [20] comprend aussi une description des “faces” (ou plus précisément, des points extrêmes du polaire) d’un polyèdre tropical en termes de transversaux minimaux dans un hypergraphe.

D’autre part, dans un travail de Gaubert et Sergeev [127], on réduit le problème spectral tropical de type faisceaux, \( Ax = \lambda Bx \), à un jeu paramétrique (ce qui permet de calculer le spectre en temps pseudo-polyomial).

Enfin, dans un travail de Gaubert, Katz, et Sergeev [27], on développe un algorithme de programmation linéaire tropicale (pseudo-polynomial) basé sur cette correspondance avec les jeux répétés. Allamigeon et Sergeev ont développé récemment un logiciel prototype en ocaml afin d’expérimenter cet algorithme. Ce logiciel inclus en particulier, en guise d’oracles résolvant des problèmes de jeux répétés, l’algorithme d’itération sur les politiques de Gaubert-Gunawardena [124] dans sa version de [115], la modification de cet algorithme due à Chaloupka [98], ainsi que l’algorithme d’itérations sur les politiques de Björklund-Vorobyov [91].

French version

In [16], we show the equivalence mean payoff games and max-plus linear inequalities: testing whether a tropical polyhedron is non-empty is equivalent to checking whether a mean payoff deterministic game is winning. More generally, checking whether a mean payoff stochastic game is winning is equivalent to checking the non-emptiness of a tropical convex set defined by an infinite family of inequalities. These results are established using techniques of non-linear Perron-Frobenius theory. Then, they are applied to the study of linear independence over the tropical semiring.

The equivalence established in [16] had several consequences.

First, a work of Allamigeon, Gaubert, and Katz [20] yields a tropical analogue of Farkas theorem: we show that deciding whether a max-plus linear inequality follows from a family of such inequalities is also equivalent to solving a mean payoff game. Moreover, the work [20] comprises a characterization of the “faces” (more precisely, the extreme points of the polar) of a tropical polyhedra in terms of minimal transversals of a hypergraph.

Next, in a work of Gaubert and Sergeev [127], the tropical spectral problem for matrix pencils, \( Ax = \lambda Bx \), is reduced to a parametric game (which allows one to compute the spectrum in pseudo-polynomial time).

Finally, in a work of Gaubert, Katz, and Sergeev [27], a (pseudo-polynomial) tropical linear programming algorithm is developed, based on the same correspondence with mean payoff games. Allamigeon and Sergeev developed recently an ocaml prototype in order to experiment this method. This prototype includes in particular mean payoff game solvers, namely the version of [115] of the policy iteration algorithm of Gaubert-Gunawardena [124], the modification of this algorithm by Chaloupka [98], as well as the policy iteration algorithm of Björklund-Vorobyov [91].

6.2.3. Meilleure approximation par des semi-modules max-plus pour la métrique projective de
Hilbert/Best approximation in Hilbert’s projective metric by max-plus semimodules

Participants: Marianne Akian, Stéphane Gaubert, Viorel Nitica [West Chester University (US) and IMAR (Bucharest, Romania)], Ivan Singer [IMAR (Bucharest, Romania)].

Nous étudions les projecteurs sur des espaces linéaires max-plus, ainsi que les demi-espaces max-plus séparants. Dans [18], nous obtenons de nouvelles propriétés concernant ces objets, ce qui nous permet de déduire une formule explicite de la distance d’un point à un demi-espace max-plus pour la métrique projective de Hilbert, ainsi qu’une caractérisation de l’ensemble des points minimisants de cette distance. Nous obtenons aussi un algorithme de type projection cyclique permettant de résoudre des systèmes d’inégalités linéaires max-plus. Ce travail est effectué dans le cadre d’un projet LEA Math mode.
We are studying projectors on max-plus linear spaces, as well as separating half-spaces over the max-plus semiring. In [18], we establish new results, and derive an explicit formula for the distance in Hilbert’s projective metric between a point and a half-space, as well as explicit descriptions of the set of minimizers of this distance. We also obtain, as a consequence of the previous results, a cyclic projection type algorithm to solve systems of max-plus linear inequalities. This work is carried out as part of a LEA Math-mode project.

6.2.4. Miscellanées en algèbre linéaire max-plus/Topics in max-plus linear algebra

 Participant: Sergei Sergeev.

During his one year post-doctoral stay, S. Sergeev collaborated with team members on tropical convexity and mean payoff game problems (§6.2.2 supra). In addition, the following research has been developed with several international coauthors.

- **Z-matrix equations and weakly stable matrices.** Sergeev continued his joint research with P. Butkovič and H. Schneider on traditional topics of max-plus linear algebra. They described solution set to Z-matrix equations \( \lambda x = Ax \oplus b \), comparing the results with nonnegative matrix setting (old works of H. Schneider with coauthors) and extending them to linear algebra over more general semirings. Weakly stable matrices are such that the set of vectors \( x \) whose orbit \( A^kx \) converges to an eigenvector, is exactly the eigenvector cone. The weak stability was characterized in terms of relations between spectral classes of \( A \) and the critical graph in each strongly connected component.

- **CSR expansions.** Sergeev continued his research on CSR expansions of matrix powers in max-plus algebra, see the accepted paper with H. Schneider [37], and paper with T. Nowak in preparation, which is going to be a survey on transience bounds in max-plus algebra based on CSR expansion schemes. New results may be applied in analysing the performance of reversal routing and reversal scheduling algorithms in collaboration with the team of B. Charron-Bost (École Polytechnique).

- **Ultradiscrete KdV.** Sergeev learned about the ultradiscrete KdV model at the conferences in Manchester (April 2011) and Glasgow (July 2011), and worked on the application of max-plus spectral theory to the ultradiscrete analogue of the Lax pair studied recently by R. Willox, J. Satsuma, J. Nimmo and others, based on the idea of S. Gaubert (who also took part in both conferences). In submission [66], he suggested the notion of pairs of fundamental eigenvectors associated with each soliton and showed that the problem can be reduced to finite-dimensional spectral theory with two additional constraints. In some special cases the problem can be solved by means of fundamental eigenvectors, which can be also used in the operation of undressing. However, in the case of several massive solitons the fundamental eigenvectors can never provide a solution, and more elaborate theory is needed.

- **max-Łukasiewicz algebra** Sergeev learned about some new fuzzy linear algebras during his research visit to Czech Republic. He observed that the linear algebraic problems over max-Łukasiewicz algebra, defined in the interval \([0, 1]\) with operations \( a \otimes b := \max(a + b - 1, 0) \) and \( a \oplus b := \max(a, b) \) can be reduced to some problems over max-plus algebra. This observation, being used in submission [63] has led to a new collaboration with Martin Gavalec and his group at the University of Hradec Kralove.

- **Visualization scaling and multi-objective optimization.** Sergeev continued his research on diagonal similarity scalings in max-plus algebra, see submission [67] and new collaboration with B. Benek Gursoy and O. Mason (in preparation). In this new collaboration, the authors analyze the possibilities of simultaneous matrix scaling (visualization), the special case of symmetrically reciprocal matrices, and extension to Pareto optimality.
6.3. Algèbre max-plus, déformations et asymptotiques /Max-plus algebra, deformations and asymptotic analysis

6.3.1. Introduction

Comme indiqué dans le § 3.7, l’algèbre max-plus est la limite d’une déformation de l’algèbre classique, ou plutôt du semi-corps des réels positifs. Elle peut aussi fournir des estimations de ces déformations, puisque

\[ \max(a, b) \leq \epsilon \log(e^{a/\epsilon} + e^{b/\epsilon}) \leq \epsilon \log(2) + \max(a, b). \]  

L’utilisation de ces propriétés a déjà conduit dans le passé aux travaux sur les perturbations de valeurs propres \[1\], \[2\], \[3\], ou sur les grandes déviations \[1\], \[2\]. Dans les travaux qui suivent, nous exploitons ces propriétés dans des contextes reliés ou similaires à ceux de nos travaux précédents.

**English version**

As detailed in § 3.7, max-plus algebra is the limit of a deformation of classical algebra, or more precisely of the semi-field of usual real positive numbers. It can also give estimations for these deformations using for instance (11). By using these properties, we already obtained some works on singular perturbations of matrix eigenvalues \[1\], \[2\], \[3\], or on large deviations \[1\], \[2\]. In the works described below, we are exploiting again these properties in contexts that are related or similar to those of our earlier works.

6.3.2. Aspects tropicaux des algorithmes de scaling matriciel/Tropical aspects of matrix scaling problems

**Participants:** Marianne Akian, Stéphane Gaubert, Laura Grigori, Meisam Sharify Najafabadi.

Le travail de thèse de M. Sharify \[1\] a porté sur les méthodes de mise à l’échelle utilisées en algorithmique numérique matricielle pour améliorer la précision des calculs.

Une première partie du travail, appliquant les techniques de \[1\], \[2\], porte sur les problèmes de valeurs propres. On montre notamment que l’ordre de grandeur des valeurs propres d’un faisceau matriciel est donné (sous des conditions de non-dégénérescence) par les valeurs tropicales, qui peuvent être calculées de manière robuste, et fournissent ainsi une mise à l’échelle pour calculer les valeurs propres classiques.

Une seconde partie du travail (collaboration avec L. Grigori) porte sur le calcul de mises-à-l’échelle issues de la résolution d’un problème d’affectation optimale. On a développé un algorithme dont l’idée est de voir le problème d’affectation comme une limite d’un problème de maximisation d’entropie. Ceci conduit à un préprocessing parallèle, qui permet d’éliminer a priori des coefficients qui ne participent pas aux affectations optimales, de sorte que le problème réduit devient résoluble sur une machine séquentielle. L’algorithme ainsi obtenu est étudié dans le preprint \[1\], qui comprend également des résultats expérimentaux.

**English version**

The PhD work of M. Sharify \[1\] deals with the development of scaling methods in matrix analysis to improve the accuracy of numerical computations.

A first part of the work, applying the techniques of \[1\], \[2\], deals with eigenvalue problems. We show in particular that the order of magnitude of the eigenvalues of a matrix pencil can be determined (under nondegeneracy conditions) by computing tropical eigenvalues. The latter can always be computed accurately and provide a scaling which can be combined with standard numerical methods for matrix pencils.

A second part of the work (collaboration with L. Grigori) deals with the parallel computation of scalings based on the optimal assignment problem. The latter is thought of as a limit of an entropy maximization problem. This leads to a parallel preprocessing, allowing one to eliminate a priori entries which do not belong to optimal assignment, so that the reduced problem becomes solvable on a sequential machine. This algorithm is studied in the preprint \[1\], which also comprises experimental results.
6.3.3. Mesures et applications maxitives

Participants: Marianne Akian, Paul Poncet.

Les mesures et intégrales maxitives qui ont été introduites et ré-introduites sous divers noms dans la littérature (intégrale de Shilkret, sup-mesures, mesures de possibilité, mesures idempotentes de Maslov, etc.), sont définies de manière analogue aux mesures et intégrales usuelles, en remplaçant les lois additive et multiplicative par celles d’un semi-anneau idempotent, comme par exemple le semi-anneau max-plus. Elles peuvent aussi être obtenues comme limites de mesures positives après déformation logarithmique, par le principe des grandes déviations. Entre autres motivations à l’étude de ces mesures, citons les processus max-stables et leur représentation intégrale, les processus extrêmes, ou les grandes déviations à la loi des grands nombres.

Le travail de thèse de Paul Poncet [12] est parti de ces motivations. Il traite essentiellement de ce que l’on appelle l’analyse idempotente, c’est-à-dire l’étude des espaces fonctionnels ou linéaires de dimension infinie sur l’algèbre tropicale, ou tout autre semi-anneau idempotent. Paul Poncet a développé pour cela un point de vue treillis continu comme dans [1], ou plus généralement domaines, et ses travaux pourraient donc aussi avoir des applications en informatique.

La première partie de la thèse traite des mesures maxitives. Paul Poncet a donné une revue des résultats existants concernant l’existence d’une densité cardinale ou d’une densité d’une mesure par rapport à une autre (théorème de Radon-Nikodym), et la régularité d’une mesure maxitive, tout en les comparant et les complétant. En particulier il prouve une réciproque au théorème de Radon-Nikodym pour les mesures maxitives, c’est-à-dire qu’il donne une caractérisation des mesures maxitives ayant la propriété de Radon-Nikodym, il caractérise les mesures maxitives régulières à valeurs dans un domaine, donne un théorème de décomposition des mesures maxitives aussi publié dans [34], et donne un théorème de représentation de Riesz pour les formes linéaires max-plus continues.

Une deuxième partie concerne les convexes dans les semi-treillis ou l’algèbre max-plus. Paul Poncet s’est intéressé à l’existence d’un théorème de type Krein-Milman, à sa réciproque de Milman, et à celle d’un théorème de type représentation de Choquet dans ces structures. Dans le cas des semi-treillis, certains de ces résultats se déduisent rapidement des travaux sur les semi-treillis compacts, mais d’autres sont entièrement nouveaux. Le théorème de Krein-Milman pour les convexes tropicaux, qui n’avait été établi dans la littérature qu’en dimension finie [137], [96], [125], est prouvé en dimension infinie au moyen de celui sur les semi-treillis. Le théorème de représentation de Choquet utilise les notions de mesures maxitives introduites dans la première partie. De tels résultats permettent de retrouver partiellement les résultats sur la frontière de Martin max-plus décrits dans la section 6.1.1.

Enfin dans une troisième et dernière partie, Paul Poncet étudie les semi-groupes inverses dans une tentative d’unification de l’algèbre usuelle et de l’algèbre tropicale.

English version

Maxitive measures and integrals, which have been introduced and re-introduced under different names in the literature (Shilkret integral, sup-measures, possibility measures, Maslov idempotent measures, etc.), are defined analogously to usual measures and integrals, by replacing the additive and multiplicative laws by the laws of an idempotent semiring, such as the max-plus semiring. They can also be obtained as limits of positive measures after logarithmic deformation, by the large deviation principle. Among motivations for the study of this notion, let us mention max-stable processes and their integral representations, extremal processes, or large deviations to the law of large numbers.

The PhD thesis work of Paul Poncet [12] started from these motivations. It concerns essentially what is called idempotent analysis, that is the study of infinite dimensional functional or linear spaces over tropical algebra, or any other idempotent semiring. For this aim, Paul Poncet developed the point of view of continuous lattices, as in [1], or more generally of domains, and his works may have applications in computer science.
The first part of his thesis concerns maxitive measures. Paul Poncet gave a survey of existing results concerning the existence of a cardinal density of a measure, that of a density of a measure with respect to another (Radon-Nikodym theorem), and the regularity of a maxitive measure, while comparing and extending them. In particular he proves a converse to the Radon-Nikodym theorem for maxitive measures, which lead to a characterisation of maxitive measures that have the Radon-Nikodym property, he characterizes domain valued maxitive measures that are regular, gives a decomposition theorem of maxitive measures also published in [34], and gives a Riesz representation theorem for continuous max-plus linear forms.

A second part concerns convex sets in lattices or max-plus algebra. Paul Poncet is showing a Krein-Milman type theorem, its Milman converse, or a Choquet representation type theorem in these structures. In the case of semilattices, some of these results can be deduced easily from works on compact semilattices, but some others are new. The Krein-Milman on tropical convex sets, which in the literature was established in finite dimension only [137], [96], [125], is deduced in infinite dimension from the analogous result concerning semilattices. The Choquet representation theorem uses the notions of maxitive measures introduced in the first part. Such results lead in particular to new proofs of some of the results on Martin boundaries described in Section 6.1.1.

In the third and last part, Paul Poncet is studying inverse semigroups in an attempt to unify usual and tropical algebras.

**6.4. Algorithmes/Algorithms**

**6.4.1. Méthodes multigrilles pour le contrôle stochastique et les jeux répétés à somme nulle/Multigrid methods for stochastic control and repeated zero sum games**

**Participants:** Marianne Akian, Sylvie Detournay.

L’algorithme d’itération sur les politiques est bien connu pour résoudre efficacement les équations de la programmation dynamique associées à des problèmes de contrôle stochastique avec critère à horizon infini (Howard) ou ergodique (Denardo et Fox). Récemment, il a été généralisé au cas de problèmes de jeux à deux joueurs et somme nulle dégénérés (avec paiements ergodiques et de type “multi-chaîne”), au moyen de techniques d’algèbre max-plus et de théorie du potentiel non linéaire [103]. Chaque itération de base de cet algorithme utilise la résolution d’un système d’équations linéaires dont l’opérateur est monotone, mais dont la taille peut être grande, soit parce qu’il provient d’une discrétisation fine d’une équation aux dérivées partielles, soit parce qu’il est associé à un problème discret de grande taille comme le graphe du Web.

Or, la méthode multigrille est l’une des rares méthodes permettant de résoudre, au moins dans les bons cas, des systèmes linéaires en un temps de l’ordre de la taille du système. De plus, alors que la méthode multigrille classique ne s’applique qu’à des discrétisations d’équations aux dérivées partielles elliptiques, la méthode multigrille algébrique (voir par exemple [172]) peut s’appliquer à tout système linéaire présentant des propriétés de monotonie (principe du maximum ou système avec M-matrice).

L’association entre méthodes multigrilles et itérations sur les politiques a déjà été utilisée et étudiée dans le cas de problèmes de contrôle stochastique actualisé (voir par exemple [72], [80]), ainsi que dans le cas d’un algorithme d’itération sur les politiques simplifié pour le contrôle ergodique (voir par exemple [5]), mais pour lequel il n’existe pas de preuve de convergence. La méthode multigrille algébrique a été récemment associée à des méthodes d’apprentissage (voir par exemple [187]). Nous l’avons aussi testée dans le cas de l’itération sur les politiques pour des problèmes de jeux à somme nulle actualisés au cours du stage de Shantanu Gangal en 2007.

La thèse de Sylvie Detournay a pour but de développer et d’étudier un algorithme associant un algorithme d’itération sur les politiques du type de celle introduite par Cochet-Terrasson et Gaubert dans [103] et une méthode multigrille algébrique, afin de résoudre des problèmes de jeux à somme nulle dégénérés, éventuellement posés directement sous forme discrète. Sylvie Detournay a d’abord travaillé sur le cas non dégénéré (actualisé) en codant d’abord seulement l’itération sur les politiques (en C) et appelant des codes libres de méthodes multigrilles algébriques. Ces codes n’étant pas assez souples pour être modifiés, elle a
ensuite codé elle-même certains types de méthodes multigrilles algébriques. Des tests sur des discrétisations d’équations aux dérivées partielles d’Hamilton-Jacobi-Bellman ou d’Isaacs, ou d’inéquations variationnelles ont donné de bons résultats et sont présentés dans [15].

Sylvie Detournay a travaillé cette année sur le cas de problèmes avec critère moyen en temps. Elle a implémenté et raffiné l’algorithme proposé par Cochet-Terrasson et Gaubert [103], en l’associant soit à des méthodes de résolution exacte de systèmes linéaires, soit à des méthodes multigrilles algébriques, en utilisant aussi des méthodes multigrilles multiplicatives pour le calcul de la mesure invariante de chaînes de Markov irréductibles. Ceci a permis en particulier l’obtention de résultats numériques dans le cas de discrétisations d’équations d’Isaacs associées à des jeux de poursuite déterministes ou aléatoires. Plusieurs de ces résultats ont été présentés cette année lors de 2 conférences internationales [44], [45], et devraient faire l’objet d’un article en préparation. Par ailleurs dans un article avec Jean Cochet-Terrasson et Stéphane Gaubert [54], nous présentons l’algorithme, sa convergence et des résultats numériques obtenus avec des méthodes de résolution exacte de systèmes linéaires.

English version

Policy iteration is a powerful and well known algorithm to solve the dynamic programming equation associated to one player problems. It has recently been extended to degenerate two players problems (with ergodic payoff and in “multichain” cases) using ideas from max-plus algebra and nonlinear potential theory [103]. One basic iteration of the algorithm consists in solving a linear system which operator is monotone, but which size may be large since it comes from the discretization of a partial differential equation or since it is associated to a large size discrete problem such as the Web graph.

For the solution of large size linear systems, the state of art consists of multigrid methods which are often able to solve systems in linear time. Whereas multigrid methods can only be applied to systems that come from discretizations of elliptic partial differential equations, algebraic multigrid methods (see for instance [172]) can be applied to any linear system with monotonicity properties (discrete maximum principle or system with a M-matrix).

The association of multigrid methods with policy iteration has been used and studied in the case of discounted stochastic control problems (see for instance [72], [80]), or in the case of a simplified policy iteration algorithm for ergodic control (see for instance [5]), but for which no proof of convergence is known. Some recent work combines the algebraic multigrid method with learning methods [187]. We have also tested it in the case of policy iterations for discounted zero-sum two-player games, during the internship of Shantanu Gangal in 2007.

The aim of the PhD thesis of Sylvie Detournay is to develop and study an algorithm for degenerate two player games (that may come from a discrete time and finite state space model) combining a policy iteration such as that introduced in [103] and an algebraic multigrid method (AMG). Sylvie Detournay has first worked on the nondegenerate (discounted) case, by coding first the policy iterations (in C) and using free AMG softwares. Since these softwares cannot be modified easily, she has then implemented some types of AMG algorithms (in C). Some tests on discretisations of Hamilton-Jacobi-Bellman or Isaacs partial differential equations or variational inequalities gave good results and are presented in [15].

She has worked this year on the case of problems with mean-payoff criteria. She has implemented and refined the algorithm proposed by Cochet-Terrasson and Gaubert [103], associated either to direct linear solvers, or to the AMG methods already used in the nondegenerate case, and also used multiplicative AMG methods developed in the literature for computing invariant measures of Markov chains. This allows her to obtain numerical results in the case of discretisations of Isaacs equations associated to deterministic or stochastic pursuit games. Several of these results were presented this year in 2 international conferences [44], [45] and are part of an article in preparation. Moreover, in an article with Jean Cochet-Terrasson and Stéphane Gaubert [54], we are presenting the algorithm, its convergence and numerical results obtained with direct linear solvers.

6.4.2. Algorithmique des polyèdres tropicaux/Algorithmics of tropical polyhedra
Participants: Xavier Allamigeon, Stéphane Gaubert, Eric Goubault [CEA].

X. Allamigeon, S. Gaubert, et E. Goubault, ont développé dans [82 ], [ 57 ] plusieurs algorithmes permettant de manipuler des polyèdres tropicaux. Ceux-ci correspondent aux travaux décrits dans § 6.2.1 . Ils permettent notamment de déterminer les sommets et rayons extrêmes d’un polyèdre tropical défini comme intersection de demi-espaces, ou inversement, de calculer une représentation externe à partir d’un ensemble de générateurs. Ces algorithmes sont implémentés la bibliothèque TPLib (voir § 5.3 ).

English version

X. Allamigeon, S. Gaubert, and E. Goubault, have developed in [82 ],[ 57 ] algorithms allowing one to manipulate tropical polyhedra. They correspond to the contributions described in § 6.2.1 . In particular, they can be used to determine the vertices and extreme rays of a tropical polyhedron defined as the intersection of half-spaces, or inversely, to compute an external description from a set of generators. These algorithms are implemented in the library TPLib (see § 5.3 ).

6.4.3. Problèmes d’accessibilité dans les hypergraphes orientés et leur complexité/Reachability problems in directed hypergraphs and their complexity

Participant: Xavier Allamigeon.

Les hypergraphes orientés sont une généralisation des graphes orientés, dans lesquelles chaque arc relie un ensemble de sommets à un autre. Ils jouent un rôle important dans les travaux récents sur la convexité tropicale (voir § 6.2.1 ), puisqu’ils offrent une représentation naturelle des cônes définis sur le sous-semi-anneau booléen $B = \{-\infty, 0\}$.

Dans un travail de X. Allamigeon [56 ], on étudie la complexité de problèmes d’accessibilité sur les hypergraphes orientés. Nous introduisons un algorithme de complexité presque linéaire permettant de déterminer les composantes fortement connexes terminales (qui n’accèdent à aucune autre composante si ce n’est elles-mêmes) d’un hypergraphe.

Nous établissons également une borne inférieure sur-linéaire sur la taille de la réduction transitive de la relation d’accessibilité dans les hypergraphes. Cela indique que la relation d’accessibilité dans les hypergraphes orientés est combinatoirement plus complexe que celle des graphes orientés. Cela suggère aussi que des problèmes comme le calcul des composantes fortement connexes est plus difficile sur les hypergraphes que sur les graphes. Nous mettons d’ailleurs en évidence une réduction en temps linéaire du problème du calcul des ensembles minimaux dans une famille d’ensembles donnée, vers le problème du calcul de toutes les composantes fortement connexes d’un hypergraphe. Le problème du calcul des ensembles minimaux a été largement étudié dans la littérature [163 ], [ 183 ], [ 182 ], [ 164 ], [ 165 ], [ 166 ], [ 118 ], [ 88 ], et aucune algorithme en temps linéaire n’est connu à ce jour.

English version

Directed hypergraphs are a generalization of directed graphs, in which the tail and the head of the arcs are sets of vertices. It appears that they play an important role in the recent works on tropical convexity (see § 6.2.1 ), since they offer a natural representation of cones defined over the boolean sub-semiring $B = \{-\infty, 0\}$.

In a work of X. Allamigeon [56 ], we study the complexity of reachability problems on directed hypergraphs. We introduce an almost linear-time algorithm allowing to determine the terminal strongly connected components (a component is said to be terminal when no other component is reachable from it).

We also establish a super-linear lower bound over the size of the transitive reduction of the reachability relation in directed hypergraphs. This indicates that the reachability relation is combinatorially more complex in directed hypergraphs than in directed graphs. This also suggests that reachability problems such as computing all strongly connected components are likely to be harder in hypergraphs than in graphs. Besides, we show that the minimal set problem can be reduced in linear time to the problem of computing all strongly connected components in hypergraphs. The former problem consists in finding all minimal sets among a given family of sets. It has been well studied in the literature [163 ], [ 183 ], [ 182 ], [ 164 ], [ 165 ], [ 166 ], [ 118 ], [ 88 ], and no linear time algorithm is known.
6.4.4. Approximation max-plus de fonctions valeurs/Max-plus approximation of value functions

Participants: Stéphane Gaubert, Zheng Qu, Shanjian Tang [Fudan University, Shanghai], William McEneaney [San Diego University].

La thèse de Zheng Qu, démarrée en septembre 2010, supervisée par S. Gaubert et S. Tang, porte sur le développement de méthodes tropicales en programmation dynamique approchée.

Un problème de base consiste à approcher au mieux la fonction valeur d’un problème de contrôle ou de jeux par le supremum d’un petit nombre de fonctions choisies dans un dictionnaire fixé a priori. Ce problème est abordé dans [43]. A l’aide de résultats de Grüber portant sur l’approximation de corps convexes par des polytopes, on donne tout d’abord une borne montrant le caractère inévitable de la malédiction de la dimension, pour certaines méthodes de type base max-plus, lorsque la fonction valeur est $C^2$ et strictement convexe. Ce résultat montre que ces familles de méthodes sont asymptotiquement coûteuses lorsque la précision requise tend vers 0. Elles permettent cependant d’obtenir rapidement des approximations certifiées d’une précision donnée pas trop petite (dans ce cas, la malédiction de la dimension est absente). On s’intéresse ensuite à un problème algorithmique clé sous-jacent à ces méthodes, qui consiste à éliminer dynamiquement des fonctions redondantes intervenant dans la représentation. On démontre dans [43] que ce problème est équivalent à un problème géométrique de localisation, dans lequel la métrique est non symétrique (de type Bregman). Ceci a permis d’appliquer divers algorithmes de localisation, conduisant à une amélioration de la méthode antérieure [155].

Un autre travail de Zheng Qu porte sur les équations de Riccati généralisées associées à des problèmes de contrôles stochastique avec critère quadratique, dans lesquels la dynamique comporte un terme bilinéaire en le contrôle et le bruit. Alors que le flot de l’équation de Riccati classique est contractant pour la métrique Riemannienne invariante, pour la métrique de Thompson, ainsi que pour toutes les métriques de Finsler invariantes sur le cône des matrices symétriques positives, on montre ici que le flot de l’équation de Riccati généralisée en question est seulement contractant pour la métrique de Thompson (sous des hypothèses naturelles).

6.5. Applications

The PhD work of Zheng Qu, which started in September 2010, and is supervised by S. Gaubert and S. Tang, aims in particular at developing tropical methods in approximate dynamic programming.

A basic problem consists in approximating the value function of an optimal control or game problem by a supremum of a small number of functions taken from a prescribed dictionary. This problem is addressed in [43]. By applying results of Grüber concerning the approximation of convex bodies by polytopes, we give first a negative result, showing that the curse of dimensionality cannot be avoided by a family of max-plus basis methods, when the value function is $C^2$ and strictly convex. This result shows that this family of methods is asymptotically computationally expensive when the requested precision tends to 0. However, they can be used to obtain quickly (in a curse of dimensionality free way) certified approximations with a fixed (not too small) precision. Then, we addressed a key algorithmic subproblem, consisting in trimming dynamically the redundant functions in a max-plus representation. We showed in [43] that this problem is equivalent to a geometric facility location problem, with a non symmetric Bregman type metric. This allowed us to apply several facility location algorithms, leading to an improvement of the earlier method [155].

Another work of Zheng Qu deals with the generalized Riccati equations associated to stochastic optimal control problems with quadratic cost, in which the dynamics comprises a term which is bilinear in the control and in the noise. Whereas the flow of the standard Riccati equation is known to be a contraction for the invariant Riemannian metric, the Thompson metric, and more generally, for all invariant Finsler metrics on the cone of positive definite matrices, it is shown here that the flow of this generalized Riccati equation is only contracting with respect to Thompson metric (under natural assumptions).
6.5.1. Introduction

Nous présentons maintenant plusieurs travaux de nature appliquée, touchant à des domaines variés, dans lesquels nous exploitons certaines des techniques mathématiques présentées précédemment, et particulièrement celles qui relèvent de la théorie de Perron-Frobenius non-linéaire et de la convexité tropicale. Ces applications utilisent aussi des techniques d’algèbre linéaire ou d’optimisation convexe.

**English version**

In this section, we describe several applied works in which we use some of the theoretical tools developed by the team, including non-linear Perron-Frobenius theory and tropical convexity. Some of these applications also make an intensive use of linear algebraic and convex programming methods.

6.5.2. Propriétés des valeurs propres de Perron et de Floquet, et application en chronothérapeutique/Properties of Perron and Floquet eigenvalue, with an application to chronotherapeutics

**Participants:** Frédérique Billy [Projet BANG, INRIA], Jean Clairambault [Projet BANG, INRIA], Olivier Fercoq, Stéphane Gaubert, Thomas Lepoutre [Projet BANG puis DRACULA, INRIA].

On s’intéresse à des modèles de systèmes dynamiques monotones structurés en âge représentant la croissance de populations de cellules (saines ou tumorales), à la suite de travaux de Clairambault et Perthame. Il s’agit de comprendre l’influence du contrôle circardien sur la croissance des cellules. Dans le cas stationnaire, le taux de croissance est représenté par une valeur propre de Perron. Dans le cas périodique, il s’agit d’une valeur propre de Floquet. Les travaux [42], [52], [58] portent sur l’identification de ces modèles ainsi que sur un problème de contrôle thérapeutique, consistant à minimiser le taux de croissance des cellules tumorales sous une contrainte de non-toxicité du traitement (maintien d’une population de cellules saines). Ce travail s’appuie en particulier sur un algorithme d’optimisation de la valeur propre de Perron d’une matrice développé par Fercoq dans un autre contexte [62].

**English version**

We study monotone dynamical systems representing the growth of cells (healthy or tumoral), following a work of Clairambault and Perthame. The goal is to understand how the circadian control influences the growth of cells. In the case of stationary monotone systems, this growth is measured by the Perron root. In the time periodic case, this Perron root is replaced by a Floquet multiplier.

The works [42], [52], [58] deal with the identification of these models, together with a therapeutic control problem, consisting in minimizing the growth rate of tumoral cells, under a non-toxicity constraint (preserving the population of healthy cells). This works relies in particular on a fast algorithm to optimize the Perron eigenvalue of a matrix, developed by Fercoq in a different context [62].

6.5.3. Équations aux dérivées partielles en dynamique des populations/Partial differential equations from population dynamics

**Participant:** Sepideh Mirrahimi.

Un des problèmes sur lequel on a travaillé est un modèle de propagation de populations sexuées dans l’espace [64] où on dérive un modèle étudié par des biologistes, à partir d’un modèle de populations structurées et on étudie la dynamique de ce dernier. Nous avons aussi travaillé sur un problème de protéines moteurs, où nous étudions le comportement asymptotique d’un système de deux équations couplées de Fokker-Planck dans un environnement périodique. Avec une approche d’homogénéisation et en utilisant des techniques de solutions de viscosité, on montre que les protéines se déplacent dans une direction constante [65]. De plus, en utilisant des idées venant des problèmes similaires mais discrets, avec Stéphane Gaubert nous étudions les EDPs qui décrivent la dynamique d’une population asexuée sous l’effet des mutations et de la sélection naturelle, et on cherche à déterminer la limite en temps long de la densité de population.

**English version**
One of the problems studied is a model of propagation of a sexual population in space [64], where we derive a model studied by biologists, from a structured population model. We then study the behavior of the solution to the latter model. We also have worked on a problem of motor proteins, where we study the asymptotic behavior of a time dependent, weakly coupled, Fokker-Planck system of two equations set in a periodic environment. By a homogenization approach and using viscosity solutions technics we prove that the molecules either move along a fixed filament with a constant speed and direction or remain immobile [65]. Moreover, using ideas coming from discrete models, with Stéphane Gaubert we study some PDEs that describe the dynamics of an asexual population under mutations and natural selection. We try to determine the long-time limit of the population density.

6.5.4. Identification du trafic dans les réseaux IP/Traffic identification in IP networks

Participants: Mustapha Bouhtou [Orange Labs], Stéphane Gaubert, Guillaume Sagnol.

Le travail de thèse de Guillaume Sagnol, réalisé en collaboration avec Orange Labs dans le cadre d’un contrat “CRE”, a porté sur l’identification du trafic dans des réseaux IP, problème auquel il a appliqué des d’optimisation SDP et d’optimisation sous-modulaire afin de développer des algorithmes passant à l’échelle. Cette thèse s’est achevée fin 2010 [173]. Les articles suivants relatifs au travail de thèse ont été publiés cette année: [36], [35].

English version

The PhD work of Guillaume Sagnol, done in collaboration with Orange Labs in the framework of a “CRE” research contract, dealt with the identification of the traffic in IP networks. Sagnol applied SDP and submodular optimization techniques to develop scalable algorithms (adapted to large networks). The PhD defense took place in 2010 [173]. Some contributions of the Phd have been published this year in [36],[35].

6.5.5. Analyse statique de programmes et itération sur les politiques/Static analysis of computer programs and policy iteration

Participants: Assale Adjé, Stéphane Gaubert, Eric Goubault [CEA].


Un problème important consiste à déterminer le plus petit point fixe (l’algorithme de [14] fournit un point fixe, qui peut ne pas être minimal). Ce problème est abordé dans [30], où l’approche de [14] est comparée avec une approche duale développée par Gawlitza et Seidl.

English version

The PhD work of A. Adjé [11], co-supervised by S. Gaubert and E. Goubault, applies methods from game theory and optimization (generalized duality, convex and non convex programming) to the fixed point problems arising in static analysis of programs by abstract interpretation. We introduced in [14] a new domain in static analysis, which extends to nonlinear cases the “templates” introduced by Manna, Sankaranarayanan, and Sipma [175]. This domain allows one to represent accessible sets that are non convex. These are defined by finitely many inequalities taken from a dictionary. This allows one to use in particular the information provided by Lyapunov functions, which are often known in applications arising from engineering. We showed in [14] that experimentally accurate invariants can be obtained by coupling policy iteration with Shor.
relaxation (SDP relaxation of convex programming problems). This yields accurate abstractions of some numerical programs, like linear filters with thresholds.

An important problem consists in determining the smallest fixed point (the algorithm of [14] yields a possibly non minimal fixed point). This problem is addressed in [30], in which the approach of [14] is compared with a dual approach developed by Gawlitza and Seidl.

6.5.6. Optimization du référencement sur la toile/ Optimization of web referencing

Participants: Marianne Akian, Mustapha Bouhtou [Orange Labs], Olivier Fercoq, Stéphane Gaubert.

La thèse d’O. Fercoq, co-encadrée par M. Akian, M. Bouhtou, et S. Gaubert, financée par un CRE d’Orange Labs, a pour but d’appliquer des méthodes d’optimisation et de théorie des jeux à l’optimisation de services en lignes. On a tout d’abord étudié le problème de l’optimisation du référencement, que l’on formalise en se donnant par exemple un ensemble d’hyperliens et de ressources obligatoires, dont la nature et la position sur le site web sont déterminées à l’avance par le concepteur. Cet ensemble forme en quelque sorte le squelette du site web. On se donne aussi un ensemble d’hyperliens ou de ressources facultatives, pour lesquels le concepteur du site a certains degrés de liberté (le lien ou le contenu peut être mis sur une page plutôt qu’une autre, voire être omis).

Dans [61], on aborde le problème de l’optimisation du “Pagerank” dans ce cadre, en appliquant des techniques de décision Markovienne classiques et sous-contraintes. Le problème peut en effet se ramener à un problème de contrôle ergodique ou de contrôle ergodique sous contraintes (ergodiques), selon que les contraintes sur les hyperliens sont locales à chaque page ou font intervenir plusieurs pages. On traite à la fois le cas relaxé où les probabilités de passage d’une page à une autre peuvent être des réels positifs quelconques (on peut par exemple supposer que cette probabilité dépend de la position et des caractères utilisés pour l’hyperlien correspondant) et le cas discret où ces probabilités sont uniformes parmis celles qui sont strictement positives (comme dans la modélisation classique conduisant au calcul du Pagerank). On montre que cette famille de problèmes correspondent à des problèmes de programmation dynamique avec un nombre exponentiel de contrôles, mais où les polytopes des mesures de probabilités de transition admettent des oracles de séparation polynomiaux. On obtient de la sorte des résultats de complexité, ainsi que, sous certaines hypothèses, des algorithmes adaptés à des instances de grande taille, couplant programmation dynamique et relaxation Lagrangienne. Ces algorithmes ont été testés sur un fragment du graphe du web.

Un critère de référencement classique, alternatif au pagerank, est donné par le vecteur propre de Perron. O. Fercoq a abordé le problème associé d’optimisation du référencement, qui se révèle plus difficile que celui du pagerank, en raison de l’absence de propriété de convexité. Cependant, il a développé un algorithme rapide et creux (basé sur des propriétés de rang 1 d’opérateurs intervenant dans le calcul de dérivées du critère) permettant de calculer un optimum local du référencement [62]. Il a enfin donné un algorithme analogue pour optimiser le score “HOTS” de Tomlin.

**English version**

The goal of the PhD work of O. Fercoq, cosupervised by M. Akian, M. Bouhtou, and S. Gaubert, and supported by a research contract (CRE) of Orange Labs, is to apply optimization and game theory methods to the optimization of online services. We started by investigating the problem of the optimization of referencing, which we modelled by considering a family of compulsory hyperlinks and resources (fixed in advance by the website designer, these constitute the “skeleton” of the website) and also a family of facultative hyperlink or resources (some links may be omitted or some other links may be added).

In [61], we are approaching the problem of the pagerank optimization in this framework, by applying usual and constrained Markov decision processes techniques. This problem can indeed be reduced to an ergodic control problem without or with (ergodic) constraints, depending on the fact that hyperlinks constraints are local to each web page or depend on several web pages. We study the relaxed problem where the transition probabilities from one page to another may be any positive real (one may assume for instance that this probability depends on the position and type used for the corresponding hyperlink), as well as the discrete problem where these probabilities are uniform among the positive ones (as in the usual modelisation leading
to the Pagerank). We show that these problems can be reduced to dynamic programming problems with exponentially many discrete actions, in which however the polytopes of transition probability measures admit polynomial time separation oracles. We derive from this approach polynomial time complexity results, as well as under some additional assumption, scalable algorithms (adapted to large web graphs), coupling dynamic programming and Lagrange relaxation. The latter have been tested on a real subgraph of the web.

A classical alternative ranking relies on the Perron eigenvector. O. Fercoq treated the associated optimisation problem, which turns out to be harder than in the pagerank case, due to the lack of convexity properties. However, he developed a fast (sparse) algorithm, exploiting the rank 1 properties of operators appearing when computing the derivative of the objective function, allowing one to compute a local optimum [62]. He also developed a similar method to optimize Tomlin’s “HOTS” score.

6.5.7. Gestion du revenu appliquée à la tarification de services données/Yield management applied to pricing of data services

Participants: Mustapha Bouhtou [Orange Labs], Jean-Baptiste Dumont, Stéphane Gaubert.

Le travail de thèse CIFRE de J-B. Dumont, qui a démarré en Septembre 2010, sous la supervision de M. Bouhtou et S. Gaubert, porte sur la tarification de services data et la gestion des ressources dans les réseaux mobiles, qui est abordée à l’aide de techniques de contrôle et d’optimisation stochastique. Dumont a développé un premier modèle de tarification, permettant d’inciter les clients à reporter leur demande en dehors des périodes les plus chargées.

English version

The CIFRE PhD work of J-B. Dumont started in September 2012, under the joint supervision of M. Bouhtou and S. Gaubert. It deals with the pricing of data services and resource allocation in mobile networks. This is addressed through stochastic control and stochastic optimization techniques. Dumont developed a first model of pricing, giving some incentive to the customers to move their demand from loaded to less loaded time periods.
MC2 Project-Team

6. New Results

6.1. Multi-fluid flows

- Microfluidics: (2011) Participants CHB, Johana Pinilla (doc), Sandra Tancogne (MC Reims) To handle oil recovery by chemical processes it is useful to better understand the behaviour of multi-fluids flows in a saturated soil. The porous medium is mimiced by a network of micro channels. The simulation of immiscible multi-fluids flows is then performed by means of the level-sets and the penalization methods to track the interfaces between the fluids and to get rid of the geometry difficulties. In addition the Cox law is added in the model to better move the interfaces during the simulations.

- Microfluidics: concerning visco-elastic fluids in micro-channel, one has often to compute solutions of system for which the viscosity in the Stokes part is much smaller than that involved in the extra-stress. In his thesis, V. Huber has introduced a new scheme to overcome this difficulties without changing the complexity of the scheme (PhD in progress).

6.2. Cancer modelling

In 2010, we have improved our generic mathematical models describing tumor growth. These models were then specialized for several types of cancer (thyroidal lung nodules, brain tumors). The algorithm used to recover the parameters of these models from medical images has also been greatly improved and is now adapted to run on HPC architectures.

+ Secondary tumors in the lung:
  The mathematical models describing the growth of secondary in the lungs have now settled and are well understood. The main focus of the year was to keep on using these models on patient data. New clinical case were selected by clinicians from the Institut Bergonié, there are currently under study. The model is currently able to reproduce the growth observed on 5 clinical cases. In 2011, various improvements to the calibration algorithms were made. The initial seeding of the algorithms was a weak point of the procedure. This has been much improved using a genetic algorithm. A complete rewrite of the routines was done to improve their versatility and efficiency. Previously, the numerical simulations and calibration were performed in 2D (clinicians selected the most relevant slice showing the evolution of the tumor). Work is now ongoing to switch to full 3D computations and calibration.

+ Breast cancer modeling:
  In collaboration with the University of Houston and the Methodist Hospital, a new mathematical model describing the growth of breast cancer has been developed. This multi-scale model takes many aspects of the disease into account and allows to study the interplays between the various mechanisms responsible for the evolution of the tumor.

+ Modeling glioblastomas:
  In 2011, a hierarchy of models describing the growth of brain tumors was developed (and described in a submitted paper) in collaboration with University of Alabama at Birmingham. As we wished to obtain models that could be calibrated from patient data and yet be reasonably accurate, we believe that these models are suitable trade-offs between the simplicity of the Swanson’s model (the only one used on patient data of brain tumors so far) and the accuracy of more complex models (that cannot really produce quantitative results). In particular, two models were built. The first one allows to study the efficacy of anti-angiogenic therapies. It seems to predict that the efficacy of these treatments is limited, this could be confirmed by a world-wide ongoing clinical study. The second model has been validated and we are trying to recover its parameters for a patient in 3D (which is a rather unique initiative to our knowledge).
Modelling of electrochemotherapy: Two articles related to the electrical cell modelling have been done ([58], [54]). The first one deals with the influence of the ionic fluxes on the transmembrane voltage potential and on the cell volume. The main insight of the results consists in linking the transmembrane potential with the cell volume: it has been observed experimentally that cells with a low voltage potential do divide, whereas cells with high voltage potential do not, and the obtained relationship between voltage potential and cell volume can provide an explanation. The second article deals with a new model of cell electroporation essentially based on the experimental results of the I.G.R. In this paper we describe precisely the model, which takes into account the main experimental results in the electroporation process, and we present a variational formulation inherent to the model that leads to new efficient schemes in order to numerically solve the involved P.D.E.

6.3. Newtonian fluid flows simulations and their analysis

- Simulations of water distribution systems: Water losses may constitute a large amount of the distributed total water volume throughout water distribution systems. Here, a new model method is proposed that intends to minimize the total water volume distributed through leakage reduction. Our group has worked on the derivation of advection-reaction-diffusion type equations with an explicit relationship between the local pressure and the leakage rate. An original splitting technique to solve this type of hydraulic problem was then achieved. This technique allows pressure-dependent leakage to be taken into account, whereas in most models leakage is assumed to be uniform along a pipe. Finally, a constrained optimization problem was formulated for leakage reduction in WDS. The control variable had the mean of a local head loss and is considered in the Boundary Conditions to avoid dealing with discontinuities in the governing equations. The objective function to minimize was a regularization of the total water volume distributed. Specific operational constraints were added to ensure enough pressure at consumption points. The direct solution for this minimization problem was sought with a Gradient type method. The leakage reduction was proven to be significant in a case study. The percentage of leakage reduced from 24% to 10% in the linear relationship between pressure and leakage flow rate. With other leakage exponents, the same rate of reduction was achieved. The method was applied on a real network in the South-West of France. Controlling the pressure at two different strategic points permits a significant amount of the total distributed water to be saved (5%). This work was performed in collaboration with Cemagref Bordeaux. Future work will consist of applying a sensitivity analysis of control location points to optimize the method.

- Incompressible flows: modeling and simulation of moving and deformable bodies. The incompressible Navier-Stokes equations are discretized in space onto a fixed cartesian mesh. The deformable bodies are taken into using a first order penalization method and/or second order immersed boundary method. The interface between the solid and the fluid is tracked using a level-set description so that it is possible to simulate several bodies freely evolving in the fluid. A turbulence model based on Samborinsky model has been added to the numerical code. The numerical code written in the C language is massively parallel. The large linear systems (over than 100 millions of dofs) are solved using the Petsc Library. As an illustration of the methods, fish-like locomotion is analyzed in terms of propulsion efficiency. Underwater maneuvering and school swimming are also explored. We were able to simulate the three-dimensional flow about a swimmer for realistic physical configurations. See [17]. Another application is the turbulent 3D flow around complex wind turbine (see http://www.math.u-bordeaux1.fr/bergmann/ and http://www.math.u-bordeaux1.fr/MAB/mc2/analysis.html for simulation movies). Wake flows generated by boat propellers are also modeled and simulated.

- Turbulence flow on an hemisphere: (2011) Participants CHB, Patrick Fischer, Yong Liang Xiong (PD) ANR Cyclobulle lead by Hamid Kellay Soap hemi-bubble film experiments have shown some links between the formation of vortices when the hemi-bubble is heated at the equator and the formation of tornados in the earth atmosphere. Two-dimensional simulations using a stereographic map are used to compare to these experimental results and confirm the results when Coriolis force and heat source terms are added.

- Compressible flows: Immersed boundary methods. We are concerned with immersed boundary methods, i.e., integration schemes where the grid does not fit the geometry, and among this class
of methods, more specifically with cartesian grid methods, where the forcing accounting for the
presence of boundaries is performed at the discrete level. We have developed a simple globally
second order scheme inspired by ghost cell approaches to solve compressible flows, inviscid as well
as viscous. In the fluid domain, away from the boundary, we use a classical finite-volume method
based on an approximate Riemann solver for the convective fluxes and a centered scheme for the
diffusive term. At the cells located on the boundary, we solve an ad hoc Riemann problem taking
into account the relevant boundary condition for the convective fluxes by an appropriate definition
of the contact discontinuity speed. This method can easily be implemented in existing codes and is
suitable for massive parallelization. It has been validated in 1D and 2D for Euler and Navier-Stokes
compressible equations. The order of convergence of the method is similar to those observed in the
literature: second order globally and first order locally, near the interface for Euler equations, and
second order locally and globally for Navier-Stokes equations.

We have developed a new cartesian method to solve elliptic problems with immersed interfaces.
These problems appear in numerous applications, among them: heat transfer, electrostatics, fluid
dynamics, but also tumour growth modelling, or modelling of electric potential in biological cells.
This method is second order accurate in the whole domain, notably near the interface. The originality
of the method lies on the use of additional unknowns located on interface points, on which are
expressed flux equalities. Special care is dedicated to the discretization near the interface, in order
to recover a stable second order accuracy. Actually, a naive discretization could lead to a first
order scheme, notably if enough accuracy in the discretization of flux transmission conditions is
not provided. Interfaces are represented with a distance level-set function discretized on the grid
points. The method has been validated on several test-cases with complex interfaces in 2D. A parallel
version has been developed using the PETSC library.

- Simulations of fluid-solid interactions : The interaction of an elastic structure and an fluid occurs
  in many phenomena in physics. To avoid the difficulty of coupling lagrangian elasticity with an
eulerian fluid we consider a whole eulerian formulation. The elasticity of the structure is computed
  with retrograde caracteristics which satisfy a vectorial transport equation. We derive the associated
  fluid-structure models for incompressible and compressible media. The equations are discretized on
  a cartesian mesh with finite differences and finite volumes schemes. The applications concern the
  bio-locomotions and the study of air-elastic interaction.

- Vortex methods : The aim of this work is to couple vortex methods with the penalization methods
  in order to take advantage from both of them. This immersed boundary approach maintains the
efficiency of vortex methods for high Reynolds numbers focusing the computational task on the
rotational zones and avoids their lack on the no-slip boundary conditions replacing the vortex sheet
method by the penalization of obstacles. This method that is very appropriate for bluff-body flows is
validated for the flow around a circular cylinder on a wide range of Reynolds numbers. Its validation
is now extended to moving obstacles (axial turbine blades) and three-dimensional bluff-bodies (flow
around a sphere). See [ 71 ]. Moreover, using the global properties of the penalization method, this
technique permits to include porous media simultaneously in the flow computation. We aim to adapt
the porous media flows to our new method and to apply it in order to implement passive control
techniques using porous layers around bluff-bodies.

- Domain decomposition : Domain decomposition methods are a way to parallelize the computation
  of numerical solutions to PDE. To be efficient, domain decompositions methods should converge
independently on the number of subdomains. The classical convergence result for the additive
Schwarz preconditioner with coarse grid is based on a stable decomposition. The result holds for
discrete versions of the Schwarz preconditioner, and states that the preconditioned operator has a
uniformly bounded condition number that depends only on the number of colors of the domain
decomposition, and the ratio between the average diameter of the subdomains and the overlap width.
Constants are usually non explicit and are only asserted to depend on the "shape regularity" of the
domain decomposition.
Last year, we showed the result holds the additive Schwarz preconditioner can also be defined at the continuous level and provided completely explicits estimates. This year, we established that a similar result also holds for non shape regular domain decompositions where the diameter of the smallest subdomain is significantly smaller than the diameter of the largest subdomain. The constants are also given explicitely and are independent of the ratio between the diameter of the largest subdomain and the diameter of the smallest subdomain.

6.4. Flow control and shape optimization

- Flow control : (2011) Participants CHB, Iraj Mortazavi, Emmanuel Creus (Lille), Patrick Gilliron (Renault chercheur indpendant !) An efficient active control of the two- and three-dimensions flow around the 25¡ rear window Ahmed body has been performed. A careful theoretical and numerical study of the trajectories of the vortices allows to adapt the control in order to improve its efficiency and get a better drag reduction.
5. New Results

5.1. Computational quantum chemistry

Participants: Eric Cancès, Ismaila Dabo, Virginie Ehrlacher, Salma Lahbabi, Francis Nier, Gabriel Stoltz.

In computational quantum chemistry as in most of our scientific endeavours, we pursue a twofold goal: placing the models on a sound mathematical grounding, and improving the numerical approaches.

E. Cancès, V. Ehrlacher, S. Lahbabi and G. Stoltz have addressed issues related to the modeling and simulation of local defects in periodic crystals (see [61] for a pedagogical introduction).

Computing the energies of local defects in crystals is a major issue in quantum chemistry, materials science and nano-electronics. In collaboration with M. Lewin (CNRS, Cergy), E. Cancès and A. Deleurence have proposed in 2008 a new model for describing the electronic structure of a crystal in the presence of a local defect. This model is based on formal analogies between the Fermi sea of a perturbed crystal and the Dirac sea in Quantum Electrodynamics (QED) in the presence of an external electrostatic field. The justification of this model is obtained using a thermodynamic limit of Kohn-Sham type models. In [29], E. Cancès and G. Stoltz studied the time evolution of defects within this model, in the context of linear response, which allowed them to give a rigorous meaning to the Adler-Wiser formula for the frequency-dependent dielectric permittivity of crystals. In [27] E. Cancès et V. Ehrlacher have proved that local defects are always neutral in the Thomas-Fermi-von Weizsäcker (TFW) theory. In this respect, all TFW crystals behave like metals. In collaboration with M. Lewin, E. Cancès and S. Lahbabi are working on the extension of Kohn-Sham like models to disordered systems.

On the numerical side, E. Cancès has worked with Y. Maday and R. Chakir (University Paris 6) on the numerical analysis of the electronic structure models. In [26], they have obtained optimal a priori error bounds for the the planewave approximation of the Thomas-Fermi-von Weizsäcker and the Kohn-Sham LDA models. Together with Y. Maday, E. Cancès and V. Ehrlacher have analyzed the computation of eigenvalues in spectral gaps of locally perturbed periodic Schrödinger operators [28].

Photovoltaic cells based upon organic photovoltaic (OPV) semiconductors are cost-effective, light-weight alternatives to conventional silicon and thin-film technologies. This year, in collaboration with A. Ferretti, N. Poilvert, N. Marzari (MIT and University of Oxford), M. Cococcioni (University of Minnesota), and Y. L. Li (Xiamen University), I. Dabo has worked on the prediction of the electronic spectra and electrical response of molecular and polymer compounds for OPV applications. Specifically, Y. L. Li and I. Dabo have implemented an efficient computational method for studying the electrical response of semiconducting polymers, in close agreement with more expensive local basis-set algorithms [52]. In parallel, A. Ferretti, I. Dabo, M. Cococcioni, and N. Marzari have applied recently developed electronic-structure theories, namely, orbital-dependent density-functional theories (OD-DFTs), to describe donor and acceptor levels in semiconducting organic materials, demonstrating that OD-DFT is apt at describing donor and acceptor levels within 0.1-0.4 and 0.2-0.6 eV of experiment [41]. Future work includes the prediction of the band structure of semiconductor alloys and the simulation of molecular heterojunctions.

The current works by F. Nier related to electronic structure calculations concern the mean field limit of the quantum dynamics in the bosonic setting; and the non linear modelling of Bose-Einstein condensates. A few years ago, Z. Ammari and F. Nier initiated a program about the Hamilton mean field dynamics of a large numbers of bosons. Their approach reconsiders the old idea that the mean field limit in the bosonic setting is actually a semiclassical limit in infinite dimension. In previous works, they proved results which are general with respect to the initial data but hold for bounded interactions. They also proved that the dynamics of the BBGKY hierarchy, often used within the mean field theory, is actually a projected picture of the dynamics of Wigner measures in the infinite dimensional phase-space. Recently in [10], they obtained such results for
singular pair interaction potentials, which include the Coulombic (attractive or repulsive) case in dimension $d = 3$. Regarding the modelling of Bose-Einstein condensates, F. Nier and collaborators have developed in [9] complete analysis of the nonlinear adiabatic ansatz proposed by physicists for the simulation of rotating Bose-Einstein condensates.

5.2. Computational Statistical Physics

Participants: Matthew Dobson, Claude Le Bris, Frédéric Legoll, Tony Lelièvre, Francis Nier, Stefano Olla, Grigorios Pavliotis, Giovanni Samaey, Gabriel Stoltz.

The extremely broad field of Molecular dynamics (MD) is a domain where the MICMAC project-team, originally more involved in the quantum chemistry side, has invested a lot of efforts in the recent years. Molecular dynamics may also be termed computational statistical physics since the main aim is to numerically estimate average properties of materials as given by the laws of statistical physics. The project-team studies both deterministic and probabilistic techniques used in the field.

5.2.1. Free Energy calculations

For large molecular systems, the information of the whole configuration space may be summarized in a few coordinates of interest, called reaction coordinates. An important problem in chemistry or biology is to compute the effective energy felt by those reaction coordinates, called free energy.

In the article [51], Tony Lelièvre, Mathias Rousset and Gabriel Stoltz study the application of constrained Langevin dynamics to the computation of free energy differences, by thermodynamic integration techniques and fluctuation relation (à la Jarzynski).

The work by T. Lelièvre and K. Minoukadeh on the longtime convergence of the ABF method in a particular bi-channel scenario (which was already mentioned in last year’s activity report) has been accepted for publication [50]. Likewise, the work by Nicolas Chopin (CREST, ENSAE), T. Lelièvre and G. Stoltz on application of the ABF method to Bayesian inference is about to appear, see [30].

5.2.2. Sampling trajectories

There exist a lot of methods to sample efficiently Boltzmann-Gibbs distributions. The situation is much more intricate as far as the sampling of trajectories (and especially metastable trajectories) is concerned.

In [32], T. Lelièvre and D. Pommier, in collaboration with F. Cérou and A. Guyader (INRIA Rennes, ASPI) investigated the interest of an Adaptive Multilevel Splitting algorithm to compute reactive paths, and estimate transition rates. The obtained results are very interesting. Current research aims at testing the technique on practical cases.

In [49], C. Le Bris and T. Lelièvre together with M. Luskin and D. Perez have proposed a mathematical analysis of the parallel replica algorithm, introduced by A. Voter in 1997 to efficiently simulate metastable trajectories. This work opens a lot of perspectives using a generic tool (the quasi stationary distribution) to relate continuous state space dynamics (Langevin type dynamics) to discrete state space dynamics (kinetic Monte Carlo type models). A follow-up work consists in theoretically investigating another related approach, the hyperdynamics method.

5.2.3. Nonequilibrium systems

The efficient simulation of molecular systems is known to be a much more complicated problem when the system is subjected to a non-conservative external forcing than when the system experiences conservative forces. Together with the sampling of metastable dynamics mentioned above, these are the two major research focus in MD of the project-team.
On this topic, G. Stoltz continued his long lasting collaboration with physicists at CEA/DAM on reduced models for shock and detonation waves. More precisely, he published two works applying some simulation techniques he devised to actual energetic materials of interest to physicists, namely (i) a technique to sample constraints in average and allowing to compute the Hugoniot curve efficiently, which was applied to reacting TATB [24]; and (ii) a reduced stochastic dynamics to model detonation waves, applied to a material with properties close to nitromethane and allowing an atomistic simulation of the shock-to-detonation transition [55].

F. Legoll and G. Stoltz pursued their studies of the anomalous thermal conductivity of one dimensional chains. They have investigated the case of a chain of rotors subjected to a mechanical forcing. In collaboration with A. Iacobucci and S. Olla (CEREMADE, Paris Dauphine) they have shown in [44] that the mechanical forcing can have a counter-intuitive effect and reduce the thermal current. Besides this numerical study, G. Stoltz, in collaboration with C. Bernardin (ENS Lyon) considered in [18] the issue of thermal transport in one of the simplest possible one dimensional model, a chain of oscillators whose kinetic and potential energy functions are the same, and which are subjected to a stochastic noise exchanging all the variables. The system therefore has only two conservation laws, the energy and the total length. A hydrodynamic limit consisting of a system of conservation laws can be obtained before the onset of shocks. However, the thermal transport is anomalous: this can be proved by analytical computations for harmonic interactions, or demonstrated numerically in the general case.

G. Stoltz also studied techniques to compute the viscosity of fluids using steady state nonequilibrium dynamics with an external nongradient bulk forcing, in the framework of the context of the PhD of Rémi Joubaud, see [45]. The two authors have proved a linear response result, and obtained asymptotic scalings of the viscosity in terms of the friction coefficients of the underlying Langevin dynamics. G. Stoltz and G. Pavliotis are now extending the results to the case of time dependent nongradient external forcings.

Nonequilibrium molecular dynamics simulations can also be used to compute the constitutive relation between the strain rate and stress tensor in complex fluids. This is fulfilled simulating molecular systems subject to a steady, non-zero macroscopic flow at a given temperature. Starting from a bath model, M. Dobson, F. Legoll, T. Lelievre, and G. Stoltz have derived a Langevin-type dynamics for a heavy particle in a non-zero background flow. The resulting dynamics, which is theoretically obtained when a unique large particle is considered, is numerically observed to also perform well when a system of many interacting particles within shear flow is considered.

5.2.4. Effective dynamics

For a given molecular system, and a given reaction coordinate \( \xi : \mathbb{R}^n \to \mathbb{R} \), the free energy completely describes the statistics of \( \xi(X) \) when \( X \in \mathbb{R}^n \) is distributed according to the Gibbs measure. On the other hand, obtaining a correct description of the dynamics along \( \xi \) is complicated.

F. Legoll and T. Lelièvre have continued their work on the definition and the analysis of a coarse-grained dynamics that approximates \( \xi(X_t) \), when the state of the system \( X_t \) evolves according to the overdamped Langevin equation (which is ergodic for the Gibbs measure). The aim is to get a coarse-grained description giving access to some dynamical quantities, such as residence times in metastable basins. These basins are usually assumed to be completely described by \( \xi \). They have proposed an effective dynamics, which is derived using conditional expectations. The first accuracy result, obtained in 2010, is phrased in terms of an estimate on the relative entropy between the law of \( \xi(X_t) \) and the law of its approximation, at any time \( t \) (see [63] for a description of the results in a simple case). If an appropriate time-scale separation is present in the system, then the effective dynamics is accurate in the sense of time-marginals. The obtained numerical results showed that this dynamics can also be used to accurately compute residence times in potential energy wells, and thus seem to be accurate in a much stronger sense. Together with S. Olla, they have started to analyze the pathwise accuracy of the proposed coarse-grained dynamics.

The extension of the numerical strategy to the case when the reference dynamics on the whole system is the Langevin dynamics has also been studied. Promising numerical results have already been obtained in collaboration with a master’s student (F. Galante).
5.2.5. Convergence to equilibrium

An important question for the analysis of sampling techniques is the rate of convergence to equilibrium for stochastic trajectories.

F. Nier continues to investigate the spectral properties of Witten Laplacians and Kramers-Fokker-Planck operators. In a recent collaboration with D. Le Peutrec and C. Viterbo, the Arrhenius law for metastable states, in its refined version also known as Eyring-Kramers law, was extended to $p$-forms. Very accurate analytical results have been provided for the exponentially small eigenvalues of Witten Laplacians acting on $p$-forms.

F. Nier has started a collaboration with T. Lelièvre about accurate exit laws for Smoluchowski processes, via a Witten complex approach. Investigations with G. Pavliotis and T. Lelièvre have started about non gradient diffusion systems.

5.2.6. Hamiltonian dynamics

Constant energy averages are often computed as long time limits of time averages along a typical trajectory of the Hamiltonian dynamics. One difficulty of such a computation is the presence of several time scales in the dynamics: the frequencies of some motions are very high (e.g. for the atomistic bond vibrations), while those of other motions are much smaller. This problem has been addressed in a two-fold manner.

Fast phenomena are often only relevant through their mean effect on the slow phenomena, and their precise description is not needed. Consequently, there is a need for time integration algorithms that take into account these fast phenomena only in an averaged way, and for which the time step is not restricted by the highest frequencies. In [38], M. Dobson, C. Le Bris, and F. Legoll developed integrators for Hamiltonian systems with high frequencies. The integrators were derived using homogenization techniques applied to the Hamilton-Jacobi PDE associated to the Hamiltonian ODE. This work extends previous works of the team. The proposed algorithms can now handle the case when the (unique) fast frequency depends on the slow degrees of freedom, or when there are several fast constant frequencies.

Another track to simulate the system for longer times is to resort to parallel computations. An algorithm in that vein is the parareal in time algorithm. It is based on a decomposition of the time interval into subintervals, and on a predictor-corrector strategy, where the propagations over each subinterval for the corrector stage are concurrently performed on the processors. Using a symmetrization procedure and/or a (possibly also symmetric) projection step, C. Le Bris and F. Legoll, in collaboration with X. Dai and Y. Maday, have introduced several variants of the original plain parareal in time algorithm [35]. These variants, compatible with the geometric structure of the exact dynamics, are better adapted to the Hamiltonian context.

5.3. Complex fluids

Participants: David Benoit, Sébastien Boyaval, Claude Le Bris, Tony Lelièvre.

In [48], Claude Le Bris and Tony Lelièvre have reviewed the state-of-the-art of numerical and mathematical results on micro-macro models for viscoelastic fluids.

Following previous works, Claude Le Bris and Tony Lelièvre together with Lingbing He have analyzed in [43] the longtime behaviour of nematic polymeric fluids (liquid crystals). The longtime asymptotics for such models is much richer than for flexible polymers, which were considered in previous analysis. Indeed, for these models, periodic in time behaviours are expected.

In his PhD under the supervision of Claude Le Bris and Tony Lelièvre, David Benoit studies models of aging fluids developed at the ESPCI (Ecole supérieure de physique et de chimie industrielles) and designed to take into account phenomena such as shear thinning, aging and shear banding in falling sphere experiments. The work consists on the one hand in studying the mathematical well-posedness of some macroscopic models and on the other hand in trying to understand the link between such macroscopic models and microscopic models which have been proposed to describe such fluids.
In the line of his former work [15], Sébastien Boyaval has pursued his work about the mathematical modelling of viscoelastic fluid flows. A new reduced model for thin layers of elastic gravity flows with a free surface was derived in collaboration with François Bouchut, following similar hydrostatic assumptions to those which lead to the Saint-Venant equations as a usual reduced model for shallow water flows. The model is naturally endowed with an energy but the conservative part is non-standard: the energy is not convex with respect to the conservative variables. It is convex with respect to other (more physical) variables. For the numerical simulation of possibly discontinuous solutions, a relaxation scheme is proposed in order to ensure that the numerical approximation mimicks the natural energy dissipation.

In [58], Tony Lelièvre together with Giovanni Samaey and Vincent Legat explored some numerical techniques to get closed macroscopic equations from microscopic models. The proposed method can be seen as a way to justify and extend techniques based on the so-called quasi-equilibrium approximation.

In [40], the effect of a topography on a free surface flow is studied using the free-surface Navier Stokes equations and ALE method for discretization.

### 5.4. Application of greedy algorithms

**Participants:** Sébastien Boyaval, Eric Cancès, Virginie Ehrlacher, Tony Lelièvre.

The manuscript [25] where E. Cancès, V. Ehrlacher and T. Lelièvre study the convergence of greedy algorithms to nonlinear convex problems has been accepted for publication. Current research now aims at extending such techniques to non-symmetric problems.

S. Boyaval has continued his study of a new variance reduction method introduced with T. Lelièvre. It makes an innovative use of a greedy algorithm in order to construct control variates for a parameterized family of Monte-Carlo estimator. New successful applications are typical problems in uncertainty quantification.

### 5.5. Homogenization

**Participants:** Ronan Costaouec, Claude Le Bris, Frédéric Legoll, Francis Nier, Florian Thomines.

The project-team has pursued its efforts in the field of stochastic homogenization of elliptic equations. The various contributions of the team, which aim at designing numerical approaches that both are practically relevant and keep the computational workload limited, have been presented from a unified perspective in [59].

An interesting case in that context is when the randomness comes as a small perturbation of the deterministic case. As previously shown by former works of the project-team, this situation can indeed be handled with a dedicated approach, which turns out to be far more efficient than the standard approach of stochastic homogenization. This previous analysis was performed manipulating the exact correctors, solutions to PDEs posed on \( \mathbb{R}^d \). In practice, one can only consider a truncated version of these corrector problems, which is next discretized using e.g. a Finite Element method. The previous analysis has been extended to this practical situation by R. Costaouec [31].

In the work mentioned above, the perturbation to the deterministic case is supposed to be small in the \( L^\infty \) norm (that is, almost surely small). The team has also considered the case when the perturbation is small in a weaker norm, typically a \( L^p \) norm with \( p < \infty \) (the case when only the expectation of the perturbation is assumed to be small, rather than the perturbation itself, is covered by that framework). The approach proves to be very efficient from a computational viewpoint, in comparison to the standard approach of stochastic homogenization, as shown in [13], [12]. In that setting, the computation of the homogenized matrix requires repeatedly solving a corrector-like equation for various configurations of the material. For this purpose, C. Le Bris and F. Thomines have shown how to adapt the Reduced Basis approach to the specific context, to even further reduce the computational cost [62].
The team has also proceeded to address, from a numerical viewpoint, the case when the randomness is not small. In that case, using the standard homogenization theory, one knows that the homogenized tensor, which is a deterministic matrix, depends on the solution of a stochastic equation, the so-called corrector problem, which is posed on the whole space \( \mathbb{R}^d \). This equation is therefore delicate and expensive to solve. In practice, the space \( \mathbb{R}^d \) is truncated to some bounded domain, on which the corrector problem is numerically solved. In turn, this yields a converging approximation of the homogenized tensor, which happens to be a random matrix. For a given truncation of \( \mathbb{R}^d \), R. Costaouec, C. Le Bris and F. Legoll, in collaboration with X. Blanc (CEA), have studied how to reduce the variance of this matrix, using the technique of antithetic variables, which is a method widely used in other fields of application. Its efficiency in the context of stochastic homogenization has been extensively studied, both numerically and theoretically [60], [19]. R. Costaouec, C. Le Bris and F. Legoll are currently investigating the possibility to use other variance reduction approaches, such as control variate techniques.

From a numerical perspective, the Multiscale Finite Element Method is a classical strategy to address the situation when the homogenized problem is not known, or when the scale of the heterogeneities, although small, is not considered to be zero (and hence the homogenized problem cannot be considered as a sufficiently accurate model). The extension of this strategy to the stochastic case, when the tensor describing the properties of the material is the sum of a periodic term and a small random term, has been studied by C. Le Bris, F. Legoll and F. Thomines [46]. A method with a much smaller computational cost than the original MsFEM in the stochastic setting has been proposed. Provided the stochastic perturbation is indeed small, the proposed method is as accurate as the original one. The work [46] also provides a complete analysis of the approach, extending that available for the deterministic setting. Such an analysis often relies on the rate of convergence of the two scale expansion (in the sense of homogenization theory) of the solution to the highly oscillatory elliptic partial differential equation. The result is classic for periodic homogenization. In generic stochastic homogenization, the rate can be arbitrary small, depending on the rate with which the correlations of the random coefficient vanish. In [47], such a result has been established for weakly stochastic homogenization. This result is a key ingredient for the numerical analysis of the MsFEM approach proposed in [46].

Still in the framework of the Multiscale Finite Element approach, F. Thomines has further investigated, in collaboration with Y. Efendiev and J. Galvis (Texas A&M University), the use of Reduced Basis methods. They have considered an extension of the MsFEM approach, well suited to the high contrast case, i.e. the case when the ratio between the maximum and the minimum values of the heterogeneous coefficient is large. The main idea of this extension is to complement the standard MsFEM basis functions with the eigenfunctions (associated to the first small eigenvalues) of a local eigenproblem. In [39], Y. Efendiev, J. Galvis and F. Thomines have considered the case when the problem depends on an additional parameter, and shown how to use the Reduced Basis approach to more efficiently compute the eigenfunctions mentioned above.

The theoretical results obtained by the team on variance reduction [19] and on the rate of convergence of the two scale expansion of the solution to a highly oscillatory, weakly random PDE [47], both rely on asymptotic properties of the Green function of the elliptic operator \( Lu = -\text{div} (A \nabla u) \), where \( A \) is a periodic, coercive and bounded matrix. In collaboration with X. Blanc (CEA) and A. Anantharaman, F. Legoll has established some results of this kind, which has been significantly extended in the recent years with some results already given in the literature.

All the works previously mentioned are concerned with elliptic PDEs. F. Nier has studied various problems in the context of wave propagation in random heterogeneous media. In collaboration with S. Breteaux (PhD student in Rennes), he has derived a Boltzmann type equation from first principles of quantum mechanics, using the bosonic QFT presentation of Gaussian random fields. Various extensions of this result are currently under investigation.

### 5.6. Atomistic to continuum methods

**Participants:** Matthew Dobson, Claude Le Bris, Frédéric Legoll.
The project-team has continued their theoretical and numerical efforts on the general topic of “passage from the atomistic to the continuum”. This concerns theoretical issues arising in this passage but also the development and the improvement of numerical simulations coupling the two scales.

The quasicontinuum method couples an atomistic model to a continuum approximation in order to compute deformed states of a crystalline lattice at a reduced computational cost compared to a full atomistic simulation. In collaboration with M. Luskin (University of Minnesota) and C. Ortner (Warwick), M. Dobson analyzed the use of numerical solvers for approximating solutions to the equilibrium equations of the force-based quasicontinuum method [37]. In particular, it was shown that a previously-proposed modified conjugate gradient algorithm has unstable modes since the linearized operator is generally not positive-definite. Based on observed properties of the spectrum, convergence rates are given for the GMRES method applied to the operator.

Still in the framework of quasicontinuum methods, several consistent couplings have been proposed in the literature in the past years. M. Dobson showed the impossibility of constructing higher-order consistent couplings for quasicontinuum energies [36]. The analysis is performed in the one-dimensional situation, and is based on the fact that the truncation error gives lower bounds on the global error on the deformation gradient. A consequence of this result is that the so-called quasi-nonlocal energy (which is one of the coupling schemes proposed earlier) has asymptotically optimal error bounds.

The team has also addressed questions related to the finite temperature modeling of atomistic systems and derivation of coarse-grained descriptions, such as canonical averages of observables depending only on a few variables. In the one-dimensional setting, an efficient strategy that bypasses the simulation of the whole system had been proposed in 2010. In collaboration with X. Blanc (CEA), F. Legoll has extended this strategy to the so-called membrane setting [20]: the system is composed of atoms that lie on a two-dimensional lattice, and have a unique degree of freedom, representing their height. The strategy can also be used to derive the stress-strain relation for one-dimensional chains of atoms, e.g. the relation between the elongation of the chain and the stress, at any given temperature [63].

5.7. Surface chemistry and electrochemistry

Participant: Ismaila Dabo.

Rising energy imperatives have revived strong interest in electrochemical cells (e.g., fuel cells) and photoelectrochemical cells (e.g., dye-sensitized solar cells). Understanding and improving the electrical performance of such systems entails the accurate description of electrode-electrolyte interfaces at the microscopic level. Despite recent advances in the application of computational methods to study realistic electrode-electrolyte interfaces, capturing the effect of the applied electrical voltage and solvent electrical response remains a central challenge in computational electrochemistry. In order to address this difficulty, a comprehensive model for chemical systems embedded in ionic media has been developed. The model couples a quantum description of the electrode with a continuum representation of the electrolyte.

This year, in collaboration with O. Andreussi (MIT) and N. Marzari (University of Oxford), I. Dabo has worked on improving the predictive accuracy and numerical implementation of the continuum solvation model to describe molecular systems and metallic surfaces in interaction with a solvent [14]. The revised model overcomes some of the numerical limitations encountered in existing solvation models and extends their range of applicability. The approach proceeds by recasting the problem in terms of induced polarization charges that act as a direct mapping of the continuum dielectric solvent. The model is defined in a self-consistent manner in terms of the first-principles electronic density of the solute, thereby limiting the number of numerical parameters involved in existing solvation methods. The model accounts for additional pressure and cavitation contributions. The resulting self-consistent continuum solvation (SCCS) model provides an effective, compact fit of computational and experimental data, with solvation energies in error of 0.3-0.4 kcal/mol. The model is implemented in the widely used open-source program QUANTUM ESPRESSO, exploiting a numerical approach that is intrinsically parallel, robust, and straightforward to adapt to most electronic-structure codes.
In terms of practical applications in surface science and electrochemistry, the adsorption of arsenate (a severe contaminant in drinking water) at the surface of biogeochemical minerals has been studied in collaboration with M. Blanchard and G. Morin (Université Paris 6) [22]. The simulation of infrared spectroscopic experiments for carbon monoxide (a severe catalytic poison) at catalytic surfaces has also been addressed in [33]. Future work includes the generalization of the SCCS model to surfaces in contact with a reservoir of electrons, and molecular dynamics simulations of electrode surfaces.
5. New Results

5.1. Mixture models

5.1.1. Taking into account the curse of dimensionality

**Participant:** Stéphane Girard.

**Joint work with:** Bouveyron, C. (Université Paris 1), Celeux, G. (Select, INRIA).

In the PhD work of Charles Bouveyron (co-advised by Cordelia Schmid from the INRIA LEAR team) [53], we propose new Gaussian models of high dimensional data for classification purposes. We assume that the data live in several groups located in subspaces of lower dimensions. Two different strategies arise:

- the introduction in the model of a dimension reduction constraint for each group
- the use of parsimonious models obtained by imposing to different groups to share the same values of some parameters

This modelling yields a new supervised classification method called High Dimensional Discriminant Analysis (HDDA) [4]. Some versions of this method have been tested on the supervised classification of objects in images. This approach has been adapted to the unsupervised classification framework, and the related method is named High Dimensional Data Clustering (HDDC) [3].

In collaboration with Gilles Celeux and Charles Bouveyron, we have designed an automatic selection of the discrete parameters of the model [12]. Also, the description of the R package is submitted for publication [44].

5.1.2. A new family of multivariate heavy-tailed distributions with variable marginal amounts of tailweight: Application to robust clustering

**Participants:** Florence Forbes, Darren Wraith.

We proposed a family of multivariate heavy-tailed distributions that allow variable marginal amounts of tailweight. The originality comes from the eigenvalue decomposition of the covariance matrix in the traditional Gaussian scale mixture representation. By contrast to most existing approaches, the derived distributions can account for a variety of shapes and have a simple tractable form with a closed-form probability density function whatever the dimension. We examined a number of properties of these distributions and illustrate them in the particular case of Pearson type VII and $t$ tails. For these latter cases, we provided maximum likelihood estimation of the parameters and illustrated their modelling flexibility on clustering examples for several simulated and real data sets.

5.2. Markov models

5.2.1. Variational approach for the joint estimation-detection of Brain activity from functional MRI data

**Participants:** Florence Forbes, Lotfi Chaari, Thomas Vincent.

**Joint work with:** Michel Dojat (Grenoble Institute of Neuroscience) and Philippe Ciuciu from Neurospin, CEA in Saclay.
In standard fMRI within-subject analysis, two steps are generally performed separately: detection and estimation. Because these two steps are inherently linked, we proposed in this work a joint detection-estimation procedure. We adopt the so-called region-based Joint Detection Estimation (JDE) framework that deals with spatial dependencies between voxels belonging to the same functionally homogeneous parcel in the mask of the 3D brain. After building a spatially adaptive General Linear Model, prior information is introduced and a hierarchical Bayesian model is established. In contrast to previous works that use Markov Chain Monte Carlo (MCMC) techniques to approximate the resulting intractable posterior distribution, we recast the JDE into a missing data framework and derive a Variational Expectation-Maximization (VEM) algorithm for its inference. It follows a new algorithm that exhibits interesting properties compared to the previously used MCMC-based approach. Experiments on artificial and real data show that VEM-JDE is robust to model mis-specification and provides computational gain while maintaining good performance. Corresponding papers [27], [38], [26].

5.2.2. Adaptive experimental condition selection in event-related fMRI

Participants: Florence Forbes, Christine Bakhous, Lotfi Chaari, Thomas Vincent, Thomas Vincent.

Joint work with: Michel Dojat (Grenoble Institute of Neuroscience) and Philippe Ciuciu from Neurospin, CEA in Saclay.

Standard Bayesian analysis of event-related functional Magnetic Resonance Imaging (fMRI) data usually assumes that all delivered stimuli possibly generate a BOLD response everywhere in the brain although activation is likely to be induced by only some of them in specific brain areas. Criteria are not always available to select the relevant conditions or stimulus types (e.g. visual, auditory, etc.) prior to estimation and the unnecessary inclusion of the corresponding events may degrade the results. To face this issue, we propose within a Joint Detection Estimation (JDE) framework, a procedure that automatically selects the conditions according to the brain activity they elicit. It follows an improved activation detection that we illustrate on real data.

5.2.3. Finding Audio-Visual Events in Informal Social Gatherings

Participant: Florence Forbes.

Joint work with: Xavier Alameida-Pineda and Radu Horaud from the INRIA Perception team.

In this work [21] we addressed the problem of detecting and localizing objects that can be both seen and heard, e.g., people. This may be solved within the framework of data clustering. We proposed a new multimodal clustering algorithm based on a Gaussian mixture model, where one of the modalities (visual data) is used to supervise the clustering process. This was made possible by mapping both modalities into the same metric space. To this end, we fully exploited the geometric and physical properties of an audio-visual sensor based on binocular vision and binaural hearing. We proposed an EM algorithm that is theoretically well justified, intuitive, and extremely efficient from a computational point of view. This efficiency makes the method implementable on advanced platforms such as humanoid robots. We described in detail tests and experiments performed with publicly available data sets that yield very interesting results.

5.2.4. Spatial risk mapping for rare disease with hidden Markov fields and variational EM

Participants: Lamiae Azizi, Florence Forbes, Senan James Doyle.

Joint work with: David Abrial and Myriam Garrido from INRA Clermont-Ferrand-Theix.

We recast the disease mapping issue of automatically classifying geographical units into risk classes as a clustering task using a discrete hidden Markov model and Poisson class-dependent distributions. The designed hidden Markov prior is non standard and consists of a variation of the Potts model where the interaction parameter can depend on the risk classes. The model parameters are estimated using an EM algorithm and the mean field approximation. This provides a way to face the intractability of the standard EM in this spatial context, with a computationally efficient alternative to more intensive simulation based Monte Carlo Markov Chain (MCMC) procedures. We then focus on the issue of dealing with very low risk values and small numbers of observed cases and population sizes. We address the problem of finding good initial parameter values in
this context and develop a new initialization strategy appropriate for spatial Poisson mixtures in the case of not so well separated classes as encountered in animal disease risk analysis. Using both simulated and real data, we compare this strategy to other standard strategies and show that it performs well in a lot of situations. Corresponding papers and communications [43], [24], [37], [25].

5.2.5. Probabilistic model definition for physiological state monitoring
Participants: Laure Amate, Florence Forbes.

Joint work with: Catherine Garbay, Julie Fontecave-Jallon and Benoit Vettier from LIG.

Assessing the global situation of a person from physiological data is a well-known difficult problem. In previous work, we proposed a system that does not produce a diagnosis but instead follows a set of hypotheses and decides of an alarming situation with this information. In this work [22], we focus on data processing part of the system taking into account the complexity and the ambiguity of the data. We propose a statistical approach with a global model based on Hidden Markov Model and we present data models that rely on classical physiological parameters and expert’s knowledge. We then learn a model that depends on the person and its environment, and we define and compute confidence values to assess the plausibility of hypotheses.

5.2.6. Solder Paste Inspection

This is joint work with VI-Technology.

The majority of defects in PCB manufacture are attributed to the stencil printing process. Stencil printing is the process where solder paste bricks are deposited on the PCB pads. Solder paste deposition is required to be accurate and repeatable, however complex physical process make this problematic. Components are placed, and their leads are pushed into the solder paste. The solder paste is then melted using, for example, reflow soldering.

Inspection can be performed before the solder paste is melted, and it is more economical to identify defects at this stage.

The evaluation of solder paste joint quality involves the analysis of a number of indicative measurements. From these measurements, potential faults are identified and inspected manually. The general challenge is to reduce of the number of potential faults by better analyzing the indicative factor measurements. That is, to improve the first pass yield (FPY) which is the percentage of total solder deposits that are good, and that do not require manual inspection. However, the ability to catch defects must be retained. Another aspect to consider is the temporal nature of the process; The mechanism for identifying faults needs to be retrained after a period of time, and so a solution must be capable of using a small training dataset.

It is important to understand and identify the factors that influence quality. The industry standard factor for measuring quality is solder volume. The precise volume is not directly observable, and so is estimated. Often, height is used as a proxy measure for solder bricks of equal area and shape. There are many other contributing factors, however not all of these can be measured directly, making accurate quality determination difficult.

Stencil printing process control attempts to adjust machine parameters according to informative factors. Online printing process control faces a similar challenge of using a limited number of measurements to inform on the quality of solder paste deposition.

We used statistical techniques to analyze such measurements. The exact nature of the work is confidential.

5.2.7. PCB defect detection
Participants: Florence Forbes, Kai Qin, Huu Giao Nguyen.

This is joint work with VI-Technology.

The objective is to detect defective components in PC Boards from image data. The exact nature of the work is confidential.
5.2.8. Statistical characterization of tree structures based on Markov Tree Models and multitype branching processes, with applications to tree growth modeling.

**Participant:** Jean-Baptiste Durand.

**Joint work with:** Pierre Fernique (Montpellier 2 University and CIRAD) and Yann Guédon (CIRAD), INRIA Virtual Plants.

The quantity and quality of yields in fruit trees is closely related to processes of growth and branching, which determine ultimately the regularity of flowering and the position of flowers. Flowering and fruiting patterns are explained by statistical dependence between the nature of a parent shoot (e.g. flowering or not) and the quantity and natures of its children shoots – with potential effect of covariates. Thus, better characterization of patterns and dependencies is expected to lead to strategies to control the demographic properties of the shoots (through varietal selection or crop management policies), and thus to bring substantial improvements in the quantity and quality of yields.

Since the connections between shoots can be represented by mathematical trees, statistical models based on multitype branching processes and Markov trees appear as a natural tool to model the dependencies of interest. Formally, the properties of a vertex are summed up using the notion of vertex state. In such models, the numbers of children in each state given the parent state are modelled through discrete multivariate distributions. Model selection procedures are necessary to specify parsimonious distributions. We developed an approach based on probabilistic graphical models to identify and exploit properties of conditional independence between numbers of children in different states, so as to simplify the specification of their joint distribution. The graph building stage was based on a Poissonian Generalized Linear Model for the contingency tables of the counts of joint children state configurations. Then, parametric families of distributions were implemented and compared statistically to provide probabilistic models compatible with the estimated independence graph.

This work was carried out in the context of Pierre Fernique’s Master 2 internship (Montpellier 2 University and AgroParisTech). It was applied to model dependencies between short or long, vegetative or flowering shoots in apple trees. The results highlighted contrasted patterns related to the parent shoot state, with interpretation in terms of alternation of flowering (see paragraph 5.2.9). This work will be continued during Pierre Fernique’s PhD thesis, with extensions to other fruit tree species and other strategies to build probabilistic graphical models and parametric discrete multivariate distributions including covariates and mixed effects.

5.2.9. Statistical characterization of the alternation of flowering in fruit tree species

**Participant:** Jean-Baptiste Durand.

**Joint work with:** Jean Peyhardi and Yann Guédon (Mixed Research Unit DAP, Virtual Plants team), Evelyne Costes and Baptiste Guitton (DAP, AFEF team), Catherine Trottier (Montpellier University)

The aim of this work was to characterize genetic determinisms of the alternation of flowering in apple tree progenies. Data were collected at two scales: at whole tree scale (with annual time step) and a local scale (annual shoot or AS, which is the portions of stem that were grown during the same year). Two replications of each genotype were available.

To model alternation of flowering at AS scale, a second-order Markov tree model was built. The ASs were of two types: flowering or vegetative. Generalized Linear Mixed Models (GLMMs) were used to model the effect of year, replications and genotypes (with their interactions with year or memories of the Markov model) on the transition probabilities. This work was the continuation of the Master 2 internship of Jean Peyhardi (Bordeaux 2 University) and was carried out in the context of the PhD thesis of Baptiste Guitton.

This PhD thesis also comprised the study of alternation in flowering at individual scale, with annual time step. To relate alternation of flowering at AS and individual scales, indices were proposed to characterize alternation at individual scale. The difficulty is related to early detection of alternating genotypes, in a context where alternation is often concealed by a substantial increase of the number of flowers over consecutive years. To separate correctly the increase of the number of flowers due to aging of young trees from alternation in flowering, our model relied on a parametric hypothesis on the base effect random slopes specific to genotype.
and replications), which translated into mixed effect modelling. Different indices of alternation were then computed on the residuals. Clusters of individuals with contrasted patterns of bearing habits were identified. Our models highlighted significant correlations between indices of alternation at AS and individual scales. The roles of local alternation and asynchronism in regularity of flowering were assessed using an entropy-based criterion, which characterized asynchronism.

As a perspective of this work, patterns in the production of children ASs (numbers of flowering and vegetative children) depending on the type of the parent AS must be analyzed using branching processes and different types of Markov trees, in the context of Pierre Fernique’s PhD Thesis (see paragraph 5.2.8).

### 5.3. Semi and non-parametric methods

#### 5.3.1. Harmony Search with Differential Mutation Based Pitch Adjustment

**Participants:** Kai Qin, Florence Forbes.

Harmony search (HS), as an emerging metaheuristic technique mimicking the improvisation behavior of musicians, has demonstrated strong efficacy of solving various numerical and real-world optimization problems. This work [36] presents a harmony search with differential mutation based pitch adjustment (HSDM) algorithm, which improves the original pitch adjustment operator of HS using the self-referential differential mutation scheme that features differential evolution - another celebrated metaheuristic algorithm. In HSDM, the differential mutation based pitch adjustment can dynamically adapt the properties of the landscapes being explored at different searching stages. Meanwhile, the pitch adjustment operator's execution probability is allowed to vary randomly between 0 and 1, which can maintain both wild and fine exploitation throughout the searching course. HSDM has been evaluated and compared to the original HS and two recent HS variants using 16 numerical test problems of various searching landscape complexities at 10 and 30 dimensions. HSDM consistently demonstrates superiority on most of test problems.

#### 5.3.2. Dynamic Regional Harmony Search Algorithm with Opposition and Local Learning

**Participants:** Kai Qin, Florence Forbes.

To deal with the deficiencies associated with the original Harmony Search (HS) such as premature convergence and stagnation, a dynamic regional harmony search (DRHS) algorithm incorporating opposition and local learning is proposed [35]. DRHS utilizes the opposition-based initialization, and performs independent HS with respect to multiple groups that are randomly recreated on a fixed period basis. Besides the traditional harmony improvisation operators, an opposition based harmony creation scheme is introduced to update the group memory. Any prematurely converged group will be restarted with the doubled size to further augment its exploration capability. Local search is periodically applied to exploit promising regions around top-ranked candidate solutions. The performance of DRHS has been evaluated and compared to HS using 12 numerical test problems at 10D and 30D, which are taken from the CEC2005 benchmark. DRHS consistently demonstrate superiority to HR over all the test problems at both 10D and 30D.

#### 5.3.3. Evolutionary algorithms with CUDA

**Participants:** Kai Qin, Federico Raimondo.

Evolutionary algorithms (EAs), inspired by natural evolution processes, have demonstrated strong efficacy for solving various real-world optimization problems, although their practical use may be constrained by their computation efficiency. In fact, EAs are inherently parallelizable due to the operations at the individual element level and population-wise evolution. However, most of the existing EAs are designed and implemented in the sequential manner mainly because hardware platforms supporting parallel computing tasks and software platforms facilitating parallel programming tasks are not prevalently available.
In recent years, the graphics processing unit (GPU) has emerged as a powerful general-purpose computation device that can favorably support massively data parallel computing tasks carried out on its hundreds of cores. The compute unified device architecture (CUDA) technology invented by NVIDIA provides an intuitive way to express parallelism and to implement parallel programs using some popular programming languages, such as C, C++ and FORTRAN. Accordingly, we can simply write a program for one data element, which gets automatically distributed across hundreds of cores for thousands of threads to execute. Although the CUDA programming model is easy-to-use, the computation efficiency of CUDA parallel programs crucially depends on careful consideration of hardware characteristics of GPUs during algorithmic design and implementation, especially about memory utilization and thread management (to maximize the occupancy of streaming multiprocessors). Without proper considerations, the parallel programs may even run slower than their sequential counterparts.

The objectives of our project are to: 1. Redesign state-of-the-art EAs using CUDA under thorough consideration of GPU’s hardware characteristics. 2. Develop a generic hardware-self-configurable EA framework, which allows automatically configuring available hardware computing resources to maximize the computation efficiency of the EA.

Currently, we had developed a memory-efficient parallel differential evolution algorithm, which features maximally utilizing the available shared memory in GPU while maximally reducing the use of the global memory in GPU considering its very limited access bandwidth. Compared with two recent parallel differential evolution algorithms implemented with CUDA in 2010 and 2011, our algorithm demonstrated significantly faster computation speed. We had also investigated the parallel implementation of test problems and provided a guideline on how to implement any user-defined test problem and combine it with an existing parallel EA framework. To the best of our knowledge, this is the first research work on this topic.

5.3.4. Modelling extremal events

Participants: Stéphane Girard, Laurent Gardes, Jonathan El-methni, El-Hadji Deme.

Joint work with: Guillou, A. (Univ. Strasbourg).

We introduced a new model of tail distributions depending on two parameters \( \tau \in [0, 1] \) and \( \theta > 0 \) [16]. This model includes very different distribution tail behaviors from Fréchet and Gumbel maximum domains of attraction. In the particular cases of Pareto type tails (\( \tau = 1 \)) or Weibull tails (\( \tau = 0 \)), our estimators coincide with classical ones proposed in the literature, thus permitting us to retrieve their asymptotic normality in an unified way. The first year of the PhD work of Jonathan El-methni has been dedicated to the definition of an estimator of the parameter \( \tau \). This permits the construction of new estimators of extreme quantiles. The results are submitted for publication [48]. Our future work will consist in proposing a test procedure in order to discriminate between Pareto and Weibull tails.

We are also working on the estimation of the second order parameter \( \rho \) (see paragraph 3.3.1). We proposed a new family of estimators encompassing the existing ones (see for instance [62], [61]). This work is in collaboration with El-Hadji Deme, a PhD student from the Université de Saint-Louis (Sénégal). El-Hadji Deme obtained a one-year mobility grant to work within the Mistis team on extreme-value statistics. The results are submitted for publication [46].

5.3.5. Conditional extremal events

Participants: Stéphane Girard, Laurent Gardes, Gildas Mazo, Jonathan El-methni.

Joint work with: J. Carreau, A. Lekina, Amblard, C. (TimB in TIMC laboratory, Univ. Grenoble I) and Daouia, A. (Univ. Toulouse I)

The goal of the PhD thesis of Alexandre Lekina is to contribute to the development of theoretical and algorithmic models to tackle conditional extreme value analysis, i.e., the situation where some covariate information \( X \) is recorded simultaneously with a quantity of interest \( Y \). In such a case, the tail heaviness of \( Y \) depends on \( X \), and thus the tail index as well as the extreme quantiles are also functions of the covariate. We combine nonparametric smoothing techniques [58] with extreme-value methods in order to obtain efficient
estimators of the conditional tail index and conditional extreme quantiles. When the covariate is random (random design) and the tail of the distribution is heavy, we focus on kernel methods [14]. We extension to all kind of tails in investigated in [45].

Conditional extremes are studied in climatology where one is interested in how climate change over years might affect extreme temperatures or rainfalls. In this case, the covariate is univariate (time). Bivariate examples include the study of extreme rainfalls as a function of the geographical location. The application part of the study is joint work with the LTHE (Laboratoire d’étude des Transferts en Hydrologie et Environnement) located in Grenoble.

More future work will include the study of multivariate and spatial extreme values. With this aim, a research on some particular copulas [1] has been initiated with Cécile Amblard, since they are the key tool for building multivariate distributions [64]. The PhD theses of Jonathan El-methni and Gildas Mazo should address this issue too.

5.3.6. Level sets estimation

Participants: Stéphane Girard, Laurent Gardes.

Joint work with: Guillou, A. (Univ. Strasbourg), Stupfler, G. (Univ. Strasbourg), P. Jacob (Univ. Montpellier II) and Daouia, A. (Univ. Toulouse I).

The boundary bounding the set of points is viewed as the larger level set of the points distribution. This is then an extreme quantile curve estimation problem. We proposed estimators based on projection as well as on kernel regression methods applied on the extreme values set, for particular set of points [10].

In collaboration with A. Daouia, we investigate the application of such methods in econometrics [41]: A new characterization of partial boundaries of a free disposal multivariate support is introduced by making use of large quantiles of a simple transformation of the underlying multivariate distribution. Pointwise empirical and smoothed estimators of the full and partial support curves are built as extreme sample and smoothed quantiles. The extreme-value theory holds then automatically for the empirical frontiers and we show that some fundamental properties of extreme order statistics carry over to Nadaraya’s estimates of upper quantile-based frontiers.

In the PhD thesis of Gilles Stupfler (co-directed by Armelle Guillou and Stéphane Girard), new estimators of the boundary are introduced. The regression is performed on the whole set of points, the selection of the “highest” points being automatically performed by the introduction of high order moments. The results are submitted for publication [51].

5.3.7. Quantifying uncertainties on extreme rainfall estimations

Participants: Laurent Gardes, Stéphane Girard.

Joint work with: Carreau, J. (Hydrosciences Montpellier) and Molinié, G. from Laboratoire d’Etude des Transferts en Hydrologie et Environnement (LTHE), France.

Extreme rainfalls are generally associated with two different precipitation regimes. Extreme cumulated rainfall over 24 hours results from stratiform clouds on which the relief forcing is of primary importance. Extreme rainfall rates are defined as rainfall rates with low probability of occurrence, typically with higher mean return-levels than the maximum observed level. For example Figure 2 presents the return levels for the Cévennes-Vivarais region obtained in [14]. It is then of primary importance to study the sensitivity of the extreme rainfall estimation to the estimation method considered.

The obtained results are published in [13].

5.3.8. Retrieval of Mars surface physical properties from OMEGA hyperspectral images.

Participant: Stéphane Girard.

Joint work with: Douté, S. from Laboratoire de Planétologie de Grenoble, France and Saracco, J (University Bordeaux).
Visible and near infrared imaging spectroscopy is one of the key techniques to detect, to map and to characterize mineral and volatile (e.g. water-ice) species existing at the surface of planets. Indeed the chemical composition, granularity, texture, physical state, etc. of the materials determine the existence and morphology of the absorption bands. The resulting spectra contain therefore very useful information. Current imaging spectrometers provide data organized as three dimensional hyperspectral images: two spatial dimensions and one spectral dimension. Our goal is to estimate the functional relationship $F$ between some observed spectra and some physical parameters. To this end, a database of synthetic spectra is generated by a physical radiative transfer model and used to estimate $F$. The high dimension of spectra is reduced by Gaussian regularized sliced inverse regression (GRSIR) to overcome the curse of dimensionality and consequently the sensitivity of the inversion to noise (ill-conditioned problems). We have also defined an adaptive version of the method which is able to deal with block-wise evolving data streams [28].

5.3.9. Statistical modelling development for low power processor.

**Participant:** Stéphane Girard.

**Joint work with:** A. Lombardot and S. Joshi (ST Crolles).

With scaling down technologies to the nanometer regime, the static power dissipation in semiconductor devices is becoming more and more important. Techniques to accurately estimate System On Chip static power dissipation are becoming essential. Traditionally, designers use a standard corner based approach to optimize and check their devices. However, this approach can drastically underestimate or over-estimate process variations impact and leads to important errors.

The need for an effective modeling of process variation for static power analysis has led to the introduction of Statistical static power analysis. Some publication state that it is possible to save up to 50% static power using statistical approach. However, most of the statistical approaches are based on Monte Carlo analysis, and such methods are not suited to large devices. It is thus necessary to develop solutions for large devices integrated in an industrial design flow. Our objective to model the total consumption of the circuit from the probability distribution of consumption of each individual gate. Our preliminary results are published in [18].
6. New Results

6.1. Intermediate dependency generative models

Participants: Christophe Biernacki, Matthieu Marbac-Lourdelle, Vincent Vandewalle.

Defining generative models for dealing with possibly correlated categorical variables is at the core of the MODAL activity. We start by noticing that it is straightforward to build a full independent distribution $\hat{p}$ and also a full dependent one $\hat{\hat{p}}$ in the categorical situation. However, both are usually too crude for modeling most of real situations.

Our idea is to combine both extreme distributions $\hat{p}$ and $\hat{\hat{p}}$ in order to obtain a new distribution called $\tilde{p}(i)$ which is an intermediate dependent situation between full independence and full dependence and (ii) which is not degenerate. As a consequence, $\tilde{p}$ is a meaningful distribution because its particular “positioning” between $\hat{p}$ and $\hat{\hat{p}}$ directly models and reveals strength of dependency between variables.

In addition, since both $\hat{p}$ and $\hat{\hat{p}}$ are easily available for most variables types, we expect to be able to design a distribution $\tilde{p}$ for most variables types, and not also the categorical ones.

A PhD thesis started on October’11 on this topic in continuation of the Master’s thesis of Matthieu Marbac-Lourdelle [37].

6.2. Transfer learning in model-based clustering

Participants: Christophe Biernacki, Alexandre Lourme.

In many situations one needs to cluster several datasets, possibly arising from different populations, instead of a single one, into partitions with identical meaning and described by similar features. Such situations involve commonly two kinds of standard clustering processes. The samples are clustered traditionally either as if all units arose from the same distribution, or on the contrary as if the samples came from distinct and unrelated populations. But a third situation should be considered: As the datasets share statistical units of same nature and as they are described by features of same meaning, there may exist some link between the samples.

We propose a linear stochastic link between the samples, what can be justified from some simple but realistic assumptions, both in the Gaussian and in the $t$ mixture model-based clustering context ([15] and a paper in revision). In the general context (categorical or heterogeneous variables), we propose to use alternatively an entropic link between populations [17]. All these works are related to the Lourme’s PhD thesis [11].

A chapter of book about transfer learning (including clustering, classification and regression) is currently submitted for publication (joint work with Julien Jacques and Alexandre Lourme).

6.3. Block regression for variable clustering: Application to genetic data

Participants: Christophe Biernacki, Julien Jacques, Loïc Yengo.

Genome Wide Association (GWA) studies have proved the implication of numerous single nucleotides polymorphisms (SNP) in the etiology of common diseases. Nevertheless, only a small part of the expected heritability of those diseases is explained by the most significantly associated SNPs. Many researches that have been lately investigating this missing heritability have considered interactions between genes and/or environmental factors as a plausible and promising explanation. Considering all if not a large number (hundreds of thousands) of variants altogether stresses the problem of the high dimensionality that most regression-based methods cannot afford. To solve this issue one either reduces the number of variants to be analyzed (shrinkage approaches) or groups them according to a certain similarity. We introduce here a regression model that simultaneously clusterizes the variants sharing close effect size while selecting the most informative clusters. The estimation of the model parameters is proposed by maximizing the likelihood. The challenges of this research rely on finding efficient algorithms for the clustering part while studying the consistency of our estimators for which the classical asymptotic theory does not apply [33], [40].
6.4. Label switching in mixtures

Participants: Christophe Biernacki, Vincent Vandewalle.

During the last fifteen years there has been an increasing interest for using Bayesian methods in mixtures models. However, one of the principal issues of these methods is the non-identifiability of components caused by symmetric prior (whatever be the kind of variables), which makes the Gibbs outputs useless for inference; this problem is known as label switching. We propose to condition the posterior distribution by a particular numbering, not on the parameter as it is usually done, but rather on a latent partition, for which the posterior distributions are not any more strictly invariant up to a renumbering of the partition [26], [19], [32]. The importance of this asymmetry depends on the choice of partition space cutting. The challenge we address is to choose a particular cutting which is justified and also easy to compute. The idea is to use some properties of the (unavailable) completed posterior distribution.

6.5. Degeneracy in Gaussian mixtures

Participant: Christophe Biernacki.

In the case of Gaussian mixtures, unbounded likelihood is an important theoretical and practical problem. Using the weak information that the latent sample size of each component has to be greater than the space dimension, we derive a simple non-asymptotic stochastic lower bound on variances. We prove also that maximizing the likelihood under this data-driven constraint leads to consistent estimates. Currently, such results are proved in the univariate case [34]. The challenge is now not only to extend them in the multivariate situation but also to complete these theoretical results with some practical strategies for properly avoiding degeneracy in softwares devoted to such mixture estimations.

This is a joined work with Gwénaëlle Castellan.

6.6. Wavelet based clustering using mixed effects functional models

Participant: Guillemette Marot.

Curve clustering in the presence of inter-individual variability has longly been studied, especially using splines to account for functional random effects. However splines are not appropriate when dealing with high-dimensional data and can not be used to model irregular curves such as peak-like data. We propose a wavelet based clustering procedure [23] and apply it to high dimensional data. We suggest a dimension reduction step based on wavelet thresholding adapted to multiple curves and using an appropriate structure for the random effect variance, we ensure that both fixed and random effects lie in the same functional space even when dealing with irregular functions that belong to Besov spaces. In the wavelet domain our model resumes to a linear mixed-effects model that can be used for a model-based clustering algorithm and for which we develop an EM algorithm for maximum likelihood estimation. An R package curvclust implementing this procedure is under building and should be posted to the CRAN, the official website of the R software, before Dec. 2011. An article has been submitted once to Biometrics and received good reports. This paper should also be submitted again to Biometrics once curvclust is on the CRAN.

6.7. Comparison of normalisation procedures in RNA-sequencing before differential analysis

Participant: Guillemette Marot.

The continuing technical improvements and decreasing cost of next-generation sequencing technologies have made RNA sequencing (RNA-seq) a popular choice for gene expression studies in recent years. Because the data collected from such studies differ considerably from those measured using microarray technology, the statistical tools used for analysis must be adapted accordingly. In particular, several methods for the normalization of RNAseq data (removal of errors due to the small number of samples, corrections for sequence composition) have been proposed in recent years. With the Statomique Consortium, we have compared seven normalisation methods. First results are given in [28].
6.8. Comparison of peak finding methods applied to tiling array experiments

Participant: Guillemette Marot.

Scan statistics are widely used to detect peaks in tiling array experiments. An extensive analysis study of real biological data is being performed with Florent Sebbane and David Hot teams (Institut Pasteur, Lille) for the study of the Yersinia Pestis bacteria in order to find new small RNAs. First results have been presented in [31]. Given a signal composed of intensities ordered along the genome, the statistical problem is to detect peaks, taking into account the irregular design of the chips, which the biologists had chosen a few years ago. A master student (D. Thuillier) has compared different normalisation methods during a 6 months internship and improved the first analysis results presented in [31]. We also propose a local score procedure, which seems promising according to first biological results obtained. The next step is to work with Alain Célisse in order to choose a generative model on the normalised data which would enable to give appropriate initial values to the local score procedure and associate p-values to local scores.

6.9. Model-based clustering for functional data

Participants: Julien Jacques, Cristian Preda.

Two procedures for clustering functional data have been developed. The first one, published in [14], is based on a functional latent mixture model which fits the functional data in group-specific functional subspaces. By constraining model parameters within and between groups, a family of parsimonious models is exhibited which allows to fit onto various situations. An estimation procedure based on the EM algorithm is proposed for estimating both the model parameters and the group-specific functional subspaces. Experiments on real-world datasets show that the proposed approach performs better or similarly than classical clustering methods while providing useful interpretations of the groups.

The second procedure, currently submitted, is a model-based clustering procedure, defined on the basis of an approximation of the density of functional random variables [36]. As previously, the EM algorithm is used for parameter estimation and the maximum a posteriori rule provides the clusters. Simulation study and real data application illustrate the interest of this methodology.

6.10. Generative models and random graphs

Participant: Alain Célisse.

The aim is to study consistency of variational and maximum likelihood estimates built from a particular generative model of random graph where independence between the ridges of the graph is not assumed. These results are established from concentration inequalities. They have a great practical interest since they justify a posteriori intensive use of variational methods in this context.

It is a joint collaboration with Jean-Jacques Daudin and a paper is submitted [35].

6.11. Resampling and learning

Participant: Alain Célisse.

This aim is to study the $k$ nearest neighbors algorithm in binary classification in two different cases: Passive and active learning. The choice of $k$ is addressed by cross-validation (resampling). In particular, we try to discover the influence of the cutting parameter on which depends the cross-validation with the retained $k$ value.

It is a joined work with Tristan Mary-Huard [16].
6. New Results

6.1. Discontinuous Galerkin methods for Maxwell’s equations

6.1.1. DGTD-$P_p$ method based on hierarchical polynomial interpolation

Participants: Loula Fezoui, Joseph Charles, Stéphane Lanteri.

The DGTD (Discontinuous Galerkin Time Domain) method originally proposed by the team for the solution of the time domain Maxwell’s equations [14] relies on an arbitrary high order polynomial interpolation of the component of the electromagnetic field, and its computer implementation makes use of nodal (Lagrange) basis expansions on simplicial elements. The resulting method is often denoted by DGTD-$P_p$, where $p$ refers to the interpolation degree that can be defined locally i.e. at the element level. In view of the design of a $hp$-adaptive DGTD method, i.e. a solution strategy allowing an automatic adaptation of the interpolation degree $p$ and the discretization step $h$, we now investigate alternative polynomial interpolation and in particular those which lead to hierarchical or/and orthogonal basis expansions. Such basis expansions on simplicial elements have been extensively studied in the context of continuous finite element formulations (e.g. [46]) and have thus been designed with global conformity requirements (i.e. $H_1$, $H(\text{rot})$ or $\text{(div)}$) whose role in the context of a discontinuous Galerkin formulation has to be clarified. This represents one of the objectives of this study.

![Figure 3. Scattering of a plane wave by a business aircraft geometry computed by DGTD-$P_1$ method on a tetrahedral mesh. Contour lines of the amplitude of the electric field.](image)

6.1.2. DGTD-$P_pQ_k$ method on multi-element meshes

Participants: Clément Durochat, Stéphane Lanteri, Claire Scheid, Mark Loriot [Distene, Pôle Teratec, Bruyères-le-Chatel].
In this work, we study a multi-element DGTD method formulated on a hybrid mesh which combines a structured (orthogonal) quadrangulation of the regular zones of the computational domain with an unstructured triangulation for the discretization of the irregularly shaped objects. The general objective is to enhance the flexibility and the efficiency of DGTD methods for large-scale time domain electromagnetic wave propagation problems with regards to the discretization process of complex propagation scenes. As a first step, we have designed and analyzed a DGTD-$P_p Q_k$ method formulated on conforming hybrid quadrangular/triangular meshes for the solution of the 2D Maxwell’s equations.

![Figure 4. Scattering of a plane wave by a perfectly conducting cylinder computed by a DGTD-$P_2 Q_4$ method on a hybrid triangular-quadrangular mesh.](image)

### 6.1.3. DGTD-$P_p$ method for dispersive materials

**Participants:** Claire Scheid, Maciej Klemm [Communication Systems & Networks Laboratory, Centre for Communications Research, University of Bristol, UK], Stéphane Lanteri.

This work is undertaken in the context of a collaboration with the Communication Systems & Networks Laboratory, Centre for Communications Research, University of Bristol (UK). This laboratory is studying imaging modalities based on microwaves with applications to dynamic imaging of the brain activity (Dynamic Microwave Imaging) on one hand, and to cancerology (imaging of breast tumors) on the other hand. The design of imaging systems for these applications is extensively based on computer simulation, in particular to assess the performances of the antenna arrays which are at the heart of these systems. In practice, one has to model the propagation of electromagnetic waves emitted from complex sources and which propagate and interact with biological tissues. In relation with these issues, we study the extension of the DGTD-$P_p$ method originally proposed in [14] to the numerical treatment of electromagnetic wave propagation in dispersive media. We consider an approach based on an auxiliary differential equation modeling the time evolution of the electric polarization for a dispersive medium of Debye type (other dispersive media will be considered subsequently). This work comprises both theoretical aspects (stability and convergence analysis) of the resulting DGTD-$P_p$ method for the time domain Maxwell equations for dispersive media, and application aspects [35].

### 6.1.4. DGFD-$P_p$ method for the frequency domain Maxwell equations

**Participants:** Victorita Dolean, Mohamed El Bouajaji, Stéphane Lanteri, Ronan Perrussel [Laplace Laboratory, INP/ENSEEIHT/UPS, Toulouse].
For certain types of problems, a time harmonic evolution can be assumed leading to the formulation of the frequency domain Maxwell equations, and solving these equations may be more efficient than considering the time domain variant. We are studying a high order Discontinuous Galerkin Frequency Domain (DGFD-P_p) method formulated on triangular meshes for solving the 2D time harmonic Maxwell equations [16]. This work is undertaken in the context of the ANR MAXWELL project whose objective is the development of an ultra wideband georadar system for imaging the subsurface. In this context, the DGFD-P_p method that we have proposed is used as the forward solver in an inversion process for the electric permittivity [17].

6.1.5. Hybridized DGFD-P_p method

Participants: Stéphane Lanteri, Liang Li, Ronan Perrussel [Laplace Laboratory, INP/ENSEEIHT/UPS, Toulouse].

One major drawback of DG methods is their intrinsic cost due to the very large number of globally coupled degrees of freedom as compared to classical high order conforming finite element methods. Different attempts have been made in the recent past to improve this situation and one promising strategy has been recently proposed by Cockburn et al. [40] in the form of so-called hybridizable DG formulations. The distinctive feature of these methods is that the only globally coupled degrees of freedom are those of an approximation of the solution defined only on the boundaries of the elements. This work is concerned with the study of such Hybridizable Discontinuous Galerkin (HDG) methods for the solution of the system of Maxwell equations in the time domain when the time integration relies on an implicit scheme, or in the frequency domain. As a first step, HDGTD and HDGFD [33] methods have been developed for the solution of the 2D propagation problems.

6.1.6. Exact transparent condition in a DGFD-P_p method

Participants: Mohamed El Bouajaji, Nabil Gmati [ENIT-LAMSIN, Tunisia], Stéphane Lanteri, Jamil Salhi [ENIT-LAMSIN, Tunisia].

In the numerical treatment of propagation problems theoretically posed in unbounded domains, an artificial boundary is introduced on which an absorbing condition is imposed. For the frequency domain Maxwell equations, one generally use the Silver-Müller condition which is a first order approximation of the exact radiation condition. Then, the accuracy of the numerical treatment greatly depends on the position of the artificial boundary with regards to the scattering object. In this work, we have conducted a preliminary study aiming at improving this situation by using an exact transparent condition in place of the Silver-Müller condition. Promising results have been obtained in the 2D case and call for an extension of this work to the more challenging 3D case.

6.2. Discontinuous Galerkin methods for the elastodynamic equations

6.2.1. DGTD-P_p method for the elastodynamic equations

Participants: Nathalie Glinsky, Fabien Peyrusse.

We continue developing high order non-dissipative discontinuous Galerkin methods on simplicial meshes for the numerical solution of the first order hyperbolic linear system of elastodynamic equations. These methods share some ingredients of the DGTD-P_p methods developed by the team for the time domain Maxwell equations among which, the use of nodal polynomial (Lagrange type) basis functions, a second order leap-frog time integration scheme and a centered scheme for the evaluation of the numerical flux at the interface between neighboring elements. Recent results concern two particular points.

The first novelty is the extension of the DGTD-P_p method initially introduced in [5] to the numerical treatment of viscoelastic attenuation. For this, the velocity-stress first order system is completed by additional equations for the anelastic functions describing the strain history of the material. These additional equations result from the rheological model of the generalized Maxwell body and permit the incorporation of realistic attenuation properties of viscoelastic material accounting for the behaviour of elastic solids and viscous fluids. In practice, one needs to add 3L additional equations in 2D and 6L in 3D, where L is the number of relaxation mechanisms of the generalized Maxwell body. This method has been implemented in 2D and validated thanks to comparisons with a FDTD method.
The second contribution is concerned with the numerical assessment of site effects especially topographic effects. The study of measurements and experimental records proved that seismic waves can be amplified at some particular locations of a topography. Numerical simulations are exploited here to understand further and explain this phenomenon. The DGTD-$\mathbb{P}_p$ method has been applied to a realistic topography of Rognes area (where the Provence earthquake occurred in 1909) to model the observed amplification and the associated frequency. Moreover, the results obtained on several homogeneous and heterogeneous configurations prove the influence of the medium in-depth geometry on the amplifications measures at the surface [26], [25].

6.3. Time integration strategies and resolution algorithms

6.3.1. Hybrid explicit-implicit DGTD-$\mathbb{P}_p$ method

Participants: Stéphane Descombes, Stéphane Lanteri, Ludovic Moya.

Existing numerical methods for the solution of the time domain Maxwell equations often rely on explicit time integration schemes and are therefore constrained by a stability condition that can be very restrictive on highly refined meshes. An implicit time integration scheme is a natural way to obtain a time domain method which is unconditionally stable. Starting from the explicit, non-dissipative, DGTD-$\mathbb{P}_p$ method introduced in [14], we have proposed the use of a Crank-Nicolson scheme in place of the explicit leap-frog scheme adopted in this method [4]. As a result, we obtain an unconditionally stable, non-dissipative, implicit DGTD-$\mathbb{P}_p$ method, but at the expense of the inversion of a global linear system at each time step, thus obliterating one of the attractive features of discontinuous Galerkin formulations. A more viable approach for 3D simulations consists in applying an implicit time integration scheme locally i.e. in the refined regions of the mesh, while preserving an explicit time scheme in the complementary part, resulting in an hybrid explicit-implicit (or locally implicit) time integration strategy. Such an approach, combining a leap-frog scheme and a Crank-Nicolson scheme, has been studied numerically in [6], showing promising results which have motivated further investigations on theoretical issues (especially, convergence in the ODE and PDE senses) [28].

6.3.2. Explicit local time stepping DGTD-$\mathbb{P}_p$ method

Participants: Joseph Charles, Julien Diaz [MAGIQUE-3D project-team, INRIA Bordeaux - Sud-Ouest], Stéphane Descombes, Stéphane Lanteri.

We have initiated this year a collaboration with the MAGIQUE-3D project-team aiming at the design of local time stepping strategies inspired from [41] for the time integration of the system of ordinary differential equations resulting from the discretization of the time domain Maxwell equations in first order form by a DGTD-$\mathbb{P}_p$ method. A numerical study in one- and two-space dimensions is underway.

6.3.3. Optimized Schwarz algorithms for the frequency domain Maxwell equations

Participants: Victorita Dolean, Mohamed El Bouajaji, Martin Gander [Mathematics Section, University of Geneva], Stéphane Lanteri, Ronan Perrussel [Laplace Laboratory, INP/ENSEEIHT/UPS, Toulouse].

We continued with the design of optimized Schwarz algorithms for the solution of the frequency domain Maxwell equations. In particular, we have analyzed a family of methods adapted to the case of conductive media [21]. Besides, we have also proposed discrete variants of these algorithms in the framework of a high order discontinuous Galerkin discretization method formulated on unstructured triangular meshes for teh solution of the 2D time harmonic Maxwell equations.

6.4. High performance computing

6.4.1. High order DGTD-$\mathbb{P}_p$ method on hybrid CPU/GPU parallel systems

Participants: Tristan Cabel, Stéphane Lanteri.
Modern massively parallel computing platforms most often take the form of hybrid shared memory/distributed memory heterogeneous systems combining multi-core processing units with accelerator cards. In particular, graphical processing units (GPU) are increasingly adopted in these systems because they offer the potential for a very high floating point performance at a low purchase cost. DG methods are particularly appealing for exploiting the processing capabilities of a GPU because they involve local linear algebra operations (mainly matrix/matrix products) on relatively dense matrices whose size is directly related to the approximation order of the physical quantities within each mesh element. We have initiated this year a technological development project aiming at the adaptation to hybrid CPU/GPU parallel systems of a high order DGTD-\$p\$ method for the numerical solution of the 3D Maxwell equations.
6. New Results

6.1. Algorithms for molecular modeling and simulation

6.1.1. Interactive quantum chemistry

Participants: Maël Bosson, Caroline Richard, Antoine Plet, Sergei Grudinin, Stéphane Redon.

We have proposed what appears to be the first algorithm for *interactive quantum chemistry simulation* at the Atom Superposition and Electron Delocalization Molecular Orbital (ASED-MO) level of theory. When drawing and editing molecular systems, interactive quantum chemistry provide immediate, intuitive feedback on chemical structures. Our method is based on the divide-and-conquer (D&C) approach, which we show is accurate and efficient for this non-self-consistent semi-empirical theory. The errors induced by the D&C approach have been studied empirically and via a theoretical study of two toy models. With this method, we have demonstrated interactive quantum chemistry simulations for systems up to a few hundred atoms on a current multicore desktop computer. As the number of cores on personal computers increases, and larger and larger systems can be dealt with, we believe such interactive simulations – even at lower levels of theory – should thus prove most useful to effectively understand, design and prototype molecules, devices and materials. This result has been published in Journal of Computational chemistry [4]. Figure 7 illustrates an interactive modeling session with a benzene molecule.

![Interactive modeling session.](image)

*Figure 7. Interactive modeling session. After breaking a benzene cycle, the user moves a hydrogen atom closer to the top carbon atom to force them to bond (a). Then, the user pulls on a carbon atom to form a fulvene molecule (b-d). Interactive electronic structure calculations allow the user to easily build plausible topologies, and get immediate feedback on the chemical structure.*

6.1.2. Adaptively Restrained Particle Simulations

Participants: Svetlana Artemova, Stéphane Redon.

Particle simulations are widely used in physics, chemistry, biology [17], [21], and even computer graphics [13]. However, many important problems still constitute significant computational challenges, including molecular docking, protein folding, diffusion across bio-membranes, fracture in metals, ion implantation, etc. Numerous methods have been developed to accelerate particle simulations, by e.g. increasing the simulation’s time step [18], [9], [14], [26], [27], [24], improving the computational complexity of the simulation [32], [15], [8], [10], [31], or simplifying the system under study [19], [29], [28], [16], [11], [31], in particular via coarse-graining methods [20], [33], [30] or multiscale and multiresolution methods [25], [22], [23], [12].
We have introduced a novel, general approach to speed up particle simulations that we call Adaptively Restrained Particle Simulations (ARPS).

Our approach works by adaptively switching positional degrees of freedom on and off repeatedly during a simulation, while letting momenta evolve. The benefits of this approach are that (a) it is mathematically grounded and is able to produce stable, energy-preserving simulations; (b) it does not require modifications to the simulated interaction potential, so that any suitable existing force-field can be directly used with ARPS; (c) under frequently-used assumptions on the interaction potential, ARPS make it possible to reduce the number of forces that have to be updated at each time step, which may significantly speed up simulations; (d) when performing constant-energy simulations, ARPS allow users to finely and continuously trade between precision and computational cost, and rapidly obtain approximate trajectories; (e) the trade-off between precision and cost may be chosen for each particle independently, so that users may arbitrarily focus ARPS on specific regions of the simulated system (e.g. a polymer in a solvent); (f) most important, when performing Adaptively Restrained Molecular Dynamics (ARMD) in the canonical (NVT) ensemble, unbiased statistics can be obtained.

We have illustrated ARPS on several numerical experiments, including a shock propagation example and a polymer-in-solvent study.

The shock propagation example demonstrates how ARPS make it possible to smoothly trade between precision and speed (Fig. 8).

The polymer-in-solvent study shows how one may collect unbiased statistics with ARPS, and demonstrates that it can be done faster than with usual (reference) simulations, The results are shown in Fig. 9.

These results have been submitted for publication.

6.1.3. Adaptive interactive quantum chemistry

Participants: Maël Bosson, Sergei Grudinin, Stéphane Redon.

We are now working on applying the adaptive paradigm to the quantum chemistry methods, to allow for interactive editing of systems of any sizes and shapes. We are developing new methods and criteria to adaptively focus the computational resources on the most relevant parts of the system. Figure 10 illustrates our recent results in this direction. In this framework, we can already achieve interactive rates and efficient virtual prototyping for systems of size up to thousand atoms on a current desktop computer.

6.1.4. Interactive molecular modeling with haptic feedback

Participants: Aude Bolopion, Barthélémy Cagneau, Stéphane Regnier, Stéphane Redon.
Figure 9. **Computing the hydrodynamic radius $R_H$ of a solvated polymer.** Traditional simulations reduce the variance more at each time step (top), but adaptively restrained (AR) simulations perform many more time steps, so that they reduce the variance faster in wall-clock time (bottom). In this example, for any target precision, AR simulations compute the hydrodynamic radius four times faster than reference simulations.

Figure 10. **Interactive electronic structure calculations in SAMSON.** In this example the system is divided in four subsystems (for which the bounded boxes are displayed). The electronic structure is adaptively updated and the geometry is being optimized while the user edits the molecular system. Because the user pulls on one atom in the left part of the system, the electronic structure is accurately recomputed for the most left subsystem (atoms in red). In the neighboring subsystem, the electronic structure is updated with a cheaper model (carbon atoms are in black and hydrogen in white). In the right part of the system, the user force do not have a sufficiently large impact and atoms as well as the electronic structure are frozen (frozen atoms are displayed in blue).
In collaboration with ISIR in Paris and LISV in Versailles, we have developed a new approach for haptic interaction with molecular systems.

Molecular interactions typically have a high dynamic range (HDR), combining short-range stiff repulsive effects with long-range, soft attractive and repulsive terms. As a result, faithful haptic rendering of such molecular interactions is both important and difficult, in particular in applications where the precise perception of molecular forces is necessary (e.g., in molecular docking simulations). Traditionally, teleoperation coupling using constant gain control schemes have limited applications since they are unable to transmit to users low attractive forces without truncating repulsive ones. Furthermore, constant scaling displacement induces either instability or time-consuming experiments (displacements are slow), which deteriorates the ease of manipulation. We have described a variable gain haptic coupling method specifically designed to render high dynamic range (molecular) forces. The proposed method has been evaluated by user tests on an experiment involving two water molecules. We have observed that variable force amplification is widely appreciated, whereas variable displacement scaling is appropriate only for users that are already familiar with haptic manipulation. A complex experiment on a HIV molecule has been carried out using this variable gain system. This approach has been published in the proceedings of the 2011 World Haptics Conference [7]. Figure 11 shows SAMSON being used with a haptic interface at ISIR.

6.2. Algorithms for molecular docking
6.2.1. Prediction of Interface Water Molecules Using a Knowledge Base

**Participants:** Georgiy Derevyanko, Sergei Grudinin.

We developed a method to predict positions for interface water molecules as part of the predicted protein-protein complex. For this purpose we used a previously developed knowledge-base scoring methodology. First, we constructed a training set of non-homologous protein complexes with interfacial water. Then, we deduced the water-protein interaction energy using this training set. And finally, we positioned water molecules around a test protein complex on a regular grid and optimized their positions according to the knowledge-based water-protein interaction energy. This method was validated in a recent CAPRI competition. Figure 12 illustrates our method on a test protein.
Figure 12. Densities of interface water molecules around a test protein computed using our water-protein knowledge-based potential.
6.2.2. Development of a Knowledge-Based Scoring Function

**Participants:** Georgiy Derevyanko, Sergei Grudinin.

We developed a new method to obtain a knowledge-based potential function for protein-protein interactions. To derive such a potential, we formulated a convex quadratic programming problem with about 1,000,000 of linear constraints and developed a fast iterative solver to solve it. We validated this scoring function in the CAPRI competition Round 24, where our prediction of the Target 50 was ranked 4th.

Figure 13 shows the use of Legendre polynomials to fit statistics obtained on the knowledge base.

![Figure 13. Using Legendre polynomials to fit protein-protein interaction statistics.](image)

6.2.3. Development of a Local Knowledge-Based Potential for Structure Optimization and Prediction of Point Mutations in a Protein

**Participants:** Petr Popov, Sergei Grudinin.

We developed and validated a method that reconstructs the shape of the binding potential function between two proteins by locations of its global minima. After, we used the obtained potential function for optimization of positions of two docking partners. We demonstrated that using our method we can significantly improve the quality of predictions of such widely-used docking algorithms as HEX and ZDOCK. We validated this method in the CAPRI competition Round 26, where our re-scoring prediction of the Target 53 was ranked 3rd.

We have also developed a method to predict the influence of point mutations on the binding affinity constant of a protein complex. First, we made point mutations and reconstructed the sidechain of the mutated residue. Then, we repacked the sidechains that are within a certain cutoff distance from the mutated residue. After, we optimized the structure of two proteins using a smooth pair-additive knowledge-based potential function. We iteratively repeated the two previous steps until convergence of the binding energy. Finally, we converted the obtained binding energy into the binding affinity constant of two proteins. We validated this method in the CAPRI competition Round 26 with the Targets 55 and 56.

Figure 14 shows an interactive docking session using a knowledge-based potential for CAPRI Round 26 Target 53.

6.3. Software engineering

6.3.1. **SAMSON’s architecture**

**Participant:** Stéphane Redon.
The data model of SAMSON has been expanded. The goal is to represent a nanosystem as the union of several interacting models: *structural models* (geometrical and topological information, to define relationships between structural elements), *dynamical models* (to define kinematical and dynamical relationships between structural components), *interaction models* (to define physical interactions between dynamical components, e.g. forces between atoms or rigid bodies), and *visual models* (visual representations, for user interaction).

All models are part of the *data graph*, which contains all the information related to the system being modeled. The referencing system has been significantly expanded, with data structures to safely handle objects creation and deletion. An *event mechanism* has been designed and added to SAMSON, so that nodes of the data graph may send messages to each other. These messages can be related to topological changes, structural changes, dynamical changes, etc.

### 6.3.2. SAMSON’s software engineering process

**Participants:** Jocelyn Gaté, Stéphane Redon.

SAMSON’s software development process has been much improved. CMAKE is used to ensure that all parts of SAMSON may easily be built on several platforms (Windows, Mac and Linux). Thanks to CMAKE, a variety of Integrated Development Environments may be used (Visual Studio, Eclipse, XCode, etc.).

CTEST and CDASH are used to test SAMSON, and the Pipol platform has been used to perform nightly builds.

### 6.3.3. Graphical User Interface design

**Participants:** Jocelyn Gaté, Stéphane Redon.

Several functionalities have been added to the graphical interface of SAMSON, including customizable toolbars (that plug-in developers will be able to modify), as well as a data graph view (Figure 15).
Figure 15. The current look of SAMSON’s user interface.

Also, because plug-ins might have complex interfaces and settings, a mechanism to save and load custom presets has been developed (Fig. 16).

Figure 16. SAMSON allows users to save and load plug-in presets.

6.4. Applications

Methods and tools developed in our group have been used in the following studies:

6.4.1. Building Blocks of Bacterial Chemoreceptor Arrays

Participant: Sergei Grudinin.
Bacterial chemoreceptors are known to cluster at the cell poles where they form partially hexagonally ordered arrays. This clusterisation is important for the function of chemotaxis system. In this study, we performed an analysis of the known structural and biochemical information on the components of chemoreceptor arrays: chemoreceptors themselves, histidine kinases and adapter proteins. Based on this analysis, we proposed a set of basic interactions within the chemotaxis system (the array building blocks) and constructed their atomistic models. The models resulting from these blocks are in agreement with experimental information and provide a basis for understanding the atomic-level structural organization of chemoreceptor arrays.

6.4.2. A Novel Dimerization Interface of Cyclic Nucleotide Binding Domain

Participant: Sergei Grudinin.

Cyclic nucleotide binding domain (CNBD) is a ubiquitous domain of effector proteins involved in signalling cascades of prokaryota and eukaryota. In this study, we described a novel CNBD dimerization interface found in crystal structures of bacterial CNG channel MlotiK1 and mammalian second messenger cAMP-activated guanine nucleotide-exchange factor Epac2. Using computational tools we demonstrated that the found interface is stable, in contrast to the dimerization interface reported previously. Comparisons with cN-bound structures of CNBD showed that the dimerization is incompatible with second messenger cAMP binding. Thus, the cAMP-dependent monomerization of CNBD may be an alternative mechanism of the cAMP sensing. Based on these findings, we proposed a model of the bacterial CNG channel modulation by cAMP.
NECS Project-Team

6. New Results

6.1. Communication and control co-design for networked systems

6.1.1. Energy-aware communication and control co-design in wireless networked control systems

Participants: C. Canudas de Wit [Contact person], N. Cardoso de Castro, F. Garin.

This work is the topic of the PhD thesis of N. Cardoso de Castro. We have considered an event-based approach to energy-efficient management of the radio chip in the sensor node of a wireless networked control system [54], [66]. Indeed, as we had pointed out in the review paper [67], the radio is the main energy consumer, and intermittent data transmission allows one to reduce the use of the radio. While the existing literature in the control community on event-based control only addresses policies using two radio-modes (Transmitting/Sleep), our work follows some considerations on the radio-chip modes well-known in the communication networks literature, and introduces some intermediate radio-modes, which consume more energy than ‘Sleep’ but allow to reach the transmitting mode consuming less energy in the transition. We propose an event-based radio-mode switching policy, which allows to perform a trade-off between energy saving and performance of the control application. To this end, a switched model describes the system, taking into account control and communication. The optimal switching policy is computed using Dynamic Programming. This work is described in [66] and in a journal paper (in preparation). A further research direction is the exploration of receding-horizon techniques (Model Predictive Control), to solve a slightly modified formulation of the same problem. This research is in collaboration with Dr. Daniel Quevedo, senior lecturer at the University of Newcastle, Australia, in particular during a three-months visit of N. Cardoso de Castro at University of Newcastle.

6.1.2. Adaptive Delta Modulation in Networked Controlled Systems with bounded disturbances

Participants: C. Canudas de Wit [Contact person], F. Gomez-Estern [University of Sevilla], F. R. Rubio [University of Sevilla].

In the context of communication and control co-design for networked systems, this work investigates the closed-loop properties of the differential coding scheme known as Delta Modulation when used in feedback loops within the context of linear systems controlled through a communication network [19]. We propose a new adaptive scheme with variable quantization step, by defining an adaptation law exclusively in terms of information available at both the transmitter and receiver. With this approach, global asymptotic stability of the networked control system is achieved for a class of controllable (possibly unstable) linear plants. Moreover, thanks to the globally defined switching policy, this architecture enjoys a disturbance rejection property that allows the system to recover from any finite–time unbounded disturbance or communication loss.

6.1.3. Control, communication, computation (3C) co-design: Multi-level classification and formulation

Participants: C. Canudas de Wit [Contact person], A. Farhadi [University of Melbourne].

We introduce here an integration framework for Control/Communication/Computation (3C) co-design based on the motivating example of fleet control of Autonomous Underwater Vehicles (AUVs) [35], [75]. Specifically, we address the problem of almost sure stability of an unstable system with multiple observations over packet erasure channel, with emphasis on coding computational complexity. We look at the tradeoff between duty cycle for feedback channel use, coding computational complexity, and performance. We compare coding computational complexity and performance for two cases: a) No feedback channel at all, and b) Feedback channel all the time. It is shown that the strategy of using feedback channel results in a better performance.
6.2. Collaborative distributed consensus algorithms for control and estimation

6.2.1. Distributed Control

**Participants:** A. Seuret [Contact person], C. Canudas de Wit, L. Briñón Arranz, G. Rodrigues de Campos, K. H. Johansson [KTH].

The first contribution in this area deals with the source-seeking problem in which the task is to locate the source of some signal using a fleet of autonomous underwater vehicles. The objective is here to use the underwater vehicles equipped with appropriate sensors as a mobile sensors network. In [28] and [29], we present a method which allows estimating the gradient of the signal propagation using a distributed consensus filters [27]. To do so, we consider a group of vehicles uniformly distributed in a fixed circular formation. We then show that this distributed consensus algorithm converges to good approximation of the gradient of the signal propagation. The algorithm takes into account the communication constraints and depends on direct signal measurements. Our approach is based on the previous results in formation control to stabilize the fleet to elastic formations which can be time-varying [29] and in a collaborative source-seeking algorithm proposed earlier by members of the team. The results are supported through computer simulations.

The second contribution on collaborative control concerns the design and analysis of a distributed algorithm whose goal is symmetric robot deployment. This activity results from the collaboration between INRIA and KTH provided by the visit of G. Rodrigues de Campos (PhD student) at KTH during six month. The objective is here to propose a hierarchical control strategy composed of two stages. The first one corresponds to an algorithm for swarm dispersion and a second concerns the design of a additional algorithm which minimizes the inter-agent angles. In this context, the behavior of each vehicle depends only on the relative positions of agents it can sense. The article submitted to ICRA’12 [84], presents some simulation examples for different configuration support the derived theoretical results.

6.2.2. Distributed Estimation

- Collaborative protocols for estimation and control

**Participants:** A. Kibangou [Contact person], A. L. F. Almeida [Universidade Federal do Ceara].

In wireless communication systems, spatial diversity plays a key role in combating signal fading arising from multipath propagation. As long as the transmitter is equipped with multiple antennas, it is well known that spatial diversity can be exploited further at the transmitter by means of space-time coding [88]. In contrast to conventional (single-user) space-time coding/decoding, when dealing with cooperative wireless networks, spatial diversity must resort to distributed space-time coding/decoding, where a collection of distributed antennas belonging to multiple terminals work in a coordinated way to encode/decode the transmitted information [85].

For this purpose, we have extended the Khatri-Rao Space time coding method proposed in [86] to cooperative networks (see Fig. 5). For cooperating nodes having a single antenna, these nodes constitute a virtual antenna array at both transmitting and receiving front-end. At each node, the received data can be viewed as slices of a third-order tensor. Therefore, retrieving the informative data is achieved by means of a CP tensor decomposition using an Alternating Least Squares (ALS) algorithm for example. When all the slices cannot be gathered at the same node, for storage resources limitations for example, a distributed ALS method can be used as in [77], which is an average consensus based method. In a consensus problem, a group of network nodes try to reach agreement on a given quantity of interest that depends on their local values [79]. Instead of using a standard consensus method where convergence is achieved asymptotically, we have proposed a finite time average consensus approach that relies on the knowledge of the graph topology. The proposed algorithm and its performance evaluation by means of simulations are described in [37].

- Kalman filtering based distributed fault detection and isolation

**Participants:** A. Kibangou [Contact person], F. Garin, S. Hachour, A. Esna Ashari.

This year, we have started a research activity related to distributed fault detection and isolation. Our first work has been the master thesis of S. Hachour on the monitoring of a solar farm.
Indeed, stimulated by increasing energy demand and ecological concerns, clean energy production with renewable resources is a key research topic that presents a largely unexplored potential of development. For this purpose, solar farms constitute power plants of the future. In such systems, electricity is produced thanks to the combined action of a large number of interconnected modules (solar panels). Each module individually produces energy, but only their interconnection allows reaching the global task of a relevant energy production. Due to the interconnection topology, a local fault on a given module can induce damageable effects on the whole network. In order to detect and localize a fault, a sensor network can be deployed over the farm. Thanks to the recent advances in wireless communications, the sensors can be equipped with wireless devices, creating a network of communicating sensors. A classical way to exploit such a network would be to create a hierarchical (tree-like) structure which conveys all measurements to a centralized computer which would analyze all data. However, a failure in a communication link or in the centralized computer would result in breakdown of the whole fault detection system, which is an unacceptable risk in an application domain of strategic importance such as a power plant. Therefore we have proposed a decentralized approach that relies on local data aggregation using the computing and communicating resources of the sensor nodes. As a consequence, nodes cooperation produces a global decision, available at each point of the network, and computable even in the case where a few sensors or links are unavailable. The monitoring procedure is achieved in a distributed way using a Kalman filtering approach. Now, by considering the sensor network as the system of interest, we try to derive distributed estimation methods that are robust to malicious entities and monitoring methods to detect anomalies in the system due to these malicious entities or malfunctioning of the network. These issues are addressed by A. Esna Ashari during his post-doctoral stay in our team.

6.2.3. Distributed Consensus

- Finite-time distributed average consensus on sensor networks
  **Participant:** A. Kibangou [Contact person].

Nowadays, several distributed estimation algorithms are based on the average consensus concept. Average consensus can be reached using a linear iterations scheme where each node repeatedly...
updates its value as a weighted linear combination of its own value and those of its neighbors. The main benefit of using a linear iterations scheme is that, at each time-step, each node only has to transmit a single value to each of its neighbors. Based on such a scheme, several algorithms have been proposed in the literature. However, in the majority of the proposed algorithms the weights are chosen so that all the nodes asymptotically converge to the same value. Sometimes, consensus can be embedded as a step of more sophisticated distributed algorithm as it is the case for the Distributed Kalman filter [80] and the Distributed Alternating Least Squares algorithm [77]. Obviously, asymptotic convergence is not suitable for these kinds of distributed methods. Even though, speed convergence of consensus algorithm have been explored in [78] and [91] with the goal to derive fast consensus algorithms, running standard consensus in finite-time constitute a source of errors not easily quantifiable. Sometimes, bounds can be derived. Therefore, it is interesting to address the question of exact consensus in finite-time.

For time-invariant network topologies and in the perfect information exchange case, i.e. without channel noise nor quantization, we have shown that the finite-time average consensus problem can be solved as a matrix factorization problem with joint diagonalizable matrices depending on the Graph Laplacian eigenvalues [38], [48]. Moreover, the number of iterations is equal to the number of distinct nonzero eigenvalues of the graph Laplacian matrix. Then, by periodically restarting the consensus algorithm, we have also shown that, in the noisy case, exact average consensus can be achieved asymptotically.

- Quadratic indices for performance evaluation of consensus algorithms

**Participants:** F. Garin [Contact person], S. Zampieri [Università di Padova], E. Lovisari [Università di Padova].

Traditional analysis of linear average-consensus algorithms studies, for a given communication graph, the convergence rate, given by the essential spectral radius of the transition matrix (i.e., the second largest eigenvalues’ modulus). For many graph families, such analysis predicts a performance which degrades when the number of agents grows, basically because spreading information across a larger graph requires a longer time. However, when considering other well-known quadratic performance indices (involving all the eigenvalues of the transition matrix), the scaling law with respect to the number of agents can be different. This is consistent with the fact that, in many applications, for example in estimation problems, it is natural to expect that a larger number of cooperating agents has a positive, not a negative effect on performance. It is natural to use a different performance measure when the algorithm is used for different purposes, e.g., within a distributed estimation or control algorithm. Examples of various relevant costs can be found in the book chapter [50] and in the references therein.

We are interested in evaluating the effect of the topology of the communication graph on performance, in particular for large-scale graphs. Motivated by the study of wireless sensor networks, our main objective is to understand the limitations which arise when agents are limited to truly local interactions, i.e., the neighborhoods are determined by being ‘near’ in a geometric (Euclidean) way, differently from graphs with few but possibly ‘distant’ connections, such as in small world models. At first [18] we consider graphs which are regular lattices (infinite lattices, or grids on tori, or grids on hyper-cubes), which are examples of geometrically local interactions, but also have a very rich structure: their symmetries allow to exploit powerful algebraic tools, such as the discrete Fourier transform over rings, to compute their eigenvalues, and then find bounds on the associated costs. Then, we extend the results to a more general class of graphs, thus showing that the behavior of lattices is mainly due to the local nature of interactions and not to the symmetries. To do so, we exploit the analogy between reversible Markov chains and resistive electrical networks. This latter work is part of the Ph.D. thesis of E. Lovisari at University of Padova, Italy.

- Distributed averaging over digital noisy networks

**Participants:** F. Garin [Contact person], R. Carli [Università di Padova], G. Como [MIT], P. Frasca [Politecnico di Torino].
We study iterative distributed averaging algorithms for networks whose nodes can communicate through memoryless erasure broadcast channels. In order to compare the performance of different algorithms, we define suitable complexity measures, which account for the number of channel transmissions (communication complexity), and, respectively, of in-node computations (computational complexity) required to achieve a desired precision. These performance measures are particularly relevant, as they allow for directly estimating the energy consumption of such distributed computation systems, as well as their time-complexity.

The algorithms we propose combine the classical iterative linear consensus algorithm (where at each iteration, each agent receives the states of its neighbors and takes a suitable convex combination of them), with source-channel coding schemes for the reliable transmission of real numbers on noisy channels. Our algorithms involve a sequence of transmission phases, of increasing duration, in which the agents attempt to broadcast their state, i.e. their current estimate of the global average, to their neighbors, alternated to averaging steps, in which the agents’ states are updated. These algorithms are fully distributed, and they do not require the agents to have any global knowledge of the network structure or size. Our main result shows that such algorithms drive the agents to state agreement (consensus) which can be made arbitrarily close to the true average. The number of channel transmissions and in-node computations is shown to grow at most poly-logarithmically in the desired precision. We also show how communication feedback, when available, allows one to modify the algorithms, achieving asymptotic average consensus (i.e., state agreement on the average of the initial observations), and reducing the computational and communication complexities. This work is presented in the paper [15]. In the paper [22], we present and analyze a modified algorithm, which can be used when source coding (compression) and channel coding (error correction) are performed by two separate encoders. Such algorithm takes into the account the fact that, even without any channel feedback, the part of the error which is due to lossy compression and not to channel noise is perfectly known by the transmitter, and a compensation can be introduced, thus improving performance.

- Distributed Consensus algorithms
  **Participant:** A. Seuret [Contact person].

Concerning this problem, we address the classical issues of the stability analysis of consensus algorithm in continuous-time. The objective of the present work is to show that the performances classical consensus algorithms can be improved using an appropriate memory of the controlled state. We want to design a novel type of consensus algorithm which uses not only the current state of the algorithm but also a sampled version of it. The key problems are here the design of the best parameters, i.e., the sampling period and the ratio between the contributions of the current and the sampled states. It has to be noticed that a usual intuition is to say that using past values of the state a reduction of performances or to instability. However, our contributions show that this is not always in single and double integrator consensus algorithms [41]. These contributions is based on an LMI framework and based on algebraic communication matrix structure. The efficiency of the method is tested for different network communication schemes.

### 6.2.4. Distributed real-time Simulation of numerical models

**Participants:** D. Simon [Contact person], A. Ben Khaled [IFPEN], M. Ben Gaid [IFPEN].

To allow real-time simulation of high fidelity engine models, different techniques have to be applied in order to fulfill the real-time constraints. Real-time simulation involves trade-offs between several aspects, such as real-time constraints, models computational complexity and integration accuracy. Traditionally HIL designers consider that every step of the simulation must be real-time and deterministic, leading to strongly synchronized systems, at the cost of ineffective computation burdens. It has been shown that adequately splitting the plant’s model into weakly synchronized sub-systems allows for efficiently using variable steps numerical integrators, simulation speed-ups and subsequent effective parallel versions of HIL systems [25].
6.3. Stability and control design of asynchronous interconnected systems

6.3.1. New approaches for stability conditions design

- Stability for asynchronous sampled data systems
  **Participants:** A. Seuret [Contact person], C. Briat [KTH], J. Gomès Da Silva Jr. [UFRGS], W. Jiang, M. M. Peet [Illinois Institute of Technology].
  During the last year an important effort has been devoted to controlled systems under communication constraints. In particular a novel approach to assess stability of continuous linear systems with sampled-data inputs has been provided for the first time in [21]. The main contribution of this article is to make the bridge between the discrete-time and the continuous-time approaches to ensure stability of the closed loop system. The interest of the method remains in the application of the discrete-time Lyapunov theorem using the continuous-time model without introducing exponential. This method suggests the introduction of particular types of functionals of several shapes: using an adaptation of classical time-delay functionals [21]; using a discretization method [26]; or using SOS [44].
  Then extensions to uncertain systems, time-varying parameter systems [21]; or non linear systems (for instance with saturations [36], [43]) become straightforward in comparison to the discrete-time approaches. The stability conditions are expressed in terms of linear matrix inequalities. Sufficient conditions for asymptotic and exponential stability are provided dealing with synchronous and asynchronous samplings and uncertain systems. An additional stability analysis is provided for the cases of multiple sampling periods and packet losses in [21]. Moreover this method has also been extended to the case of sampled-data systems with additional input delay [42], [46] and to the case of impulsive systems (several papers are submitted on this topic, for example [65]).

- Stability of control under weakened real-time constraints
  **Participants:** D. Simon [Contact person], A. Seuret, P. Andrianaiaina [AIRBUS].
  A weakened implementation scheme for real-time feedback controllers is proposed to reduce the conservatism due to traditional worst-cases considerations. To save wasted computing resources, new real-time scheduling scenarios allowed for reducing the time slots allocated to control tasks below the value of the Worst Case Execution Time which is traditionally used to implement embedded control software. The stability of the control system under occasional deadlines miss is assessed using robustness arguments, using Lyapunov-Krasovskii functionals and LMIs solving based on [46]. The methodology has been successfully assessed for a fighter aircraft pitch controller, which show that the stability of the plant can be kept (and even improved) using the new scheduling schemes using less computing resources than traditional implementations [24], [63].

6.3.2. Control for asynchronous sampled data systems

- Control architecture and tools
  **Participants:** D. Simon [Contact person], R. Pissard [SED], S. Arias [SED].
  During the development of control systems, hardware-in-the-loop that is showed in Fig. 6 takes place between design level simulations and costly experiments with the real plant. Using the prototype of ORCCAD V4 several HIL real-time simulators have been set up. These simulators combine multi-threaded/multi-rate real-time controllers running control algorithms synchronized with a variable step numerical integrator running a model of the plant [49]. These simulators have been further used to implement and test several kind of feedback/ flexible scheduling schemes related to the FeedNetBack project, in particular real-time controllers subject to (m,k)-firm scheduling, Kalman filters modified to account for data loss and varying sampling controllers [62]. Finally a collaboration with SARDIES about the integration of discrete (logical) control loops on top of continuous control tasks has been carried on. In this architecture ORCCAD is used to design the low level continuous controllers while the reactive parts are designed and synthesized using the BZR language [23].
• Event-based control algorithms

**Participants:** A. Seuret [Contact person], N. Marchand [Contact person], S. Durand.

Asynchronicity is becoming more and more meaningful in modern control architectures and some new control strategies are being developed by some research teams in the world. The principle of these control laws is to compute a new control signal only when some event occur, where a event characterizes a change in the system and therefore a need for a new control. These approaches are supposed to reduce the number of times the control is computed (and consequently the CPU utilization) and to remove the real-time hard constraint on the computational system. Some works around this domain have been proposed by some members of the NeCS team.

In [45], one may look at the problem of reducing the amount of information to be sent to the actuators through the Network. Indeed the controller may be able to trigger the information to be sent. The main idea is to let the controller decide if the system needs an update of its control input. This class of control algorithms is called event-triggered. An algorithm is suggested to sample the control input based on the behavior of a Lyapunov-like function in [45]. This algorithm is event-triggered since the Lyapunov-like function directly depends on the state of the systems.

In [69] and [73], we firstly proposed to remove the safety limit condition introduced by K-E. Årzén in his event-based PID controller [93]. In this paper, the control signal is updated only when required from a performance point of view, that is when the measurement crosses a given level. Årzén also suggested to enforce an event when the sampling interval achieves a given maximal amount of time. This safety limit was added to prevent the system to be sampled less than what Shannon theorem requires but, in fact, we showed that the Shannon sampling condition is no more consistent in the context of event-based systems. Moreover, a practical implementation of a cruise control mechanism on a small radio-controlled vehicle was recently suggested [34]. Moreover, we are interested on updating the control signal only when required from a stability point of view. Such a solution consists for instance in enforcing an event when a Lyapunov function crosses a given level. Based on the seminal work from M. Velasco in [90], we suggested a simple Lyapunov sampling in [33].

• Control with adaptive sampling

**Participants:** D. Simon [Contact person], O. Sename, E. Roche.

Control and real-time computing have been associated for a long time, for the control of industrial plants and in embedded or mobile systems, e.g. automotive and robotics. However both parts, control and computing, are often designed with poor interaction and mutual understanding. We propose here an integrated control and scheduling co-design approach, where closing the loop between the control
performance and the computing activity is promising for both adaptivity and robustness issues. We developed during the last years a variable sampling control methodology based on the LPV (Linear Parameter Varying) framework and $H_\infty$ control synthesis, where the sampling interval is used as a known and controlled variable [8]. Few assumptions about sampling are needed for this control design: the main point is that the control interval is known and lies between the predefined bounds $[h_{\text{min}}, h_{\text{max}}]$, whatever the origin of the control interval variations, its speed and its frequency. Another approach is proposed now to design sampling varying gain-scheduled controller for LPV systems, based on the Linear Fractional Representation (LFR). This method has already been studied concerning the synthesis of discrete-time gain-scheduled controller, depending only on the sampling period. The method has been extended to deal with the design of a gain scheduled LFT controller w.r.t the sampling interval and w.r.t system’s parameters, given a discrete-time Linear Fractional Representation (LFR) of the LPV varying sampling model [13]. The approach comes from the robust control theory and consists in separating the LTI part $P$ (not depending on the set of parameters) from the varying part $\Delta$ (parameters or uncertainties), as shown on Fig. 7. The matrix $\Delta$ represents the influence of the set of parameters $\rho(\cdot)$ on the plant. From this model, a gain scheduled controller can be computed, depending on the same set of parameters $\Delta$, or a subset of $\Delta$. The LFT approach proposed in [83] has been extended to set a LFR model that accounts for system and sampling parameters. In that case the uncertainty matrix is as depicted in Fig. 7 on the right, where $\delta$ is the deviation of the sampling interval w.r.t. the nominal sampling rate and $\Delta$ account for the plant’s uncertainty. A controller is then synthesized using the $H_\infty$ framework, where weighting functions allow to shape the control system’s response [82].

In the framework of the FeedNetBack IST project, the LFR formulation previously presented is applied to an Autonomous Underwater Vehicle (AUV) for the control of its altitude $z$ [40]. The global control structure is presented on Fig. 8. To control the altitude, two controllers are considered ($\hat{K}_z(\delta)$ and $\hat{K}_\theta(\delta, \rho)$) based on two models ($\hat{G}_z(\delta)$ and $\hat{G}_\theta(\delta, \rho)$). The LFR formulation is used here to keep some varying parameters into the model formulation. Indeed, in previous works, the limits of a simple linearization around a fixed equilibrium point have appeared. When the pitch angle is too far from 0 (the value chosen for $\theta$ at equilibrium), the linearized model becomes too different from the linearized one, which leads to bad performances. The $\Delta$ block contains the varying part of the model, which depends on the linearization point ($\theta_{eq}$). The two considered varying parameters are $\rho_1 = \cos(\theta_{eq})$ and $\rho_2 = \sin(\theta_{eq})$. The model is then discretized and the sampling period is added to the $\Delta$ parameter block. Compared with the previous approach using a linearized plant model, the new one shows (in simulation) improved altitude tracking and a better utilization of the available range of the actuators [82].

![Figure 7. System under LFT form.](image-url)
6.4. Vehicular transportation systems

6.4.1. Traffic modelling, estimation and control

Participants: C. Canudas de Wit [Contact person], C. Irinel-Constantin [CRAN], D. Pisarski, L. Leon Ojeda.

This part is related to the developed work within the Network of Excellence HYCON2 (Highly-Complex and Networked Control Systems). It interested to problems of modelling, estimation and control in traffic.

In [81], the problem of equilibrium sets for the Cell Transmission Model is studied. The objective is to design the homogeneous distribution of density on the freeway, where the input flows are the decision parameters to be determined. For the proper design of the balanced density the extensive analysis on the structures of equilibria is crucial. The analysis is carried out for the two different cases, where all sections of the freeway are assumed to be free or congested, respectively. The necessary conditions for the existence of balanced equilibria are formulated. These conditions show, that the key for the design of the balanced states may be the variable speed limiting, which strictly cooperates with the ramp metering. The computational algorithm for the input flows in case of free balance is proposed. In order to illustrate the results, the numerical example is provided.

In [39], the authors are interested to the highway traffic model-based density estimation. A strategy is proposed for real-time density estimation for traffic networks. To this aim, we introduced a deterministic constrained macroscopic model which reduce the number of possible affine dynamics of the system and preserve the number of vehicles in the network. This model is used to recover the state of the traffic network and precisely localize the eventual congestion front. The state of the network is recovered using what we call forward/backward observers. We pointed out that during unobservable modes the estimation error is preserved due to vehicle conservation law. Numerical simulations show the efficiency of the proposed strategy.

In [30], the problem of front congestions control is treated. For this, we have introduced a new traffic lumped model with only two cells (one free, and another congested) the cells have variable length, and a variation law for the front congestion completes the 3-dimensional model. In opposition to fixed-length cell models that are commonly represented by a set of linear state-dependent switching systems, our model results in a lower dimensional nonlinear system which solutions are continuous. Based on this model, we have designed a “best-effort” control strategy using variable speed limits. The notion of best effort control is here linked to the physical variable speed limit constraints which limits its size and as well as its rate variation. This results in a relative simple control in closed-form that can be implemented by using only information about the front congestion location.

Other work is under development and is related to the traffic show case application and the achievements reached that correspond to the operation of the freeway network around the Grenoble area (Grenoble South Ring). We started by designing the general network architecture, after specifying sensors and actuators locations along the highway and finally setting the platform of an interface between Matlab and our micro simulator "Aimsun". We have also carrying out some simulations from a real life application on Grenoble South ring of a deterministic state estimation technique. Using constrained macroscopic model which in fact reduces the number of dynamic states and preserves the conservation law (number of vehicles in the network).
6.4.2. Vehicle control

- Tele-operated control
  Participants: C. Canudas de Wit [Contact person], W. Jiang, J. Dumon, O. Sename.
  A mathematical driver model in the spatial equation form has been introduced for analysis of drivers’ behavior [31]. In the model, a previewed distance is taken into account. First, optimal control is applied. For the ideal case without driver’s reaction delay concerned, for both long distance preview and shorter one, the vehicle tracks well the path. Whereas, when time-delay added into the system, too short preview distance cause the instability of the system. The simulation result corresponds with the real driving experience. Then, Lyapunov-Krasovskii functional approach is applied deal with stability problem with the driver’s delay. In this case, when the delay becomes greater, the longer preview distance will be needed. The average derivation of the steering input is calculated for different preview distance as well as the time-delay cases, which well verifies our driver model. The main contribution of this model is that the preview effect only depends on the path information and it does not affected by the vehicle speed, so the result is more neutral

- Electric Power Steering Systems
  Participants: C. Canudas de Wit [Contact person], V. Ciarla, J. Tordesillas Illàn [UJF].
  This part presents several aspects of modeling, observation and control towards a new generation of Electrical Power Steering (EPS) systems [47]. In particular we design an optimal control to reject oscillations of the steering column, then we device a new observer to estimate the internal state variables of the steering column, the driver applied torque (steering wheel torque), and the load torque (tire/ground contact friction). Finally, we also revisited the LuGre tire dynamic friction model by improving the transient behavior between the sticking phases and the dynamic ones. Simulation of the proposed control and observer are shown at the end of the paper using the improved LuGre-tire friction model. Index Terms—Electric Power Steering systems (EPSs), LQ control, LuGre friction model, observer.

6.5. Energy-aware control of systems on chip

Participants: N. Marchand [Contact person], S. Durand.

Achieving a good compromise between computing power and energy consumption is one of the challenges in embedded architectures of the future. This management is especially difficult for 45 nm or 32 nm known to be at the limit of the scalability, i.e. with a high process variability. This is a key point in the ARAVIS project. Automatic control loops have therefore to be designed to minimize the energy consumption still making the performance fit the requirement in a context of highly technological uncertainties of the chip. This issue is notably discussed in [92]. Finally, the main objective is to dynamically control the computing activity and the energy consumption using the voltage and the frequency according to the requirements of the operating system. In this way, a robust control law was developed in [72] and [32] in order to minimize the high voltage running time with a predictive technique, i.e. to minimize the energy consumption while ensuring good computational performance. This control was initially done for one node (i.e. a processor) but in ARAVIS SoC, the chip is composed of several clusters with several nodes each (see Fig. 9). Thus, the energy controller has to manage the voltage level (one voltage domain by cluster) and the frequency for all nodes: a maximal frequency is performed for critical node and then a ratio of this frequency could be apply for the other nodes. Thus, a multicore control strategy with low computational needs was also proposed [71]. This work yields two patents: the first one for the monocore case [70] and the second one for the multicore case [74].
Figure 9. Architecture du SoC ARAVIS.
5. New Results

5.1. Model-free control

Participants: Cédric Join, Riachy Samer, Lotfi Belkoura.

After the successful implementation of model-free control [75], [81] for several concrete situations:

- Throttle control for IC engines (with APPEDGE and PSA) [94];
- Stop-and-go automotive control strategy (in collaboration with the École des Mines de Paris and PSA) [72], [114], [115];
- Hydroelectrical dams modeling and control (in collaboration with EDF) [92], [93];
- Shape memory actuators (collaboration with the team directed by Prof. E. Delaleau at the École Nationale des Ingénieurs de Brest [88], [89]);
- Model-free control involves the design of the so-called "intelligent" PID controllers [75], [82], and a mathematical explanation via "intelligent" PID controllers of the strange ubiquity of PIDs has been developed in [74], and the simulations confirm the superiority of the new intelligent feedback design;
- Application of model-free control method to set Delta hedging [76];
- Model-free control of "Planar Vertical Take-Off and Landing" (PVTOL) aircraft [108];
- Model-free control for power converters [103];
- The longitudinal control of the electrical vehicle by using model-free control technique [73];
- Model-free control for automatic water level regularization [93] and [92];
- [49] simplifies several aspects of the practical implementation of the newly introduced model-free control and of the corresponding intelligent PID controllers. Four examples with their computer simulations permit to test our techniques.

More achievements have been made in 2011, listed as follows:

- Shape Memory Alloys (SMA) are more and more integrated in engineering applications. These materials with their shape memory effect permit to simplify mechanisms and to reduce the size of actuators. Most of successful control strategies applied to SMA actuator are not often suitable for industrial applications. In [17], an application of the new framework of model-free control to a SMA spring based actuator was proposed. This control strategy is based on new results on fast derivatives estimation of noisy signals, its main advantages are: its simplicity and its robustness. Experimental results and comparisons with PI control are exposed that demonstrate the efficiency of this new control strategy.
- The regulation of freeway traffic flow, which is a complex nonlinear system, is achieved via the newly introduced model-free control in [39] and [64]. Several computer simulations are validating our control strategy, which is easy to implement and shows good robustness properties with respect to perturbations;
- After numerous successful applications, [36] has revisited some points of model-free control. The numerical differentiation of noisy signals may be replaced by a real time parameter identification which is much simpler. The strange ubiquity of classic PIDs is explained as well as the almost universal utilization of ultra-local models of order 1. We show that even with a partially known model the utilization of an intelligent PI controller remains profitable.
- The Ph.D. work of Y. El Afou allowed for experimental results on climate control in greenhouse.
5.2. Algebraic technique for estimation, differentiation and its applications

**Participants:** Cédric Join, Mamadou Mboup, Wilfrid Perruquetti, Lotfi Belkoura, Olivier Gibaru, Zoran Tiganj, Dayan Liu.

Elementary techniques from operational calculus, differential algebra, and noncommutative algebra lead to a new algebraic approach for estimation and detection. It is investigated in various areas of applied sciences and engineering. The following lists only some applications:

- [28] presents a partial derivatives estimation method for multidimensional signals. On a small interval the signal is represented by a truncated Taylor expansion. Then the application of multivariate Laplace transform together with adequate algebraic manipulations enabled us to express the desired partial derivative of any order as a function of iterated integrals of the noisy signal. Several recurrence relations and structural properties were provided. An interpretation of the estimators as least square minimization is also done by expressing the estimators in an orthogonal basis constituted by Jacobi polynomials. This projection enabled us not only to show a spatial shifting inherent to a specific class of estimators but also to synthesize a new class of estimators minimizing the truncation remainder of the Taylor local model. We provided also another class of estimators minimizing the noise influence. Finally we provided a numerical implementation scheme in the form of a finite impulse digital filters.

- A fast identification algorithm is proposed in [20] for systems with delayed inputs. It is based on a non-asymptotic distributional estimation technique initiated in the framework of systems without delay. Such a technique leads to simple realization schemes, involving integrators, multipliers and piecewise polynomial or exponential time functions. Thus, it allows for a real time implementation.

- A new approach to estimate vehicle tire forces and road maximum adherence is presented in [30]. Contrarily to most of previous works on this subject, it is not an asymptotic observer based estimation, but a combination of elementary diagnosis tools and new algebraic techniques for filtering and estimating derivatives of noisy signals. In a first step, instantaneous friction and lateral forces will be computed within this framework. Then, extended braking stiffness concept is exploited to detect which braking efforts allow to distinguish a road type from another. A weighted Dugoff model is used during these “distinguishable” intervals to estimate the maximum friction coefficient. Very promising results have been obtained in noisy simulations and real experimentations for most of driving situations.

- [27] proposes a diagnosis approach of sensor and actuator modeled as structured signals acting on a particular class of uncertain linear dynamical systems. The main advantage of this approach is that it is possible under certain assumptions, to detect, isolate and identify faults using only input and output measurements without having to identify model parameters. The method is based on the generation and analysis of analytical redundancy relations and exploits the fact that a structured signal satisfies a differential equation. The decision rule is based entirely on the temporal behaviour of the estimates of some fault characteristics.

- Numerical causal derivative estimators from noisy data are essential for real time applications especially for control applications or fluid simulation so as to address the new paradigms in solid modeling and video compression. By using an analytical point of view, [23] revisited the $n$th order algebraic derivative estimators. Thanks to a given noise level and a well-suitable integration length window, we analyzed the derivative estimator error;

- Recent algebraic parametric estimation techniques provide an estimate of the derivatives by using iterated integrals of a noisy observation signal. These algebraic parametric differentiation techniques give derivative estimations which contain two sources of errors: the bias term error and the noise error contribution. In order to reduce these errors, [25] extends the parameter domains used in the estimators, and studies some error bounds which depend on these parameters. This allows us to minimize these errors. It is shown that a compromise choice of these parameters implies an “optimized” error among the noise error contribution, the bias term error and the time delay.
The numerical differentiation by integration method based on Jacobi polynomials originally introduced by Mboup, Fliess and Join is revisited in [24] for the central case where the used integration window is centered.

[59] proposed new algebraic techniques to estimate the amplitude, frequency and phase of a biased and noisy sinusoidal signal. The methods which are popular today seem unable to obtain a robust estimation of those parameters within a fraction of the signal’s period. The efficiency of our approach is illustrated by several computer simulations;

A “practical” comparison between high-order sliding modes and the recently introduced model-free control is made in [56]. The perfect knowledge of the relative degree of the output variable, which is a standard assumption for sliding modes, is assumed. The comparisons are based on two concrete case-studies and on numerous computer simulations. The smoothness of the input variables, the robustness with respect to noises and the straightforward extendibility of the model-free controllers to MIMO systems are highlighted.

[43] and [66] present a parameter estimation algorithm for a magnetic bearing. Such process are inherently unstable systems with strongly nonlinear dynamics. A simplified model of the magnetic bearing is developed, which enables to obtain a linear expression with respect to the unknown parameters. These parameters are measurable with difficulties, and may slightly vary over time. The expression of the estimates is written as a function of integrals of the inputs and outputs of the system. The simulations and the experiments show a fast and robust on-line identification.

Estimators of the frequency, amplitude and phase of a noisy sinusoidal signal with time-varying amplitude by using the algebraic parametric techniques is studied in [52], in which a similar strategy to estimate these parameters by using modulating functions method is applied. The convergence of the noise error part due to a large class of noises is studied to show the robustness and the stability of these methods. We also show that the estimators obtained by modulating functions method are robust to “large” sampling period and to non-zero-mean noises.

In the framework of the SYSIASS project, a single landmark based localization algorithm for non-holonomic mobile robots is studied in [58]. In the case of a unicycle robot model, the localization problem is equivalent to the system observability. Based on this observation, the proposed localization method consists in finding a vector function which depends on the measurement vector and its derivatives, for which a numerical differentiation method is used in [58].

5.3. Observability and observer design for nonlinear systems

Participants: Jean-Pierre Barbot, Wilfrid Perruquetti, Lotfi Belkoura, Thierry Floquet, Gang Zheng.

Observability analysis and observer design are important issues in the field of control theory. Some recent results are listed below:

[32] investigates the observability and observer design for a class of single output switched systems with high frequency switching, where classical observers cannot be applied directly since the high frequency switching signals are not derivable. By assuming that these signals are integrable in the less restrictive way, and defining a new output, this study shows that algebraic observer can be adopted to estimate the states of the studied switched systems. Although the main idea is explained via normal forms, it can be easily extended to treat generic switched systems with high frequency switching.

Observability of a class of switched systems with Zeno phenomenon or high switching frequency is treated in [31]. Particularly, three observability forms are proposed and the observability for each form with knowledge of filtered switching signal is analyzed. Meanwhile, sufficient and necessary conditions for the existence of a diffeomorphism to transform a class of switched systems into one of such forms are presented.
• A triangular canonical form for a class of 0-flat nonlinear systems is studied in [13]. Necessary and sufficient geometrical conditions are given in order to guarantee the existence of a local diffeomorphism to transform the studied nonlinear systems into the proposed 0-flat canonical form, which enables us to compute the flat output as well.

• A fault tolerant control for induction motors based on backstepping strategy is designed in [44]. The proposed approach permits to compensate both the rotor resistance variations and the load torque disturbance. Moreover, to avoid the use of speed and flux sensors, a second order sliding mode observer is used to estimate the flux and the speed. The designed observer converges in a finite time and gives a good estimate of flux and speed even in the presence of rotor resistance variations and load torque disturbance.

• [42] studies the observability problem of a general class of singular linear systems with unknown inputs. It is shown that, under some assumptions, the problem is equivalent to study the observability of a standard linear system with unknown inputs satisfying algebraic constraints. We obtain necessary and sufficient conditions for observability in terms of the zeros of the system matrix.

• [53] is concerned with the study of observability properties of systems without inputs via homogeneous approximations. This approximation is induced by a filtration on the space of observation. A corresponding filtration on a Lie algebra of vector fields is defined and allows to construct the approximation that preserve observability properties. An explicit construction is given in [53].

5.4. Time-delay systems

Participants: Jean-Pierre Richard, Jean-Pierre Barbot, Thierry Floquet, Gang Zheng, Denis Efimov, Wilfrid Perruquetti.

• [22] considers a networked control loop, where the plant is a “slave” part, and the remote controller and observer constitute the “master”. Since the performance of Networked Control Systems (NCS) depends on the Quality of Service (QoS) available from the network, it is worth to design a controller that takes into account qualitative information on the QoS in realtime. The goal of the design is to provide a controller that guarantees two things: 1) high performances (here expressed by exponential decay rates) when the QoS remains globally the same; 2) global stability when the QoS changes. In order to guarantee the global stability, the controller will switch by respecting a dwell time constraint. The dwell time parameters are obtained by using the switched system theories and the obtained conditions are Linear Matrix Inequalities (LMI).

• Causal and non-causal observabilities are discussed in [33] for nonlinear time-delay systems with unknown inputs. Using the theory of non-commutative rings and the algebraic framework, the nonlinear time-delay system is transformed into a suitable canonical form to solve the problem. A necessary and sufficient condition is given to guarantee the existence of a change of coordinates leading to such a form.

• The notion of homogeneity is extended to the time-delay nonlinear systems in [45]. It is shown that under some conditions the stability of homogeneous functional systems on a sphere implies the global stability of the system. The notion of local homogeneity is introduced, the relations between stability of the locally approximating dynamics and the original time-delay system are established.

• [38] addresses the problem of the position/force tracking in tele-operation system and proposes a haptic proxy control scheme, in which communication delays are assumed to be both time-varying and asymmetric, and the response of the synchronization and the transparency are improved. The control design is performed using Linear Matrix Inequality (LMI) optimization based on Lyapunov-Krasovskii functionals (LKF) and $H_1$ control theory.

• Stability and synchronization of systems with time-varying delays is studied in [37], in which a novel control scheme with position/velocity information channel on the basis of Lyapunov-Krasovskii functional (LKF) and $H_1$ control theory by using Linear Matrix Inequality (LMI) is proposed. The proposed solution is efficient for different working conditions, such as abrupt motion and wall contact illustrated by various simulations.
• Embedded systems can benefit from all results on variable sampling for delayed systems [18], [19] and [47].
6. New Results

6.1. Mathematical analysis and control of macroscopic traffic flow models

6.1.1. Vehicular traffic

Participants: Maria Laura Delle Monache, Paola Goatin, Mauro Garavello [Piedmont University, Italy].

Concerning road traffic, the research activity during 2011 focused on the mathematical analysis of traffic flow models on road networks or subject to unilateral constraints. In particular, [ 34 ] is devoted to a hyperbolic 2nd order model for traffic flow with local flux constraint. We describe two admissible Riemann solvers and we construct ad hoc finite volume numerical schemes to compute these solutions. The paper [ 59 ] is devoted to the study of a traffic flow model on a network composed by an arbitrary number of incoming and outgoing arcs connected together by a node with a buffer. We define the solution to the Riemann problem at the node and we prove existence and well posedness of solutions to the Cauchy problem. Finally, a general traffic flow model with phase-transition is proposed and described in [ 28 ].

M.L. Delle Monache just started her doctoral thesis in the same topic. More precisely she will study hyperbolic models of traffic flow and associated optimization problems.

6.1.2. Crowd motion

Participants: Nora Aïssiouene, Régis Duvigneau, Nader El Khatib, Jihed Joobeur, Paola Goatin, Massimiliano D. Rosini [ICM, Warsaw University, Poland].

Concerning pedestrian motion modeling, we are interested in the optimization of facilities design, in order to maximize pedestrian flow and avoid or limit accidents due to panic situations. To this aim, we are now studying a macroscopic model for crowd movements consisting in a scalar conservation law accounting for mass conservation coupled with an Eikonal equation giving the flux direction depending on the density distribution. From the theoretical point of view, and as a first step, we are studying the problem in one space dimension (for applications, this case corresponds to a crowd moving in a corridor). In collaboration with M. Rosini (supported by the project CROM3, funded by the PHC Polonium 2011), we have established entropy conditions to select physically relevant solutions, and we have constructed explicit solutions for some simple initial data (these results are presented in [ 54 ]). We are now studying existence and uniqueness of solutions of the corresponding initial boundary value problem. From the numerical point of view, we are implementing the model in two space dimensions on triangular meshes on the Num3sis platform. This was partly done by N. El-Khatib (postdoc at INRIA from January to August 2011), and will be completed soon by Nora Aïssiouene. This will provide a performing numerical tool to solve the related optimization problems arising in the optimization of facilities design, such as the position and size of an obstacle in front of (before) a building exit in order to maximize the outflow through the door and avoid or limit over-compression. Moreover, jointly with the PULSAR team, we have supervised J. Joobeur’s internship, which was devoted to pedestrian data collection from real-word video recordings (Turin metro station). The density data will serve to validate the model.

The above researches were partially funded by the ERC Starting Grant "TRAM3 - Traffic management by macroscopic models".

6.2. Optimum design in fluid dynamics and its couplings

In computational sciences for physics and engineering, Computational Fluid Dynamics (CFD) are playing one of the major roles in the scientific community to foster innovative developments of numerical methodologies. Very naturally, our expertise in compressible CFD has led us to give our research on numerical strategies for optimum design a particular, but not exclusive focus on fluids.
6.2.1. Cooperation and competition in multidisciplinary optimization

**Participants:** Étienne Baratchart [ENSI2B MATMÉCA], Jean-Antoine Désidéri, Régis Duvigneau, Adrien Zerbinati.

The framework of our research aims to contribute to numerical strategies for PDE-constrained multiobjective optimization, with a particular emphasis on CPU-demanding computational applications in which the different criteria to be minimized (or reduced) originate from different physical disciplines that share the same set of design variables. These disciplines are often fluids, as a primary focus, coupled with some other discipline, such as structural mechanics.

Our approach to competitive optimization is based on a particular construction of Nash games, relying on a split of territory in the assignment of individual strategies. A methodology has been proposed for the treatment of two-discipline optimization problems in which one discipline, the primary discipline, is preponderant, or fragile. Then, it is recommended to identify, in a first step, the optimum of this discipline alone using the whole set of design variables. Then, an orthogonal basis is constructed based on the evaluation at convergence of the Hessian matrix of the primary criterion and constraint gradients. This basis is used to split the working design space into two supplementary subspaces to be assigned, in a second step, to two virtual players in competition in an adapted Nash game, devised to reduce a secondary criterion while causing the least degradation to the first. The formulation has been proved to potentially provide a set of Nash equilibrium solutions originating from the original single-discipline optimum point by smooth continuation, thus introducing competition gradually. This approach has been demonstrated over a testcase of aero-structural aircraft wing shape optimization, in which the eigen-split-based optimization reveals clearly superior [33].

While the two-discipline method is currently being applied to various complex physical multiobjective situations (see in particular 6.2.2, 6.2.6, 6.2.7, 6.2.8), the method has been extended to situations involving more than two objectives when the initial point is Pareto-optimal. Then, a particular convex combination of the criteria is locally stationary, and the two-discipline strategy can be applied using this combination as preponderant criterion, and a particular other criterion as secondary one. Whence, the proposed split of territory produces a continuum of Nash equilibrium points tangent to the Pareto set. This theoretical result has been illustrated in the context of a simpler numerical experiment by E. Baratchart during his internship [53], see Fig. 3.

![Figure 3. Combination of cooperative and competitive optimization algorithms: in red the Pareto set, in blue MGDA steps directed to the Pareto set, in green steps by Nash games with split of territory tangent to the Pareto set.](image-url)
Our approach to cooperative optimization is based on a result of convex analysis established for a general unconstrained multiobjective problem in which all the gradients are assumed to be known. The theorem [58] states that in the convex hull of the gradients, there exists a unique vector of minimal norm, \( \omega \); if it is nonzero, the vector \( \omega \) is a descent direction common to all criteria; otherwise, the current design point is Pareto-optimal. This result led us to generalize the classical steepest-descent algorithm by using the vector \( \omega \) as search direction. We refer to the new algorithm as the multiple-gradient descent algorithm (MGDA). The MGDA yields to a point on the Pareto set, at which a competitive optimization phase can possibly be launched on the basis of the local eigenstructure of the different Hessian matrices. This general formulation fosters several connected studies detailed in 6.2.3.

6.2.2. Virtual games for coupling global to local shape optimization

**Participant:** Régis Duvigneau.

In several engineering problems, the system to optimize is characterized by some parameters that define global shape properties, while remaining parameters define local shape modifications. Of course, these two sets of parameters do not play the same role and have not the same impact on the cost functional value. Therefore, we are studying how to construct an efficient optimization strategy that takes benefit of this global / local splitting of parameters.

A typical aerodynamic shape optimization problem has been studied, that consists of a lift-constrained drag minimization for a transonic wing, whose sections are defined by two B-Spline curves whereas global shape characteristics are defined by five parameters (span, roottip length ratio, angle of attack, twist angle, sweep angle). It has been found that the naive simultaneous optimization of all parameters failed, due to the multimodality of the problem. Alternatively, the use of a virtual game strategy, based on a splitting between the local and global parameters, yields a satisfactory result for a moderate cost [47].

6.2.3. Multiple-Gradient Descent Algorithm (MGDA)

**Participants:** Jean-Antoine Désidéri, Régis Duvigneau, Adrien Zerbinati.

6.2.3.1. Basic experiments and validation

In multi-objective optimization, the knowledge of the Pareto set provides valuable information on the reachable optimal performance. A number of evolutionary strategies (PAES, NSGA-II, etc), have been proposed in the literature and proved to be successful to identify the Pareto set. However, these derivative-free algorithms are very demanding in terms of computational time. Today, in many areas of computational sciences, codes are developed that include the calculation of the gradient, cautiously validated and calibrated.

In [50], MGDA has been tested over a number of classical multiobjective-optimization testcases, and found successful to converge to Pareto-optimal solutions in situations of either convex or concave Pareto sets. Additionally, MGDA and PAES [61] were found to have complementary merits, making a hybrid method promising.

6.2.3.2. Metamodel-supported CFD optimization by MGDA

Using MGDA in a multi objective optimization problem requires the evaluation of a large number of points with regard to criteria, and their gradients. In the particular case of a CFD problems, each point evaluation is very costly since it involves a flow computation, possibly the solution of an adjoint-equation. To alleviate this difficulty, we have proposed to construct metamodels of the functionals of interest (lift, drag, etc) and to calculate approximate gradients by local finite differences. These metamodels are updated throughout the convergence process to the evaluation of the new design points by the high-fidelity model, here the 3D compressible Euler equations.

This variant of MGDA has been tested successfully over a problem of external aerodynamic optimum-shape design of an aircraft wing consisting of reducing wave-drag, and augmenting lift. After only a few cycles of database updates, the Pareto front visibly forms, and this result is achieved at a very moderate computational cost. This variant is currently being tested and extended to an internal flow optimization problem related to an automobile air-conditioning system and governed by the Navier-Stokes equations. This more difficult problem has been proposed by Renault within the OMD2 ANR project.
6.2.3.3. MGDA in functional setting

One aspect of the theoretical result concerning the minimal-norm element $\omega$ is that, regardless the possibly-functional setting of the problem in case of a distributed system, the descent-direction $\omega$ is identified in the standard $n$-dimensional vector space $\mathbb{R}^n$ ($n$: the number of objective functions).

This observation has led to examine the application of MGDA in the functional setting of domain-decomposition methods (DDM) in which a functional criterion and a functional control can be defined at each interface independently permitting to formulate the DDM problem as a multi-objective optimization. On-going research in this area is related to the necessary preconditioning, or normalization procedure, of the gradients.

6.2.4. Flow control

Participants: Régis Duvigneau, Jérémie Labroquère.

Shape optimization methods are not efficient to improve the performance of fluid systems, when the flow is characterized by a strong unsteadiness related to a massive flow separation. This is typically the case for the flow around an automotive body or a wing in stall condition. To overcome this difficulty, flow control strategies are developed, that aim at manipulating vortex dynamics by introducing some active actuators, such as periodic blowing/suction jets. In this context, the choice of the control parameters (location, amplitude, frequency) is critical and not straightforward. Therefore, a numerical study is conducted to i) improve the understanding of controlled flows ii) develop a methodology to determine optimal control parameters by coupling the controlled flow simulation with optimization algorithms. Two research axes have been considered:

- the solution of the unsteady sensitivity equations derived from the state equations, to exhibit the dependency of the flow dynamics with respect to the control;
- the optimization of control parameters using a statistical metamodel-based strategy. First results show the efficiency of such an approach for laminar flow problems [31], [44].

6.2.5. Optimum shape design in aerodynamics by the adjoint method

Participants: Manuel Bompard, Sébastien Bourasseau, Jean-Antoine Désideri, Jacques Peter [Research Engineer, ONERA/DSNA].

At ONERA, compressible flow simulations governed by the Euler or Navier Stokes (RANS) equations are conducted with the software elsA [57] that admits both structured and unstructured-grid formulations. Local aerodynamic optimizations are made with a version that includes the calculation of the shape gradient via the solution of an adjoint equation. The discrete adjoint is calculated formally step-by-step to include the various derivative terms involved, and is being enhanced gradually to account for more complex models. In particular, for RANS computations, this gradient today includes the differentiation of the turbulence model.

6.2.5.1. Metamodels including derivative information.

In this context, to alleviate the cost of an optimum-shape design in aerodynamics, M. Bompard in his thesis [26], has examined how metamodels, firstly based on functional values only, could be used to determine shortcuts in the convergence process. Second, when the gradient w.r.t. the design parameters is known, the gradients of functionals of interest, that is, most commonly, aerodynamic coefficients, are calculated. Thus, these derivative informations can also used to construct more elaborate metamodels. Such constructions have also been studied systematically and used efficiently in global optimizations [37]; in particular co-Kriging and Support-Vector Regression, for which a technique to adjust automatically the free parameters has been proposed based on a simplification of the leave-one-out test.

6.2.5.2. Parameterization-free local optimization

When the derivatives of the functionals w.r.t. the volume geometry, $dJ/dX$, have been calculated, it is also possible to calculate the gradient w.r.t. surface coordinates, $dJ/DS$. Since the surface deformation steers the entire mesh movement, often through analytical dependencies, M. Bompard [26] has also examined how could $dJ/DS$ be used directly in a local aerodynamic optimization. However, it is well-known that the distribution of $dJ/DS$ is very irregular, and its usage in the optimization loop necessitates that adequate smoothing procedures be elaborated. Partial success was achieved in this area, still subject to research.
6.2.6. **Aero-structural optimization**  
**Participants:** Gérald Carrier [Research Engineer, ONERA/DAAP], Jean-Antoine Désideri, Imane Ghazlane.

In industry, aircraft wings are designed by accounting for several multidisciplinary couplings. Certainly of greatest importance is the coupling, or concurrency, between aerodynamic optimization and structural design. At ONERA, in the former thesis of M. Marcelet, the aerodynamic gradient has been extended to account for (the main terms of) static fluid-structure interaction, commonly referred to as the “aeroelastic gradient”.

In her thesis, I. Ghazlane has extended M. Marcelet’s work to take into account, in the aeroelastic gradient, the terms originating from the differentiation of the wing-structural model. In this development, the wing structure is treated as an equivalent Euler-Bernoulli beam. These formal extensions have been validated by an extensive experimentation. Additionally, special post-processing procedures are applied to evaluate accurately the various physical contributions to drag. As a result, the numerical tools necessary to conduct a very realistic aircraft wing optimization are now set up and are being exploited [38]. It is also envisaged to conduct a two-objective optimization (drag and mass reduction) via a Nash game using our optimization platform FAMOSA.

6.2.7. **Sonic boom reduction**  
**Participants:** Gérald Carrier [Research Engineer, ONERA/DAAP], Jean-Antoine Désideri, Andrea Minelli, Itham Salah El Din [Research Engineer, ONERA/DAAP].

When an aircraft flies at supersonic speed, it generates at ground level an N-shaped shock structure which can cause serious environmental damage (“sonic boom”). Thus a problem of interest in aerodynamic optimization is to design such an aircraft to reduce the intensity of the sonic boom while maintaining the aerodynamic performance (drag minimization under lift constraint). Andrea Minelli’s aimed at contributing to this two-discipline optimization problem. In the first part of his work, an inverse problem has been formulated and solved for “shaped sonic boom” and found in excellent agreement with the George-Seebass-Darden theory [60] for the calculation of the Whitham function corresponding to the lowest-boom (axisymmetric) shape. The method is currently being extended to account for more general geometries. Besides, aero-acoustic optimizations have been realized successfully by coupling the aerodynamic optimizer (based on Euler calculations by the elsA software) with the sonic-boom computation in a Nash game formulation. These experiments, conducted with our optimization platform FAMOSA, have demonstrated that starting from the shape optimized aerodynamically, one could retrieve smoothly a shape corresponding to nearly-optimal sonic-boom reduction.

6.2.8. **Helicopter rotor blade optimization in both situations of hovering and forward flight**  
**Participants:** Michel Costes [Research Engineer, ONERA/DAAP], Jean-Antoine Désideri, Arnaud Le Pape [Research Engineer, ONERA/DAAP], Enric Roca Leon.

E. Roca Leon has recently started at ONERA a CIFRE thesis supported by EUROCOPTER, Marignane. This thesis follows the doctoral thesis of A. Dumont in which the adjoint-equation approach was used to optimize a rotor blade in hovering flight. The goal of this new thesis is to solve a two-objective optimization problem in which the hovering-flight criterion is considered preponderant, but a new criterion that takes into account the forward-flight situation is also introduced, concurrently. The thesis work includes the set up of a hierarchy of models from low to high fidelity, in order to calibrate appropriate functional criteria. Secondly, our Nash game approach to competitive optimization will be implemented, using our optimization platform FAMOSA, and comparisons with the results by A. Dumont will be made.

6.2.9. **Optimum design in naval hydrodynamics**  
**Participants:** Régis Duvigneau, Louis Blanchard.

Naval hydrodynamics field has recently shown a growing interest for optimum design methods. The computational context is especially complex because it implies unsteady two-phase turbulent flows, with possibly very high Reynolds number (up to $10^9$). The use of automated design optimization methods for such problems requires new developments to take into account the large CPU time necessary for each simulation and the specificity of the geometries considered.
In collaboration with GALAAD Project-Team, some developments have been initiated on the geometrical modeling of hull shapes by parametric surfaces. The objective is to be able to modify existing hull shapes by controlling a small number of parameters, that are meaningful for naval architects. Two test cases are considered: the bow shape for trawler ships (see Fig. 4) and the whole hull shape for canoes, in collaboration with the Fédération française de Canoe-Kayak.

Figure 4. Initial shape (left) and deformed shape to generate a bow (right) for the trawler ship, to reach two line targets (in red).

6.3. Optimum design in structural mechanics

6.3.1. Shape Optimization in Multidisciplinary Non-Linear Mechanics

Participants: Aalae Benki, Jean-Antoine Désidéri, Abderrahmane Habbal.

In collaboration with the ArcelorMittal’s Center for Research in Automotive and Applications, we study the multidisciplinary shape and parameter design of highly non linear mechanical 2D and 3D structures. We have developed methods adapted to the approximation of Pareto Fronts such as Normal Boundary Intersection NBI and Normalized Normal Constraint Method NNCM. Due to the time consuming cost evaluation, the use of cheap to evaluate surrogate models is mandatory. We have studied the consistency of the approach NBI or NNCM plus surrogates, which turned out to be successful for a broad panel of standard mathematical benchmarks. The application of this approach for the case of beverage cans which undergo elastoplastic deformation under high pressure is ongoing.

6.3.2. Optimization of Addendum Surfaces in Stamping

Participants: Fatima Zahra Oujebbour, Jean-Antoine Désidéri, Abderrahmane Habbal.

Within the OASIS Consortium (ArcelorMittal, ErDF, INRIA, UTC, EURODECISION, ESILV, NECS, Delta-CAD, SCILAB-DIGITEO), Opale Project leads the Optimization task. Our aim is to develop decentralized decision-making algorithms dedicated to find efficient solutions (Pareto optimal) in a complex multidisciplinary framework (forming, stamping, welding non-linear processes, spring-back, vibration, in-function linear processes, crash and fatigue non linear and non differentiable processes) for several (between three and five) criteria. An important difficulty when trying to identify the Pareto Front, even when using adapted methods such the Normal Boundary Intersection, is that the criteria involved (thanks to the high nonlinearity in the mechanical models) exhibit many local optima. So one must use global optimization methods. We have studied the hybrid approach Simulated Annealing with Simultaneous Perturbation SASP for a suite of mathematical test-cases. To envisage the application of our method to the complex CPU time consuming stamping process, we lead an intermediate phase dedicated to the validation of the SASP method for the minimization of the spring-back that follows the stamping of a metal sheet, the design variable being the thickness distribution.
6.4. Application of shape and topology design to biology and medicine

6.4.1. Mathematical modeling of dorsal closure DC

Participants: Abderrahmane Habbal, Luis Almeida [University of Nice-Sophia Antipolis], Patrizia Bagnerini [Genova University], Fanny Serman [University of Nice-Sophia Antipolis], Stéphane Noselli [University of Nice-Sophia Antipolis], Glenn Edwards [Duke University].

A mathematical model for simulation of actin cable contraction, during wound closure for Drosophila embryo, which contains an extra term in addition to the curvature flow is developed. The basic mathematical model introduced and validated in [27] is extended in order to include the non-homogeneous wound healing or non-homogeneous dorsal closure [52].

6.5. Particular applications of simulation methods

6.5.1. Analysis of a two-level parameterization optimization for antenna design

Participants: Benoît Chaigne [Doctoral student, 2007-2010], Jean-Antoine Désideri.

Similar to the discretization of ordinary or partial differential equations, the numerical approximation of the solution of an optimization problem is possibly subject to numerical stiffness. In the framework of parametric shape optimization, hierarchical representations of the shape can be used for preconditioning, following the idea of Multigrid (MG) methods. By analogy with the Poisson equation, which is the typical example for linear MG methods, we have addressed a parametric shape inverse problem. The ideal cycle of a two-level algorithm can be defined and adapted to shape optimization problems that require appropriate transfer operators. With the help of a symbolic calculus software we have shown that the efficiency of an optimization MG-like strategy is ensured by a small dimension-independent convergence rate. Numerical examples are worked out and corroborate the theoretical results. Applications to antenna design have been realized. Finally, some connections with the direct and inverse Broyden-Fletcher-Goldfarb-Shanno preconditioning methods have been shown [29].

6.5.2. Mesh qualification

Participants: Jean-Antoine Désideri, Maxime Nguyen, Jacques Peter [Research Engineer, ONERA/DSNA].

M. Nguyen Dinh is conducting a CIFRE thesis at ONERA supported by AIRBUS France. The thesis topic is the qualification of CFD simulations by anisotropic mesh adaption. Methods for refining the 2D or 3D structured mesh by node movement have been examined closely. Secondly, it is investigated how could the local information on the functional gradient \( \|dJ/dX\| \) be exploited in a multi-block mesh context. This raises particular questions related to conservation at the interfaces.

6.5.3. Hybrid meshes

Participants: Sébastien Bourasseau, Jean-Antoine Désideri, Jacques Peter [Research Engineer, ONERA/DSNA], Pierre Trontin [Research Engineer, ONERA/DSNA].

S. Bourasseau has started a CIFRE thesis at ONERA supported by SNECMA. The thesis is on mesh adaption in the context of hybrid meshes, that is, made of both structured and unstructured regions. Again, the aim is to exploit at best the function gradient provided by the adjoint-equation approach. Preliminary experiments have been conducted on geometries of stator blade yielding the sensitivities to global shape parameters.

6.5.4. Nash game approach to image processing

Participants: Abderrahmane Habbal, Rajae Aboulaich [Mohamed V University of Rabat], Maher Moakher [University of Tunis], Moez Kallel [University of Tunis], Anis Theljani [University of Tunis].
We have started in 2011 to study the application of game modeling to image processing problems. We propose an original game theory approach to simultaneously restore and segment noisy images [56]. We define two players: one is restoration, with the image intensity as strategy, and the other is segmentation with contours as strategy. Cost functions are the classical relevant ones for restoration and segmentation, respectively. The two players play a static game with complete information, and we consider as solution to the game the so-called Nash Equilibrium. For the computation of this equilibrium we present an iterative method with relaxation. The results of numerical experiments performed on some real images show the relevance and efficiency of the proposed algorithm. Based on a similar idea, we formulated well known data completion (Cauchy) problems for Laplace equation as Nash games [55] and obtained results of existence, uniqueness and stability of a Nash equilibrium which turns out to be the Cauchy solution when the Cauchy data are compatible. With A. Theljani, we study the extension of the Nash data completion approach to nonlinear parabolic equations with application to image inpainting.

### 6.6. Isogeometric analysis and design

**Participants:** Louis Blanchard, Régis Duvigneau, Bernard Mourrain [Galaad Project-Team], Gang Xu [Galaad Project-Team].

Design optimization stands at the crossroad of different scientific fields (and related software): Computer-Aided Design (CAD), Computational Fluid Dynamics (CFD) or Computational Structural Dynamics (CSM), parametric optimization. However, these different fields are usually not based on the same geometrical representations. CAD software relies on Splines or NURBS representations, CFD and CSM software uses grid-based geometric descriptions (structured or unstructured), optimization algorithms handle specific shape parameters. Therefore, in conventional approaches, several information transfers occur during the design phase, yielding approximations that can significantly deteriorate the overall efficiency of the design optimization procedure. Moreover, software coupling is often cumbersome in this context.

The isogeometric approach proposes to definitely overcome this difficulty by using CAD standards as a unique representation for all disciplines. The isogeometric analysis consists in developing methods that use NURBS representations for all design tasks:

- the geometry is defined by NURBS surfaces;
- the computation domain is defined by NURBS volumes instead of meshes;
- the solution fields are obtained by using a finite-element approach that uses NURBS basis functions
- the optimizer controls directly NURBS control points.

Using such a unique data structure allows to compute the solution on the exact geometry (not a discretized geometry), obtain a more accurate solution (high-order approximation), reduce spurious numerical sources of noise that deteriorate convergence, avoid data transfers between the software. Moreover, NURBS representations are naturally hierarchical and allows to define multi-level algorithms for solvers as well as optimizers. In this context, some research axes have been developed in collaboration with GALAAD Project-Team:

- Methods for adaptive parameterization including a posteriori error estimate for elliptic problems [36], [35], [42];
- Numerical schemes based on Spline functions for 2D inviscid compressible flow simulations;
- Optimization methods for structural elasticity, based on shape-gradient concept, and fluid-structure interactions [48] (in collaboration with Technical University of Munich).

### 6.7. Resilient workflows for distributed multidiscipline optimization

**Participants:** Toan Nguyen, Laurentiu Trifan.
A distributed platform based on the YAWL workflow management system has been designed and implemented to deploy HPC applications on the Grid5000 network infrastructure. The goal is to provide a generic environment for the design of complex applications that require HPC resources for large-scale fault-tolerant applications, see Fig. 2 and [39].

The platform provides application-level fault-tolerance, i.e., resilience, in order to restart the workflow execution whenever abnormal behavior or system-level errors occur. This allows a variety of errors to be taken into account, ranging from execution time-outs to out-of-bounds parameter values to be managed, with the help of user intervention when necessary [40].

The error management procedure uses exception handlers in YAWL to trigger the appropriate corrective actions, which are defined by rules invoking the adequate compensating workflows. Once defined, this can be made transparent to the users [41].

An original scheme based on asymmetric checkpoints has been designed in order to reduce overhead in both checkpointing and application restarts. It minimizes the number of required checkpoints created based on default rules and user-specific needs.

The platform is currently developed in Java on Linux workstations and should be portable on Windows and MacOS, although this has not been tested yet.

Examples are deployed on the Grid5000 national network infrastructure using the OMD2 test-cases (e.g., vehicle air-conditioner pipe optimization). The goal is here to provide a demonstrator platform that deploys large-scale optimization applications involving several (typically over five) HPC clusters distributed on the Grid5000 network. The coarse-grain definition of the application is defined by a workflow that monitors the distributed execution of the parallel component codes on the various clusters, providing resilience capabilities in case of system and application errors, see Fig. 5.

Figure 5. Application definition using YAWL.
POEMS Project-Team

6. New Results

6.1. Numerical methods for time domain wave propagation

6.1.1. High Order Theta Scheme for the linear wave equation. 

Participants: Juliette Chabassier, Sébastien Impériale.

We have pursued our work about a new class of high order implicit three time step schemes for semi-discretized wave equations of the form

\[
\frac{d^2}{dt^2} u_h + K_h u_h = 0, \quad u_h(0) = u_{0,h}, \quad \frac{du_h}{dt}(0) = u_{1,h},
\]

(6)

where \(K_h\) is a symmetric positive definite matrix. For such problems, explicit three time step schemes generally show good performances but present two major drawbacks, in some situations, that have not yet been completely overcome:

- If the mesh presents elements of quite different sizes, the time step must be adapted to the smallest one because of the CFL condition.
- If the mass matrix is non diagonal or non block-diagonal, its inversion (at least one time per iteration) can lead to a dramatic over cost of the schemes, which is obviously not the case with implicit schemes for which a matrix has to be inverted in any case.

A natural way to avoid this restriction is to use local time stepping techniques which divides into two categories. First, the locally implicit technique, which is optimal in term of CFL restriction but “only” second order accurate in time, and requires the inversion of interface matrices. Second, the fully explicit local time stepping, as developed that enables to achieve high order time stepping but without (up to now) a full control over the CFL condition.

This is why we construct generalized implicit \(\theta\)-scheme using the modified equation approach to obtain 4th order approximation. These schemes introduce 2 degree of freedom \(\theta\) and \(\varphi\) and can be written under a general form as

\[
\frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2} + K_h \left( \theta u_h^{n+1} + (1 - 2\theta) u_h^n + \theta u_h^{n-1} \right) + \left( 1 - \frac{\theta}{12} \right) K^2_h \left( \varphi u_h^{n+1} + (1 - 2\varphi) u_h^n + \varphi u_h^{n-1} \right) = 0.
\]

The parameters \(\theta\) and \(\varphi\) are chosen as functions of the time step through the resolution of an optimization problem: we consider that the time step is given and we try to obtain a scheme that minimizes the consistency errors still being stable. The limit problem when the time step tends is infinite gives an optimal unconditionally stable fourth order scheme. This work has been submitted for publication.

6.1.2. Discontinuous Galerkin Methods for wave equations

Participants: Patrick Joly, Antoine Tonnoir.
This has been the subject of the internship of A. Tonnoir and can be seen as a contribution to the mathematical analysis of coupled BEM / DG methods for time domain diffraction problems (see section 6.1.4). We did not pretend to make a real contribution to the field of the analysis of DG methods, but wanted to understand in sufficient detail the quite surprising observation that the non-centered (in space) schemes, provided by the use of non centered fluxes in the DG approach, leads to a better accuracy than the centered schemes issued for central fluxes! This has been done (for the simple 1D wave equation) from two points of view: the direct energy method and the Fourier analysis (or dispersion analysis) on regular grids.

6.1.3. Analysis of time domain boundary integral equations

Participants: Patrick Joly, Nicolas Le Guillarme.

This has been developed again as a part of the mathematical analysis of coupled BEM / DG methods for time domain diffraction problems (see section 6.1.4). With J. Rodríguez, we have revisited the analysis of the coercivity / continuity properties of space-time boundary integral associated to wave diffraction problem. Unlike the more traditional approach based on Laplace transform in time (cf the initial work by Bamberger-Ha Duong) we treat this question directly in the time domain by exploiting in a simple way the connection between integral equations and boundary value problems. This can be done in a very general setting but also particularized to the 1D case for which we got sharp estimates (internship of N. Le Guillarme). Furthermore, we can reinterpret the classical discretization by finite elements for retarded potentials as non conforming finite element methods. This allows to investigate in a new way the error analysis of time domain boundary element methods, which will be the subject of a future work.

6.1.4. Coupling Retarded Potentials and Discontinuous Galerkin Methods for time dependent wave propagation problems

Participant: Patrick Joly.

This topic is developed in collaboration with J. Rodríguez (Santiago de Compostela) in the framework of the contract ADNUMO with AIRBUS. Let us recall that our objective was to use time-domain integral equations (developed in particular at IMACS and EADS) as a tool for constructing transparent boundary conditions for wave problems in unbounded media. Our previous contribution of this topic concern the construction of an energy preserving method for the coupling space-time Galerkin approximation of the integral equations with discontinuous Galerkin finite elements and leap-frog time discretization for the numerical approximation of the equations inside the computational domain. For stability reasons, we had to use central fluxes. The drawback of central fluxes is that they do not lead to an optimal accuracy (see also the paragraph on the analysis of DG methods) which is traduced in practice by high frequency spurious oscillations.

Our objective this year was to look for another discretization approach that would overcome these problems. The approach we propose consists in playing on the time discretization procedure. For this we decompose the stiffness bilinear form as the sum of a conservative term corresponding to the use of central fluxes and a ("small") dissipative term due to off-centering and involving jumps across interfaces of the discrete solution. We propose a scheme which treats the conservative part of the equation in a centered way (leap-frog type) and the dissipative term in a non centered way (backward Euler type). Doing so, the overall accuracy in time of the scheme is preserved (with respect to the case of central fluxes) and the stability is maintained at the price of a (slightly) more constraining CFL conditions, which does not seem that much penalizing for the application. The stability is analyzed through the decay of an discrete energy. As a consequence, we can adapt the discretization of the coupling terms between integral and interior equations in order to preserve the stability of the fully coupled scheme under the same CFL condition. Numerical validations are in progress.


Participants: Eliane Bécache, Aliénor Burel, Sébastien Impériale, Patrick Joly.
This topic is the subject of the first part of the PhD thesis of A. Burel and has been developed for a part in collaboration with Marc Duruflé. Decomposing the displacement field into potentials is a well-known tool in elastodynamics, and it expresses the decoupling of the pressure wave and the shear wave inside a homogeneous isotropic media. Although this tool is classically used when searching for analytic solutions, it does not seem to have been exploited for numerical computation using finite elements for instance. However, this is a priori attracting since, contrary to a displacement field approach for instance, it allows to decouple the approximation of P and S waves and to adapt the discretization process (mesh size, order of elements) to the dynamics of each type of wave, which is a priori particularly interesting when S-waves propagate much more slowly than P-waves (soft materials such as rubber). The main difficulty is to cope with the coupling of the different types of waves (the so-called conversion of modes) which occurs, due to wave reflections and transmissions, at interfaces between homogeneous media or at physical boundaries. The simplest situation where this phenomenon appears is the propagation of elastic waves in a homogeneous domain with clamped boundary. The main difficulty is to guarantee the stability of the treatment of the boundary condition. For this, we have proposed a variational formulation in which the stiffness bilinear form appears as a sum of a decoupled volumic bilinear form and a coupled surfacic bilinear form. This formulation is is compatible to an energy conservation result. After space discretization with finite elements spaces well adapted to each type of waves (using different meshes and/or polynomial degrees), we propose a discrete energy preserving numerical scheme, based on an explicit discretization of the volume terms and an implicit discretization of the surface terms. The resulting scheme is mainly explicit (only a sparse boundary linear system has to be solved at each time step) and stable under a CFL stability condition that is not affected by the presence of the boundary. An approach based on a modification of this scheme has been proposed for treating the free surface condition. This approach appears to give satisfactory results in the frequency domain (on the basis of numerical simulations) but fails in time domain due to (apparently unconditional) instabilities.

6.1.6. Mathematical and numerical modeling of piezoelectric sensors.

Participants: Sébastien Impériale, Patrick Joly.

This research, which constitutes the subject of the PhD thesis of S. Imperiale (which will be defended in January 2012) is developed in the framework of a collaboration with CEA-LIST about the numerical modeling of non-destructive testing experiments using ultrasonic waves.

More precisely, we have concerned during the past three years by mathematical and numerical questions related to the simulation of non destructive testing experiments using piezoelectric devices. In particular, we have investigated the modeling of piezoelectric sensors that are used to generate and record ultrasonic waves in a solid material: such waves are typically used to investigate in a non invasive way the possible presence of defects in manufactured items. Such an issue has already been tackled in the engineering literature but not, to our knowledge, by way of rigorous applied mathematics. As we are arriving at the conclusion of this work, let us summarize our main contributions during these three years (these have been published in M2AN)

- The full equations of piezoelectricity couple Maxwell’s equations with linear elastodynamics equations which corresponds to a coupled hyperbolic system. This system presents quite different time scales due to the very large ratio between the speed of light and the sound speed, which makes it impossible to treat by a direct numerical approach. To overcome this problem, we have given a rigorous justification, via asymptotic analysis, of the so-called quasi-static approximation model in which the electric unknowns are reduced to a scalar electric potential: the reduced model appears as a coupled elliptic-hyperbolic system.

\[
\begin{align*}
\rho \partial_t u - \text{div} \ C \varepsilon(u) - \text{div} \ e \nabla \varphi &= 0 & \text{in } & \Omega_S \ (\text{the solid domain}), \\
\nabla \cdot \varepsilon \nabla \varphi - \nabla \cdot e^T \varepsilon(u) &= 0 & \text{in } & \mathbb{R}^3.
\end{align*}
\]

- We have next justified the reduction of the computation of this electric potential to the piezoelectric parts of the computational domain, considering the large contrast of permittivity between piezoelectric materials and surrounding polymers. This relies again on a limit process: the ratio between permittivities is the small parameter.
A particular attention has to be devoted to the modeling of the electric supply process: more precisely the nonlocal (in space) boundary conditions on the electrodes (used to model the emission and reception regimes), as well as the modeling of the coaxial cable connecting the sensor to the electric generator (see also section 6.4.1).

Concerning the numerical approximation, an energy preserving finite element / finite differences numerical scheme is developed. A specific procedure is used for treating the nonlocal boundary conditions on the electrodes. Unbounded media have been treated via PML techniques that are dealt with using an original mortar element technique (see section 6.3.1 for more details). Various numerical results in academic or more realistic situations have been obtained. For instance, we have been able to model how one can exploit multi-component devices (see figure 1) to produce well focused waves (see figure 2).

A computational code issued from our research, named Ondomatic (12 000 lines in C++), has been implemented. This code uses the finite element library FEMME (15 000 lines in C++) that essentially relies upon the use of spectral finite element on hexahedral meshes.

Figure 1.

Figure 2.
6.1.7. Maxwell’s equations in Lorentz materials  
**Participants:** Patrick Ciarlet, Patrick Joly, Valentin Vinoles.

This is the time-domain counterpart of the research done at Poems about metamaterials (see also the section 6.2.7) and has been the subject of the internship of V. Vinoles. Lorentz materials are particular non dissipative dispersive materials which behave as metamaterials (with negative index of refraction) in some range of frequencies. Their constitutive laws (links between electric and magnetic fields and the related inductions) are described in terms of ordinary differential equations (harmonic oscillators). We have studied Maxwell’s equations in such materials and in particular proposed an energy preserving space-time discretization scheme based on an extension of classical methods (mixed finite elements in space, leap-frog type schemes in time). This has been implemented in a code including PMLs for the treatment of unbounded domains. Various numerical experiments have been performed. They illustrate the spectacular dispersive effects of Lorentz materials and allow us to recover the expected focalisation phenomena by an interface between a metamaterial and a standard dielectric one. This code will be an essential tool for the further investigation of more theoretical (and challenging) issues such as the limiting amplitude principle in this context.

6.1.8. Evolution problems in perturbed infinite periodic media  
**Participants:** Julien Coatléven, Sonia Fliss.

First, as a part of the PhD of J. Coatléven, based on the former method to solve linear evolution problems in locally perturbed infinite periodic strips through the construction of transparent boundary conditions, a method has been developed to solve its natural geometric extension, i.e. the case of a locally perturbed line defect in an infinite periodic media. The method is again based on a semi-discretization in time of the problem on the whole infinite periodic media, and a generalization of the Lippmann-Schwinger equation approach we have developed for the treatment of this kind of geometries but for harmonic wave problems. At each time step, the solution is written as the sum of the solution of a time step in the unperturbed line defect and a contribution of the perturbation due not only to the current time step but also to all time steps involved in the numerical scheme. This intrication of time steps requires a careful treatment of the Lippmann-Schwinger equation, and in particular of the source term. Using the Floquet-Bloch transform, the computation of all the quantities involved can be reduced to the resolution of a few time steps of linear evolution problems in locally perturbed infinite periodic strips, where we can use the former method. As in the harmonic case, the discretization of the inverse Floquet-Bloch transform is done using appropriate quadrature rules, whereas the space discretization requires classical finite elements. The theoretical basis as well as the numerical analysis of this method are well understood, and the method has been successfully tested numerically. For instance, it can be shown, and checked numerically, that in the particular case of the wave equation, if one uses enough quadrature points (depending on the length of the time interval), then this quadrature creates no error of approximation.

For parabolic problems set in locally perturbed periodic media, we have developed another approach to determine the time-domain DtN operator. The principle is to apply the Laplace Transform in time to the equation and use the construction of the DtN operator for stationary equations. The main difficulty is the computation of the inverse of the Laplace Transform, more precisely to understand how to deal with the unbounded interval of integration and the choice of the discretization of the laplace variable. To deal with the first difficulty for waveguide problem, we have studied the asymptotic behavior of the DtN operator in the laplace domain when the laplace variable tends to $p_0 \pm \infty$. To deal with the second difficulty, we have used the Z-Transformation and its properties. The numerical study is still in progress. This work enters in the framework of the ANR PRoject MicroWave (Sonia Fliss is an external collaborator), in collaboration with Karim Ramdani (Institut Elie Cartan de Nancy, UMR CNRS 7502), Christophe Besse and Ingrid Violet (Laboratoire Paul Painlevé, UMR CNRS 8524).

6.1.9. Modeling and numerical simulation of a piano.  
**Participants:** Juliette Chabassier, Patrick Joly.
This work is developed in collaboration with Antoine Chaigne (UME, ENSTA). The purpose of this work is the time domain modeling and numerical simulation of a piano. We aim at explaining the vibratory and acoustical behavior of the piano, by taking into account the main elements that contribute to sound production. The soundboard is modeled as a bidimensional thick, orthotropic, heterogeneous, frequency dependent damped plate, using Reissner Mindlin equations. The vibroacoustics equations allow the soundboard to radiate into the surrounding air, in which we wish to compute the complete acoustical field around the perfectly rigid rim. The soundboard is also coupled to the strings at the bridge, where they form a slight angle from horizontal. Each string is modeled by a one dimensional damped system of equations, taking into account not only the transversal waves excited by the hammer, but also the stiffness thanks to shear waves, as well as the longitudinal waves arising from geometric nonlinearities. The hammer is given an initial velocity that projects it towards a choir of strings, before being repelled. The interacting force is a nonlinear function of the hammer compression.

The final piano model that is discretized is a coupled system of partial differential equations, each of them exhibiting specific difficulties (nonlinear nature of the string system of equations, frequency dependent damping of the soundboard, great number of unknowns required for the acoustic propagation), in addition to couplings’ inherent difficulties. On the one hand, numerical stability of the discrete scheme can be compromised by nonlinear and coupling terms. A very efficient way to guarantee this stability is to construct a numerical scheme which ensures the conservation (or dissipation) of a discrete equivalent of the continuous energy, across time steps. A major contribution of this work has been to develop energy preserving schemes for a class of nonlinear systems of equations, in which enters the string model. On the other hand, numerical efficiency and computation time reduction require that the unknowns of each problem’s part, for which time discretization is specific, hence different, be updated separately. To achieve this artificial decoupling, adapted Schur complements are performed after Lagrange multipliers are introduced.

The potential of this time domain piano modeling is emphasized by realistic numerical simulations. Beyond greatly replicating the measurements, the program allows us to investigate the influence of physical phenomena (string stiffness or nonlinearity), geometry or materials on the general vibratory behavior of the piano, sound included. Spectral enrichment, << phantom partials >> and nonlinear precursors are clearly revealed when large playing amplitudes are involved, highlighting how this approach can help better understand how a piano works.

The main contributions of this year have been the following:

- to write the acoustic propagation as a first order system of equation, involving the physical sound pressure (as opposed to before, when its primitive had to be considered) and the acoustical velocity: this allowed us to artificially bound the computation domain with Perfectly Matched Layers.
- to decrease the numerical computation times via a massive parallelization of the code: parallel modal search for the soundboard, parallel dense matrix-vector product for the vibroacoustic equations in the modal basis, parallel resolution of the 3D acoustic propagation, multi-threaded computation of the strings’ problem.
- to perform realistic computations, which have provided physically relevant numerical simulations (see Figure 3).

Several measurements have also been conducted on a grand piano in order to provide realistic values of parameters and calibrated data to compare simulation with.

### 6.2. Time-harmonic diffraction problems

#### 6.2.1. Numerical computation of variational integral equation methods

**Participants:** Marc Lenoir, Nicolas Salles.
The dramatic increase of the efficiency of the variational integral equation methods for the solution of scattering problems must not hide the difficulties remaining for an accurate numerical computation of some influence coefficients, especially when the panels are close and almost parallel.

The formulas have been extended to double layer potentials and, for self influence coefficients, to affine basis functions. Their efficiency for the solution of Maxwell equations has been proved in the framework of a collaboration with CERFACS. The redaction of a paper devoted to the case of parallel panels has been completed and submitted to SIAM J. Sci. Comp. Some preliminary work on the numerical integration of the regular part of the integrand has been undertaken in the context of an internship.

6.2.2. Fast Multipole Method for Viscoelastodynamics

**Participants:** Marc Bonnet, Stéphanie Chaillat.

This work is done in collaboration with Eva Grasso (LMS, Ecole Polytechnique) and Jean-François Semblat (IFSTTAR). We have extended the single- and multi-domain time-harmonic elastodynamic multi-level fast multipole BEM (Boundary Element Method) formulations to the case of weakly dissipative viscoelastic media [21]. The underlying boundary integral equation and fast multipole formulations are formally identical to that of elastodynamics, except that the wavenumbers are complex-valued due to attenuation. Attention was focused on evaluating the multipole decomposition of the visco-elastodynamic fundamental solution, involving complex-valued wavenumbers. As a result, a damping-dependent modification of the selection rule for the multipole truncation parameter was proposed and assessed on 3D single-region and multi-region visco-elastodynamic examples involving up to about $3 \times 10^5$ boundary nodal unknowns.

6.2.3. Formulation and Fast Evaluation of the Multipole Expansions of the Elastic Half-Space Fundamental Solutions

**Participants:** Marc Bonnet, Stéphanie Chaillat.
This ongoing work is concerned with a formulation and computation algorithm for the elastodynamic Green’s tensor for the traction-free half-space allowing its use within a Fast Multipole Boundary Element Method (FM-BEM). Due to the implicit satisfaction of the traction-free boundary condition achieved by the Green’s tensor, discretization of (parts of) the free surface is no longer required. Unlike the full-space fundamental solution, the elastodynamic half-space Green’s tensor cannot be expressed in terms of usual kernels such as \( e^{ikr}/r \) or \( 1/r \). Its multipole expansion thus cannot be deduced from known expansions, and is formulated in this work using a spatial two-dimensional Fourier transform approach. The latter achieves the separation of variables which is required by the FMM. A key numerical issue, upon which current work is focused, is concerned with the definition of an efficient numerical quadrature for the evaluation of the inverse Fourier transform, whose integrand is both singular and oscillatory, as classical Gaussian quadratures would perform poorly, fail or require unacceptably large number of quadrature points.

### 6.2.4. Multiple scattering by small scatterers

**Participants:** Maxence Cassier, Christophe Hazard.

We consider the scattering of an acoustic time-harmonic wave by an arbitrary number of sound-soft obstacles located in a homogeneous medium. When the size of the obstacles is small compared with the wavelength, the numerical simulation of such a problem by classical methods (e.g., integral equation techniques or methods based on a Dirichlet-to Neumann map) can become highly time-consuming, particularly when the number of scatterers is large. In this case, the use of an asymptotic model may reduce considerably the numerical cost. Such a model was introduced by Foldy and Lax in the middle of the last century to study multiple isotropic scattering in a medium which contains randomly distributed small scatterers. Their asymptotic model is based on the fact that the scattered wave can be approximated by a wave emitted by point sources placed at the centers of the scatterers; the amplitudes of the sources are calculated by solving a linear system which represents the interactions between the scatterers. Nowadays, the Foldy–Lax model is still used in numerous physical and numerical applications to approximate the scattered wave in a deterministic media. But to the best of our knowledge, there was no mathematical justification of this asymptotic model. We have proposed such a justification which provides local error estimates for the two-dimensional problem in the case of circular obstacles. An article on this subject has been recently submitted to Wave Motion.

### 6.2.5. Harmonic wave propagation in locally perturbed infinite periodic media

**Participants:** Julien Coatléven, Sonia Fliss, Patrick Joly.

A part of the PhD of J. Coatléven consists in developing a method for solving harmonic wave problems with locally perturbed line defects in periodic media. For the treatment of these unbounded defects, which are structured apart from a local perturbation, a new approach has been developed, based on a perturbation principle. The solution is written as a the sum of a solution corresponding to the unperturbed line defect and a contribution of the local perturbation. This decomposition leads to a generalization of the so-called Lippmann-Schwinger equation, whose coefficients are computed through their Floquet-Bloch transform, which leads to solve wave-guide problems, these last problems being solved using the transparent boundary condition method developed during S. Fliss’s PhD. The discretization of the inverse Floquet-Bloch transform is done using appropriate quadrature rules, whereas the space discretization requires classical finite elements. The theoretical basis as well as the numerical analysis of this method are well understood, and the method has been successfully tested numerically. In particular, the theoretical convergence estimates have been checked in practice, and the method has a satisfying behavior in limit cases not fully covered by the theory, such as the non-absorbing case.

Concerning the non absorbing case, the question of the limiting absorption principle has been treated for locally perturbed periodic media with particular assumptions. In this case, we are studying the behavior of the solution at large distance of the local perturbation.

### 6.2.6. Time harmonic aeroacoustics

**Participants:** Anne-Sophie Bonnet-Ben Dhia, Jean-François Mercier.
We are still working on the numerical simulation of the acoustic scattering and radiation in presence of a mean flow. This is the object of the ANR project AEROSON, in collaboration with Florence Millot and Sébastien Pernet at CERFACS, Nolwenn Balin at EADS and Vincent Pagneux at the Laboratoire d’Acoustique de l’Université du Maine. The main recent improvements concern: the consideration of ducts with treated boundaries and the development of an alternative model to Galbrun’s equation.

**Treated boundaries**

Our aim is to extend the time harmonic equation of Galbrun to take into account acoustically treated boundaries. Such boundaries are generally described by the Myers boundary condition. Since this condition is naturally expressed in terms of Galbrun’s unknown, the displacement $u$, Galbrun’s equation easily extends to treated boundaries. However, we face a difficulty: the original equation of Galbrun leads to a non coercive problem. For rigid boundaries, considering an augmented variational formulation leads to well-posedness. But this approach does not work anymore for treated boundaries.

We have improved our understanding of this difficulty. We are now convinced that the Augmented Galbrun’s equation combined with Myers condition leads to an ill-posed problem. More precisely, source terms for which the solution of Galbrun’s equation does not belong to standard functional spaces exist. Such source terms are very particular: located on the treated boundary and singular. This is why during the numerical validations performed at Cerfacs, for "standard" source terms ($L^2$ functions compactly supported in the fluid) we did not get any problem.

To get a well-posed problem, Myers boundary condition, which just requires the normal displacement to belong to $L^2$, must be regularized. It can be achieved by requiring the tangential derivative of $u \cdot n$ to belong to $L^2$. We have also understood that less regularity is sufficient to get a well-posed problem, as it is the case if the fluid is in contact of an elastic medium (such interface cannot be described by a Myers condition since it is necessarily non-local). In particular, in the case of a uniform flow well-posedness is proved for a sufficiently slow flow.

**Alternative to Galbrun’s model**

We have kept on considering the model of Goldstein’s equations, alternative to Galbrun’s equation. The Goldstein’s equations couple two unknowns: the velocity potential $\varphi$ and a vectorial unknown $\xi$. $\varphi$ satisfies a modified Helmholtz’s equation with variable coefficients linked to the flow, in which $\xi$ is added as a source term. $\xi$ satisfies a transport equation coupled to the velocity potential. This new model facilitates the treatment of 3D problems since Galbrun’s equation requires to introduce many unknowns. Moreover, the vectorial unknown in Goldstein’s formulation vanishes in the areas where the flow is potential which is interesting since realistic flows are mainly potential, the non-potential areas being located near the boundaries or behind obstacles.

As it is the case to calculate the vorticity $\psi = \text{curl} u$, used to regularize Galbrun’s equation, a Discontinuous Galerkin (DG) discretization is used to determine $\xi$ and numerical simulations have been performed at Cerfacs. We have also developed an alternative method allowing to solve Goldstein’s equations with simple Lagrange Finite Element and this was the object of Jean-Emmanuel Lauzet’s internship. We have developed a method combining the Streamline Upwind Petrov Galerkin (SUPG) scheme to discretize the Goldstein’s model with the introduction of PML to bound the calculation domain. To test the efficiency of the method, a non-potential flow has been determined analytically. It consists of a lid-cavity flow connected to a uniform flow in a duct. The viscous cavity flow is solution of the Stokes’ equation and is determined in a rectangular domain by a modal method.

### 6.2.7. Modeling of meta-materials in electromagnetism

**Participants:** Anne-Sophie Bonnet-Ben Dhia, Patrick Ciarlet, Lucas Chesnel.
A collaboration with Eric Chung (Chinese Univ. of Hong Kong) and Xavier Claeys (ISAE).
Meta-materials can be seen as particular media whose dielectric and/or magnetic constant are negative, at least for a certain range of frequency. This type of behavior can be obtained, for instance, with particular periodic structures. Of special interest is the transmission of an electromagnetic wave between two media with opposite sign dielectric and/or magnetic constants. As a matter of fact, applied mathematicians have to address challenging issues, both from the theoretical and the discretization points of view.

The first topic we considered a few years ago was: when is the (simplified) scalar model well-posed in the classical $H^1$ framework? It turned out this issue could be solved with the help of the so-called $T$-coercivity framework. While numerically, we proved that the (simplified) scalar model could be solved efficiently by the most "naive" discretization, still using $T$-coercivity.

Recently, we have been able to provide sharp conditions for the $T$-coercivity to hold in general 2D and 3D geometries (with L. Chesnel), which involve explicit estimates in simplified geometries together with localization arguments. We then analyzed the discretization of the scalar problem with a classical, $H^1$ conforming, finite element method, and proved the convergence under the same sharp conditions (with L. Chesnel). We also showed that the problem can be solved with the help of a Discontinuous Galerkin discretization, which allows one to approximate both the field and its gradient (with E. Chung).

Last (with L. Chesnel and X. Claeys), we investigated the case of a 2D corner which can be ill-posed (in the classical $H^1$ framework). Using the Mellin transform, we showed that a radiation condition at the corner has to be imposed to restore well-posedness. Indeed there exists a wave which takes an infinite time to reach the corner: this "black hole" phenomenon is observed in other situations (elastic wedges for example).

As a second topic, we studied the transmission problem in a purely 3D electromagnetic setting from a theoretical point of view: to achieve well-posedness of this problem, we had to proceed in several steps, proving in particular that the space of electric fields is compactly embedded in $L^2$. For that, we had to assume that the interface is "sufficiently smooth", excluding in particular corners. With L. Chesnel, we have been able to remove this assumption, so that we can solve the problem around an interface with corners. It turns out the $T$-coercivity framework can be applied once more, under the same assumptions as for the scalar model. In the process, we recover more compact embedding results.

### 6.2.8. Numerical MicroLocal Analysis

**Participants:** Jean-David Benamou, Francis Collino, Simon Marmorat.

Numerical microlocal analysis of harmonic wavefields is based on a family of linear filters using Bessel functions and applied to wave data collected on a circle of fixed radius $r_0$ around the observation point $x_0$ where we want to estimate the Geometric Optics/ High Frequency components. The data can easily be reconstructed from more conventional line array or grid geometry. The output is an angular function presenting picks of amplitudes in the direction angles of rays.

The original NMLA algorithm relied on a local plane wave assumption for the data. For arbitrary waves, it meant linearization errors and accuracy limitations. Also, only the directions of the (multiple) rays are recovered but the traveltimes and amplitudes are not reliably computed. We recently introduced a new "impedant" observable which allows to prove a stability theorem. Numerical results confirm that the new NMLA filter is robust to random and correlated noise.

Using asymptotic expansion on NMLA filtered point sources data, we designed a correction method for the angle which also estimates the wavefront curvature. It can be used to correct the linearization errors mentioned above and provides a second order correction in the Taylor approximation of the traveltimes.

The parameters of the method (size of observation circle, discretization) are automatically optimized and a posteriori quantitative error on angles and curvature are available. Numerical studies validate the stability result and confirm the superior accuracy of the curvature corrected NMLA version over image processing methods.
When some bandwidth is available we can also compute the travel time. The amplitude remains polluted by phase errors. Its determination is still open.

6.3. Absorbing boundary conditions and absorbing layers

6.3.1. Perfectly matched transmission problem with absorbing layers: application to anisotropic acoustics

Participant: Sébastien Impériale.

This work has been carried out in collaboration with Edouard Demaldent from CEA-LIST. We have worked on an original approach to design perfectly matched layers (PML) for transient wave equations. This approach is based, first, on the introduction of a modified wave equation and, second, on the formulation of general “perfectly matched” transmission conditions for this equation. The stability of the transmission problem is discussed by way of the adaptation of a high frequency stability (necessary) condition, and we apply our approach to define PML suited for the anisotropic wave equation. A variational formulation of the problem is then developed. It includes a Lagrange multiplier at the interface between the physical and the absorbing domains to deal with the “perfectly matched” transmission conditions. We have carried out numerical results in 2D and 3D that first show the validity of our approach in terms of stability and accuracy and the efficiency when using constant damping coefficients combined with high order elements. This work has been submitted for publication.

6.3.2. On the stability in PML corner domains

Participant: Eliane Bécache.

In collaboration with Andres Prieto, from the University of Santiago de Compostella. We have finalized our work on the stability of the discretization of PMLs in the corners and submitted a paper (see preprint [24]).

6.3.3. High-order Absorbing Boundary Conditions for anisotropic elastodynamics

Participants: Daniel Baffet, Eliane Bécache.

This work is done in collaboration with Daniel Baffet, PhD student of Dan Givoli, at the Technion University in Haifa (Israel) and has started during a visit of Daniel at Poems.

The aim is to design new efficient and stable absorbing boundary conditions for anisotropic materials. It is known that the anisotropy introduces a specific difficulty, for ABC as well as for PMLs, in particular for models which involve inverse modes, i.e., waves for which the phase velocity and the group velocity propagate in opposite directions (with respect to the boundary). This has given rise to specific treatment for scalar models but for anisotropic elastodynamics, there are some materials for which no satisfactory solution exist. For these materials however, we can design a low order boundary condition, which is proved to be stable via an energy estimate. We have started to investigate several ways to design higher-order boundary conditions. The main difficulty is to show whether these boundary conditions are stable or not.

6.3.4. Dirichlet to Neumann map with overlap for waveguides

Participants: Anne-Sophie Bonnet-Ben Dhia, Sonia Fliss, Geoffrey Martinache, Antoine Tonnoir.

For scattering problems in acoustic waveguides, a usual approach consists in restricting the computation to a bounded domain containing the sources and the perturbations, using transparent boundary conditions on the artificial boundaries. These conditions are written by using the so-called Dirichlet-to-Neumann maps which can be expressed thanks to a modal decomposition.
An iterative solution of the related linear system can be seen as a domain decomposition formulation without overlapping, where one domain is the bounded region and the other one is infinite. This iterative method does not converge necessarily. A classical idea is to consider a domain decomposition method with overlapping. In this work, we find the equivalent of this method in terms of a new Dirichlet-to-Neumann operator which links the trace of the solution on a section of the waveguide to the normal trace on a different one. This operator can also be expressed analytically via a modal decomposition. Its main advantage is that, because of the overlapping, it becomes compact and this is exactly why we think an iterative resolution has more chance to converge. Other advantages will appear with the elasticity application. Indeed, in the formulation of the transparent boundary condition without overlapping, appears a Lagrange multiplier which makes the resolution more costly. This additional unknown will be avoided with an overlap.

For now, the theory is done for the scalar acoustic waveguide and the method has been implemented in the Melina code. The extension to the elastic case is in progress.

6.3.5. An alternative to DTN maps in waveguides

Participants: Anne-Sophie Bonnet-Ben Dhia, Guillaume Legendre.

We are interested by the treatment of the radiation condition at infinity for the numerical solution of a problem set in an unbounded waveguide. We have proposed an alternative to the classical approach involving a modal expression of Dirichlet-to-Neumann (DtN) operators. This new method is particularly simple to implement since it only requires to solve several times a boundary value problem with local boundary conditions. In the case of an acoustic waveguide, we have proved that the corresponding approximate solution is comparable in accuracy to the one obtained by truncating the infinite series in the DtN maps. The number of linear systems to invert, which has to be greater than the number of propagative guided modes, can be significantly reduced by combining the approach with the perfectly matched layer (PML) technique. It works even in elastic waveguides, despite existence of the so-called backward waves which are known to make the PMLs inefficient, when used alone.

6.4. Waveguides, resonances, and scattering theory

6.4.1. Modelling of non-homogeneous lossy coaxial cable for time domain simulation.

Participants: Sébastien Impériale, Patrick Joly.

In this work, we focus on the time-domain simulation of the propagation of electromagnetic waves in non-homogeneous lossy coaxial cables. This question has been motivated by our collaboration with CEA-LIST about the numerical modeling of non-destructive testing (for the detection of cracks in metallic bodies for instance) by ultrasound and more precisely the modeling of piezo-electric transducers (see section 6.1.6). The complete description of such devices often requires an accurate modeling of the supply process, which includes the propagation of the electric current along coaxial cables. This question appears as an independent sub-problem.

The main characteristic of such coaxial cables is that their transverse directions are very small with respect to their length as well as the wavelength. As a consequence, one would like to use a simplified 1D model as an effective (or homogenized) model for electromagnetic propagation. In this work, we construct and justify rigorously such a model by way of an asymptotic analysis of time harmonic 3D Maxwell’s equations in such a structure. The effective model appears as a generalized wave equation with additional time convolution terms that take into account electric and magnetic losses. By this way, we justify and extend some models proposed in the electrical engineering literature, in particular the well-known telegraphist’s equation. The properties of our limit model in time domain has been analyzed and a stable discretization process has been proposed. Numerical simulation in academic simulations exhibit some exotic phenomena of “dispersive dissipation”.

Our further investigations of the subject will be developed in the framework of the new ANR project SODDA in collaboration with the team Sysiphe (M. Sorine).
6.4.2. Study of lineic defect in periodic media

**Participant:** Sonia Fliss.

This works deals with one dimensional infinite perturbation - namely line defects - in periodic media. In optics, such defects are created to construct an (open) waveguide that concentrates light. The existence and the computation of the eigenmodes is a crucial issue. This is related to a selfadjoint eigenvalue problem associated to a PDE in an unbounded domain (in the directions orthogonal to the line defect), which makes both the analysis and the computations more complex. Using a Dirichlet-to-Neumann (DtN) approach, we show that this problem is equivalent to one set on a small neighborhood of the defect. In opposition to existing methods, this method is exact but there is a price to be paid: the reduction of the problem leads to a nonlinear eigenvalue problem of a fixed point nature. An article presenting the method and its properties is being written, the numerical study is in progress in collaboration with Kersten Schmidt and Dirk Klindworth from the Technische Universität Berlin.

Our further investigations of the subject will be developed in the framework of a new DGA project.

6.4.3. A new approach for the numerical computation of non linear modes of vibrating systems

**Participants:** Anne-Sophie Bonnet-Ben Dhia, Jean-François Mercier.

A collaboration with Cyril Touzé and François Blanc (Unité de Mécanique, ENSTA). The simulation of vibrations of large amplitude of thin plates or shells requires the expensive solution of a non-linear finite element model. The main objective of the proposed study is to develop a reliable numerical method which reduces drastically the number of degrees of freedom. The main idea is the use of the so-called non-linear modes to project the dynamics on invariant subspaces, in order to generate accurate reduced-order models. Cyril Touzé from the Unité de Mécanique of ENSTA has derived an asymptotic method of calculation of the non-linear modes for both conservative and damped systems. But the asymptotically computed solution remains accurate only for moderate amplitudes. This motivates the present study which consists in developing a numerical method for the computation of the non-linear modes, without any asymptotic assumption. This is the object of a collaboration with Cyril Touzé, and new results have been obtained during the post-doc of François Blanc in the Unité de Mécanique of ENSTA. The partial differential equations defining the invariant manifold of the non-linear mode are seen as a vectorial transport problem: the variables are the amplitude and the phase \((a, \varphi)\) where the phase \(\varphi\) plays the role of the time. In the case of conservative systems, a finite difference scheme is used and an iterative algorithm is written, to take into account the \(2\pi\)-periodicity in \(\varphi\) which is seen as a constraint. An adjoint state approach has been introduced to evaluate the gradient of the cost function. The method has been validated in a simple example with two degrees of freedom. Good agreement with an alternative method, the continuation of periodic solutions method, has been found. Currently the method is extended to the case of damped systems. The main difficulty is that, due to a change of variables, the \(2\pi\)-periodicity does not hold anymore and new constraints more complicated to implement must be considered. Numerical implementation is still under progress.

6.5. Asymptotic methods and approximate models

6.5.1. Effective boundary conditions for thin periodic coatings

**Participants:** Mathieu Chamaillard, Patrick Joly.

This topic is developed in collaboration with H. Haddar (DEFL, INRIA Saclay) can be seen as a continuation of the PhD thesis of B. Delourme (see the activity report of last year) on effective transmission conditions for thin rough interfaces. On this last subject, the mathematical analysis of such transmission conditions in the (difficult) case of 3D Maxwell’s equations has been completed and submitted for publication.
We are now coming back to the more traditional issue of effective or approximate boundary conditions for simulating thin periodic coatings at the surface of a diffractive obstacle. This subject has already been more widely investigated, in particular in France (see the works by Achdou, Ammari and their collaborators for instance). However, we attack, with the PhD thesis of M. Chamaillard, supported by a DGA/INRIA scholarship, various new aspects of the problem, namely:

- the treatment of surfaces of general geometry,
- the use of non standard materials, such as ferromagnetic materials, for the coating,
- the construction of higher order impedance conditions.

This is motivated by various recent progress in the domain of stealth technology and we hope to develop a collaboration in this domain with CEA-DAM (CESTA and Le Ripault).

6.5.2. Elastic wave propagation in strongly heterogeneous media

Participants: Patrick Joly, Simon Marmorat.

This subject enters our long term collaboration with CEA-LIST on the development on numerical methods for time-domain non destructive testing experiments using ultra-sounds. This is also the subject of the PhD thesis of Simon Marmorat. Our objective is to develop an efficient numerical approach for the propagation of elastic waves in a medium which is made of many small inclusions / heterogeneities embedded in a smooth (or piecewise smooth) background medium, without any particular assumption (such as periodicity) on the spatial distribution of these heterogeneities. Our idea is to exploit the smallness of the inclusions (with respect to the wavelength in the background medium) to derive a simplified approximate model in which each inclusion would de described by very few parameters (functions of time) coupled to the displacement field in background medium for which we could use a computational mesh that ignores the presence of the heterogeneities. For deriving such a model, we intend to use and adapt the asymptotic methods previously developed at Poems (such as matched asymptotic expansions).

6.5.3. Approximate models in aeroacoustics

Participants: Anne-Sophie Bonnet-Ben Dhia, Patrick Joly, Guillaume Legendre, Ricardo Weder.

This topic concerns the 2D acoustic propagation in presence of a mean flow, modeled for instance by the Galbrun equation. We had previously derived effective boundary conditions taking into account the boundary layers of the mean flow near a rigid or treated boundary. These boundary conditions are in general non local with respect to the normal coordinate inside the boundary layer. However when the Mach profile in the shear layer is piecewise linear, the condition can be replaced by a system of 1D advection equations, which are coupled with the Galbrun equation in the 2D domain. We have derived a variational formulation for this model, in time-harmonic regime and for the case of a rigid boundary. This formulation has been implemented in the Melina code: the first results are promising and the validation, by comparison to the solution obtained by a full discretization of the shear layer, is in progress.

6.6. Imaging and inverse problems

6.6.1. Sampling methods in waveguides

Participants: Laurent Bourgeois, Eric Lunéville, Alexandre Routier.

We have derived a modal formulation of sampling methods (both the Linear Sampling Method and the Factorization Method) in an acoustic waveguide when the obstacle to recover is a set of cracks. This was the subject of the Master internship of Alexandre Routier. For such particular obstacle, we have analyzed the importance of the test function we introduce in the sampling method if we a priori know the type of boundary condition we have on the lips of the crack, both from the theoretical and the numerical point of view. Besides we have proved that in our modal formulation, the Factorization Method is applicable by using the same data as those used in the Linear Sampling Method, which is a novelty as concerns sampling methods in waveguides. The Linear Sampling Method has been extended to the elastic case, for which the usual obstacle is a set of traction free cracks. This makes the choice of the test function crucial, and we have emphasized such fact on numerical examples.
6.6.2. Inverse scattering with generalized impedance boundary conditions

Participants: Laurent Bourgeois, Nicolas Chaulet.

This work is a collaboration between POEMS and DEFI projects (more precisely Houssem Haddar) and constitutes the subject of the PhD thesis of N. Chaulet. We are concerned with the identification of some obstacle and some Generalized Impedance Boundary Conditions (GIBC) on the boundary of such obstacle from far field measurements. The GIBCs are approximate models for thin coatings or corrugated surfaces. During this last year, we have completed the computation of the partial derivatives of the far field with respect to the unknowns, among which is the boundary of the obstacle, and we have implemented many numerical experiments. In particular, we have shown the efficiency of the method consisting in approximating a perfect conductor which is coated with a thin dielectric layer of variable width by a second order GIBC in order to retrieve the obstacle, as well as the refraction index and the width of the layer.

6.6.3. Detection of targets using time-reversal

Participants: Maxence Cassier, Patrick Joly, Christophe Hazard.

This topic concerns the studies started last year about time-reversal in the context of Maxence Cassier’s thesis. The main question is to generate a time-dependent wave that focuses on one given scatterer not only in space, but also in time. Our recent works concern two items. On one hand, we have proposed a way to construct such a focusing wave which does not require an a priori knowledge of the location of the obstacle. This wave is represented by a suitable superposition of the eigenvectors of the so-called time-reversal operator in the frequency domain. Numerical results show the focusing properties of such a wave. On the other hand, we try to understand how to translate the physical idea of “focusing” into mathematical terms. We have proposed an energy criterion which can be used in numerical experiments in order to evaluate the quality of the focus. The question is to relate such a criterion with the construction of the above mentioned focusing wave. Works on this topic are in progress.

6.6.4. Interior transmission problem

Participants: Anne-Sophie Bonnet-Ben Dhia, Lucas Chesnel.

This work is a collaboration with Houssem Haddar from the DEFI project. The interior transmission problem plays an important role in the inverse scattering theory for inhomogeneous media. In particular, it arises when one is interested in the reconstruction of an inclusion embedded in a background medium from multi-static measurements of diffracted fields at a given frequency. Physically, it is important to prove that, for a given frequency, there are no waves which do not scatter. Mathematically, this last property boils down to say that the frequency is not a transmission eigenvalue, that is, an eigenvalue of the interior transmission problem. An important issue is to prove that transmission eigenvalues form at most a discrete set with infinity as the only accumulation point. This is not trivial because the operator associated with this problem exhibits a sign changing in its principal part and its study is not standard. Using the T-coercivity approach, we proved the discreteness under relatively weak assumptions. In particular, the simple technique we proposed allows to treat cases, which were not covered by existing methods, where the difference between the inclusion index and the background index changes sign.

6.6.5. Flaw identification using elastodynamic topological derivative

Participant: Marc Bonnet.

In collaboration with Cédric Bellis (Columbia Univ. USA), Bojan Guzina (Univ. of Minnesota, USA). The concept of topological derivative (TD) quantifies the perturbation induced to a given cost functional by the nucleation of an infinitesimal flaw in a reference defect-free body, and may serve as a flaw indicator function. In this work, the TD is derived for three-dimensional crack identification exploiting over-determined transient elastodynamic boundary data. This entails in particular the derivation of the relevant polarization tensor, here given for infinitesimal trial cracks in homogeneous or bi-material elastic bodies. Simple and efficient adjoint-state based formulations are used for computational efficiency, allowing to compute the TD field for arbitrarily shaped elastic solids. The latter is then used as an indicator function for the spatial location of the sought
crack(s). This approach, which allows a qualitative reconstruction of cracks in terms of their location but also their orientation (utilizing the fact that the polarization tensor depends on the normal to the trial small crack), has been implemented within a conventional FEM platform. A standard Newmark unconditionally-stable time-marching scheme is used for simulating data, and for computing the free and adjoint solutions used in the evaluation of the TD field. Extensive 3D time-domain numerical experiments for the detection of cracks buried either in a homogeneous pipe-like structure or on the interface between two sandwiched plates highlight its usefulness and performance. The application of TD to flaw identification has thus far rested upon a heuristic basis. Its justification in limiting situations such as the Born approximation is currently being investigated.

6.6.6. Topological derivative in anisotropic elasticity

Participant: Marc Bonnet.

In collaboration with Gabriel Delgado (CMAP, Ecole Polytechnique). This work addresses the current lack of a comprehensive treatment of the topological derivative for anisotropic elasticity, by addressing the case where both the background material and the trial small inhomogeneity have arbitrary anisotropic elastic properties. Accordingly, a formula for the topological derivative of any cost functional defined in terms of regular volume or surface densities depending on the displacement is established, by combining small-inhomogeneity asymptotics and the adjoint solution approach. The latter feature makes the proposed result simple to implement and computationally efficient. Both three-dimensional and plane-strain settings are covered; they differ mostly on details pertaining to the elastic moment tensor. This result achieves a direct generalization to the fully anisotropic case of previously-known formulations for isotropic elasticity. Moreover, the main properties of the EMT, a critical feature of any elastic topological derivative formula, are studied for the fully anisotropic case, generalizing available results on the isotropic case. Finally, further generality is achieved by also deriving the topological derivative of strain energy-based cost functionals, which depend on the displacement gradient. This case, seldom addressed so far, requires a specific, and separate, treatment. Applications of these results include topology optimization of composite structures (a topic currently pursued by G. Delgado) or flaw identification using experimental data from nondestructive testing.

6.6.7. Energy functionals for elastic medium reconstruction using transient data

Participant: Marc Bonnet.

In collaboration with Wilkins Aquino (Cornell Univ., USA).

Energy-based misfit cost functionals, known in mechanics as error in constitutive relation (ECR) functionals, are known since a long time to be well suited to (electrostatic, elastic,...) medium reconstruction. In this ongoing work, a transient elastodynamic version of this methodology is developed, with emphasis on its applicability to large time-domain finite element modeling of the forward problem. The formulation involves coupled transient forward and adjoint solutions, a fact which greatly hinders large-scale computations. A computational approach combining an iterative treatment of the coupled problem and the adjoint to the discrete Newmark time-stepping scheme is found to perform well on large FE models, making the time-domain ECR functional a worthwhile tool for medium identification.

6.6.8. Accelerated boundary element method for diffuse optical imaging

Participant: Marc Bonnet.

In collaboration with Simon Arridge and Josias Elisee (University College London, UK). Numerical methods for calculating forward models of light propagation in tissue are extensively used in diffuse optical imaging (DOI). DOI involves a Helmholtz-type PDE with a complex-valued wavenumber. It requires modelling large optical regions whose parameters are known and piecewise constant. The boundary element method (BEM) answers this need and avoids the detailed interior meshing of these regions. The single-level Fast Multipole Method has been applied for solving the DOI governing equation, allowing substantial reduction of computational costs. The enhanced practicability of the BEM in DOI was demonstrated through test examples on single-layer problems, where two-digit reduction factors on solution time are achieved, and on a high-resolution version of a three-layered neonate’s head.
6.6.9. **High-Velocity Estimates and Inverse Scattering for Quantum N-Body Systems with Stark Effect**  
**Participants:** Ricardo Weder, Gerardo Daniel Valencia.

In an $N$-body quantum system with a constant electric field, by inverse scattering, we uniquely reconstruct pair potentials, belonging to the optimal class of short-range potentials and long-range potentials, from the high-velocity limit of the Dollard scattering operator. We give a reconstruction formula with an error term.

6.6.10. **Small-Energy Analysis for the Matrix Schrödinger Operator on the Half-Line**  
**Participant:** Ricardo Weder.

In collaboration with Tuncay Aktosun (University of Texas Arlington) and Martin Klaus (Virginia Tech). The matrix Schrödinger equation with a selfadjoint matrix potential is considered on the half line with the most general selfadjoint boundary condition at the origin. When the matrix potential is integrable and has a first moment, it is shown that the corresponding scattering matrix is continuous at zero energy. An explicit formula is provided for the scattering matrix at zero energy. The small-energy asymptotics are established also for the related Jost matrix, its inverse, and various other quantities relevant to the corresponding direct and inverse scattering problems.

6.7. **Other topics**

6.7.1. **Fast non-overlapping Schwarz domain decomposition methods for the neutron diffusion equation**  
**Participant:** Patrick Ciarlet.

A collaboration with Erell Jamelot (CEA Saclay/DEN). Investigating numerically the steady state of a nuclear core reactor can be very expensive, in terms of memory storage and computational time. In order to address both requirements, one can use a domain decomposition method, which is then implemented on a parallel computer.

We model the problem using a mixed approach, which involves a scalar flux and a vector current. The equivalent variational formulation is then discretized with the help of Raviart-Thomas-Nédélec finite elements.

The domain decomposition method is based on the Schwarz iterative algorithm with Robin interface conditions to handle communications. This method is analyzed from the continuous to the discrete point of views: well-posedness, convergence of the finite element method, optimality of the parameter appearing in the Robin interface condition and algorithms. Numerical experiments carried out on realistic 3D configurations using the APOLLO3©code (of CEA/DEN) show the parallel efficiency of the algorithm.

6.7.2. **Equivalent local boundary conditions for the Monge Kantorovitch Mass Transport problem**  
**Participant:** Jean-David Benamou.

This work is done in collaboration with Adam Oberman Brittany Froese from Simon Fraser University, Vancouver. In the last 20 years, the Monge Kantorovich Optimal Transport problem (OTP) and its relationship with Partial Differential Equations (PDE) experienced a spectacular research revival (the Fields medal was awarded to Cedric Villani in 2010 partly for his contributions to OTP). Applications appeared in fields as diverse as meteorology, medical image processing, astronomy and economy. This new area offers numerical challenges which go beyond current knowledge in the field. Novel computational tools are needed. The OTP can also be reformulated as a Monge-Ampere (MA) PDE with non standard/non local boundary conditions (BC). We would like to use the new and efficient wide-stencil Finite Difference MA solver developed by Oberman and Froese but we do not know how to deal with the OTP equivalent BCs. We pursue the investigation of the two following innovative strategies : 1. Iteratively construct Neumann BCs such that the solution of the MA equation satisfies the OTP BC in the limit. 2. Merge the data in all of space and design simplified asymptotic local BCs at infinity which can be used to formulate local transparent BC on a truncated domain.
6.7.3. **Topological Effects in Quantum Mechanics and High-Velocity Estimates**

**Participant:** Ricardo Weder.

This work is done in collaboration with Miguel Ballesteros. High-velocity -or high-energy- estimates for scattering solutions to the Schrödinger equation are important for many reasons. For example, in topological effects in quantum mechanics, where the space accessible to the particles has a non-trivial topology, like, for example, in the celebrated magnetic Aharonov-Bohm effect, where an electron is constrained to be on the exterior of a torus that contains a magnetic flux inside. Here, the solution acquires a phase if the electron travels inside the hole of the magnet and, on the contrary, it acquires no phase if the particle travels outside the hole. We obtain precise high-velocity estimates for the scattering solutions, that prove that quantum mechanics actually predicts the existence of the magnetic Aharonov-Bohm effect, under the conditions of the celebrated Tonomura et al. experiments. Moreover, in the case of the electric Aharonov-Bohm effect, we provide precise conditions for the validity of the Aharonov-Bohm Ansatz and we give a rigorous proof that quantum mechanics predicts the existence of this effect.

6.7.4. **Entanglement Creation in Low-Energy Scattering**

**Participant:** Ricardo Weder.

We study the entanglement creation in the low-energy scattering of two particles of mass, \(m_1, m_2\) in three dimensions. We consider a general class of interaction potentials that are not required to be spherically symmetric. The incoming asymptotic state, before the collision, is a product of two normalized Gaussian states with the same variance, \(\sigma\), and opposite mean momentum. After the scattering the particles are in the outgoing asymptotic state that is not a product state. We take as a measure of the entanglement created by the collision the purity of one of the particles in the outgoing asymptotic state. In the incoming asymptotic state the purity is one. We provide a rigorous explicit computation, with error bound, of the leading order of the purity at low-energy. The leading order depends strongly in the difference of the masses. The entanglement takes its minimum when the masses are equal, and it increases rapidly with the difference of the masses. It is quite remarkable that the anisotropy of the potential plays no role, on spite of the fact that entanglement is a second order effect.

6.7.5. **Open Scattering Channels in Manifolds with Horns**

**Participant:** Ricardo Weder.

This work is done in collaboration with Olaf Post (Humboldt University, Berlin) and Rainer Hempel (Mount Allison University, Sackville New Brunswick). In the framework of time-dependent geometric scattering theory, we study the existence and completeness of the wave operators for perturbations of the Riemannian metric for the Laplacian on a complete manifold of dimension \(n\). The smallness condition for the perturbation is expressed in purely geometric terms using the harmonic radius; therefore, the size of the perturbation can be controlled in terms of local bounds on the radius of injectivity and the Ricci-curvature, and no global assumption is needed. As an application of these ideas we obtain a stability result for the scattering matrix. As a consequence we find that a scattering channel which interacts with other channels preserves this property under small perturbations.
5. New Results

5.1. Numerical methods

5.1.1. Finite volume methods in curvilinear coordinates

**Participants:** Hervé Guillard, Boniface Nkonga, Afeintou Sangam.

Finite volume methods are specialized numerical techniques for the solving of divergence equations in strong conservation forms of the form

\[
\frac{\partial S}{\partial t} + \text{div}T = 0
\]  

(7)

where \( S \) is a scalar or a vector while \( T \) is a vector or a second-order tensor. Using textbook formulas for the expression of the divergence operator in curvilinear coordinates, the use of these coordinate systems instead of the Cartesian one can lead to a loss of the strong conservation form of the equations and introduce a source term in (7). Actually, this is unnecessary and one can show that whatever the system of curvilinear coordinate used, there exists a strong conservation law form of the system. However, when vector equations have to be considered (that is if \( S \) is a vector and \( T \) a tensor), it is necessary to extract from (7) scalar equations for the components of the vector \( S \) and this may destroy the strong conservation form of the equation. Following the work done in [12] where a general method (i.e that does not depend on the curvilinear system used) based on the projection of the discretized vector system have been designed, we have studied this year its application to cylindrical coordinates in the case where the geometry is a torus. This approach is robust and accurate for problems that take place for instance inside tokamak devices for magnetic confinement fusion or in toroidal plasmas occurring in stars and galaxies to take another examples. The method is now implemented in the PlaTo software.

5.1.2. Entropy preserving schemes for conservation laws

**Participants:** Christophe Berthon [University of Nantes], Bruno Dubroca [CEA/DAM/CESTA and University of Bordeaux 1], Afeintou Sangam.

Entropy preserving schemes for conservation laws

In collaboration with C. Berthon of University of Nantes, and B. Dubroca of CEA/DAM/CESTA and University of Bordeaux 1, we have established a new technique that proves discrete entropy inequalities of finite volume methods to approximate conservation laws. This technique is free of additional numerical models such as kinetic and relaxation schemes. Moreover, our results leads to a full class of entropy preserving schemes for general Euler equations [11]. This proposed thechnique has been successfully applied to two intermediates states scheme for 10-moments equations with laser source-term in context of Inertial Fusion Confinement. Moreover, the derived procedure is now extended to Saint-Venant model.

5.1.3. Mesh adaptation Methods

**Participants:** Anca Belme [Projet Tropics], Hubert Alcin [Projet Tropics], Alain Dervieux, Frédéric Alauzet [Projet Gamma, INRIA-Rocquencourt].

This activity results from a cooperation between Gamma, Tropics, Pumas, and Lemma company. See details in Tropics and Gamma activity reports. Its concerns Pumas’s subject through the current applications of mesh adaptation to flows with interfaces and the starting application of mesh adaptation to Large Eddy Simulation. It is also planned to use mesh adaptation for simplified plasma models in the context of ANEMOS ANR project.
5.1.4. Parallel CFD algorithms

**Participants:** Hubert Alcin [Tropics], Olivier Allain [Lemma], Anca Belme [Tropics], Marianna Braza [IMF-Toulouse], Alexandre Carabias [Tropics], Alain Dervieux, Bruno Koobus [Université Montpellier 2], Carine Moussaed [Université Montpellier 2], Hilde Ouvrard [IMF-Toulouse], Stephen Wornom [Lemma].

Pumas is associated to the ANR ECINADS project started in end of 2009, devoted to the design of new solution algorithms for unsteady compressible flows, adapted to scalable parallelism and to reverse (adjoint) Automatic Differentiation. See in the activity report of Tropics. The newer two-level deflation algorithm is currently applied to a simplified plasma model in the context of ANEMOS ANR.

5.2. Plasma physics

5.2.1. Analysis of the drift approximation

**Participants:** Hervé Guillard, Afeintou Sangam, Philippe Ghendrih [IRFM, CEA Cadarache], Yanick Sarazin [IRFM, CEA Cadarache], Patrick Tamain [IRFM, CEA Cadarache].

Drift approximation considers the slow evolution of the fields in the vicinity of a tokamak equilibrium. These models are typically used to study the micro-instabilities that are believed to be responsible of turbulent transport in tokamaks. Since the drift asymptotic uses a “slow” scaling of the velocity field, the resulting models are significantly different from the full MDH models. This is particularly true with respect to the computation of the electric field that is given by an Ohm’s law in MHD models whereas it is computed by a vorticity-like evolution equation in drift approximations. Drift asymptotic models are extremely interesting from a computational point of view since they save substantial CPU time and computer memory. However, the mathematical and numerical properties of these models are essentially unknown. We have begun a detailed study of the derivation of these models from two-fluid Braginskii-type models in order to establish the range of applicability of these asymptotic models, understand their mathematical properties and relations with the reduced MHD models and design appropriate numerical methods for their approximations.

5.2.2. Stabilized C1-Finite Element Method for MHD

**Participants:** Boniface Nkonga, Marie Martin.

Reduced MHD models are often used in plasma physics and therefore fast compressible waves are not taken into account. In the context of Elms instabilities investigations, full and extended MHD models are to be considered within the framework of high order finite element approximation. In order to obtain predictive simulations with reduced, full and extended MHD models, it is crucial to design numerical strategies that can face some difficulties related to the use of the classical Galerkin methods for convection dominated flows. We have developed a general VMS stabilization strategy for time dependent implicit scheme, which can be applied to MHD models in order to preserved the global accuracy of the initial Galerkin formulation and enforce physical properties as monotony and positivity. Higher order of continuity shape functions are important for accuracy and also help to obtain more robustness of the stabilization as there is no more singularities on elements edges. Numerical implementation and preliminary validations has been performed using C1-Bell shapes functions for triangular meshes. Order five accuracy of the theory is recovered for specific boundary condition. Improvement for general boundary condition is still an open issue. The next step is to combine this Bell shapes functions in the poloidal direction with B-splines functions to achieve accurate representation of complex torus as ITER.

5.2.3. Two fluid modelling of the Scrape-Off-Layer Plasma

**Participants:** Audrey Bonnement, Hervé Guillard, Richard Pasquetti.
A two fluid physical model has been developped in close collaboration with researchers from IRFM. It is based on an hypothesis of stationary magnetic field and the electrostatic and electroneutrality assumptions. However the usual drift assumption, e.g. used in the CEA code TOKAM3D (thesis of P. Tamain), is not used. On the basis of the conservation equations of density, electron and ion velocities, electron and ion temperatures and electrical charges, a set of 10 nonlinear coupled partial differential equations (PDE) can be set up. Our investigations rely on the development of two solvers for this set of PDE. In the frame of her thesis, A. Bonnement (co-direction H. Guillard and R. Pasquetti, financial support of INRIA and PACA region, industrial partner ASSYSTEM) uses a Finite volume / element (FV/FE) approach. The other code, developed by R. Pasquetti, focuses on the use of high order approximations: a spectral element method (SEM) is used in the poloidal plane whereas Fourier expansions are used in the toroidal direction. Each of these codes has allowed in 2010 to solve strongly anisotropic diffusion equations in 2D and axisymmetric 3D geometries for the FV/FE code and in 2D and fully 3D for the Fourier-SEM code. The FV/FE code has also been used to carry out a study of radiative layers evolution in a 2D annular configuration but also in a realistic 3D ITER configuration.

Works carried out during the year 2011 are described hereafter: – On the basis of a Godunov scheme, the FV/FEd approach has been extended to solve the axisymmetric Euler, Navier-Stokes or Braginskii like systems in the jet-tokamak geometry. With respect to Navier-Stokes, the Braginskii system is characterized by anisotropies in the transport coefficients. The mesh is unstructured : A finite element approximation is used for the diffusion terms whereas for the convective terms a finite volume approximation is used on the dual mesh. We have especially focused on the treatment of the toroidal geometry. We have also focused on the implementation of the so-called Bohm conditions, which are enforced by imposing that at the limiter the fluid velocity is colinear to the magnetic field and that the « parallel Mach number » equals or is greater than one. In these studies the governing equations are completed by constant force terms in order to model the Lorentz forces as well as sources of mass and energy. Such forcing terms allow to preserve an equilibrium state, e.g. obtained on the basis of simulations that make use of the drift assumption (see next paragraph). We have then introduced perturbations of this equilibrium state to study the evolution of the different variables (density, velocity and temperature). Various perturbations have been used to this end, fully random in order to check the stability of the equilibrium or on the contrary localized, in order to roughly model the injection of pellets inside the SOL. It is planned that Audrey Bonnement will defend her thesis in Spring 2012.

The SEM-Fourier 3D code has been extended to solve the full set of governing equations. The unknowns are then the density, the velocity of the center of mass, the electric current and potential, the ion and electron internal energies. In time, this set of PDE is solved by using an IMEX (Implicit – Explicit) approach, based on the combination of an explicit Runge-Kutta scheme for the flux terms and on a DIRK (diagonal implicit Runge Kutta) for the Lorentz terms, which indeed lead to an unconditionnaly unstable scheme if treated explicitly. A projection method is used to enforce the divergence free constraint of the current, so that an additional solve of a Poisson equation is required to obtain the potential. The Bohm conditions are here implemented by enforcing that the ion pressure is such that the parallel Mach number shows the expected value, i.e. equal or greater than one. The main difficulty that we presently meet is that the initial condition that we use does not correspond to an equilibrium state, so that instabilities quickly develop till yielding an unsteady flow not consistent at the limiter with the Bohm condition. Detailed analyses of these results are presently carried out with Sebastian Minjeaud (new CNRS researcher of the LJAD) to provide relevant explanations of the observed phenomena.

5.2.4. Drift approximation modelling of the Scrape-Off-Layer Plasma

Participants: Marco Bilanceri, Hervé Guillard.

Based on a fluid model using the drift velocity approximation, a simulation method have been designed to compute the flow in the scrape-off-layer of a Tokamak. The variables used by the model are the particle number, the parallel (to the magnetic field) velocity and the electric potential. The spatial approximation uses a finite volume/finite element approach and is therefore easy to apply to complex geometries. Bohms boundary conditions are used on the divertor plates of the machine. The figure 1. shows the density in a poloidal cut where the influence of the separatrix can be clearly seen.
Figure 1. Density plot in a poloidal section of a jet-like Tokamak.
5.3. Fluid Turbulence

5.3.1. Hybrid RANS-LES models

Participants: Anca Belme [Tropics], Alain Dervieux, Bruno Koobus [University of Montpellier 2], Carine Moussaed [University of Montpellier 2], Hilde Ouvrard [IMF-Toulouse], Maria-Vittoria Salvetti [University of Pisa], Stephen Wornom [Lemma].

The purpose of our works in hybrid RANS/LES is to develop new approaches for industrial applications of LES-based analyses. In the foreseen applications (aeronautics, hydraulics), the Reynolds number can be as high as several tenth millions, a far too large number for pure LES models. However, certain regions in the flow can be much better predicted with LES than with usual statistical RANS (Reynolds averaged Navier-Stokes) models. These are mainly vortical separated regions as assumed in one of the most popular hybrid model, the hybrid Detached Eddy Simulation model. Here, “hybrid” means that a blending is applied between LES and RANS. The french-italian team has designed a novel type of hybrid model. This year, the new model has been adapted to very high Reynolds number. Our benchmark is the flow past a circular cylinder, an ECINADS test case. Reynolds number as high as 3 Millions could be passed with good prediction of main properties like mean drag, root mean square of lift fluctuation, base pressure.

5.3.2. Acoustics

Participants: Anca Belme [Tropics], Illya Abalakin [IMM-Moscou], Alain Dervieux [Tropics], Alexandre Carabias.

A method for the simulation of aeroacoustics on the basis of these models has been designed and developed by a cooperation between the Computational Aeroacoustics Laboratory (CAL) of Institute for Mathematical Modeling at Moscow and INRIA. Further applications has been developed by the Russian team from the two common numerical scheme, the Mixed-Element-Volume at sixth-order, and the quadratic reconstruction scheme. This year the cooperation is concentrated on the study by Alexandre Carabias of a new quadratic reconstruction scheme, which extends the one developed by Hilde Ouvrard and Ilya Abalakin. A second research topic was the calculation of acoustic propagation with unsteady mesh adaptation.

5.4. Environmental flows

5.4.1. Mobile bed and sediment transport

Participants: Hervé Guillard, Boniface Nkonga, Marco Bilanceri, Maria-Vittoria Salvetti [University of Pisa, Italy], Imad Elmahi [University of Oudja, Morocco].

The numerical approximation of a model coupling the shallow-water equations with a sediment transport equation for the morphodynamics have been studied. In shallow-water problems, time advancing can be carried out by explicit schemes. However, if the interaction with the mobile bed is weak, the characteristic time scales of the flow and of the sediment transport can be very different introducing time stiffness in the global problem. For this case, it is of great interest to use implicit schemes. The time integration strategy that we have devised is based on a defect-correction approach and on a time linearization, in which the flux Jacobians are computed through automatic differentiation. The aim of the present work is to investigate the behaviour of this time scheme in different situations related to environmental flows. This work has been presented in [18] and [19].
6. New Results

6.1. Theoretical and Methodological Developments

Participants: Cédric Joncour, Andrew Miller, Arnaud Pêcher, Pierre Pesneau, Ruslan Sadykov, Gautier Stauffer, François Vanderbeck.

We made progress in the development of theory and algorithms in the area of “Reformulation and Decomposition Approaches for MIP”, “Mixed Integer Nonlinear Programming”, and “Polyhedral Combinatorics and Graph Theory”.

6.1.1. Column Generation for Extended Formulations

Working in an extended variable space allows one to develop tight reformulations for mixed integer programs. However, the size of the extended formulation grows rapidly too large for a direct treatment by a MIP-solver. Then, one can use projection tools to derive valid inequalities for the original formulation and implement a cutting plane approach. Or, one can approximate the reformulation, using techniques such as variable aggregation or by reformulating a submodel only. Such approaches result in outer approximation of the intended extended formulation. The alternative considered in [28], [25] is an inner approximation obtained by generating dynamically the variables of the extended formulation. It assumes that the extended formulation stems from a decomposition principle: a subproblem admits an extended formulation from which an extended formulation for the original problem can be derived. Then, one can implement column generation for the extended formulation of the original problem by transposing the equivalent procedure for the Dantzig-Wolfe reformulation. Pricing subproblem solutions are expressed in the variables of the extended formulation and added to the current restricted version of the extended formulation along with the subproblem constraints that are active for the subproblem solution.

Our paper [28], [25] revisits the column-and-row generation approach. Our purpose is to show light on this approach, to emphasize its wide applicability, and to present it with a new angle as a method that is natural when considering a problem reformulation based on any extended formulation of a subproblem, whether it yields the subproblem integer hull or just an approximation of it. In the spirit of [80], column-and-row generation is viewed herein as a generalization of standard column generation, the latter being based on a specific subproblem extended formulation. This generic view not only highlights the scope of applicability of the method, but it also leads to a more general termination condition than the traditional reduced cost criteria and to theoretically stronger dual bounds (observing that solving the integer subproblems yields Lagrangian dual bounds that might be tighter than the extended formulation LP bound). We highlight a key motivation for working in the extended space: there arises natural recombinations of previously generated columns into new subproblem solutions, which results in an acceleration of the convergence. We point out that lifting the master program in the variable space of the extended formulation can be done while carrying pricing in the compact variable space of the original formulation, using any oracle.

With [28], [25], we establish the validity of the column-and-row generation algorithm in a form that encompass all special cases of the literature. The analysis therein should help practitioners to evaluate whether this alternative procedure has potential to outperform classical column generation on a particular problem. Our numerical experiments highlight a key observation: lifting pricing problem solutions in the space of the extended formulation permits their recombination into new subproblem solutions and results in faster convergence.
6.1.2. Primal Heuristics for Branch-and-Price

Our goal is to exploit global optimization decomposition approaches to retrieve very good feasible solution to large scale problem. This required extending primal heuristic paradigms to the context of dynamic generation of the variables of the model. We highlight an important fact: such generic tools typically performs better than problem specific meta-heuristics, in terms of solution quality and computing times. Based on our application specific experience with these techniques [65], [67], [86], [87], and on a review of generic classes of column generation based primal heuristics, in [58], we are developing a full blown review of such techniques, completed with new methods and an extensive numerical study. This research is being carried on in collaboration with the members of the associated team project, SAMBA.

Significant progress has been achieved in developing generic primal heuristics that made their way into commercial mixed integer programming (MIP) solvers. Extensions to the context of a column generation solution approach are considered by our team, in search for generic black-box primal heuristics for use in Branch-and-Price approaches. As the Dantzig-Wolfe reformulation is typically tighter than the original compact formulations, techniques based on rounding its linear programming solution have better chance to yield good primal solutions. The aggregated information built into the column definition and the price coordination mechanism provide a global view at the solution space that may be lacking in somewhat more “myopic” approaches based on compact formulations. However, the dynamic generation of variables requires specific adaptation of heuristic paradigms. We focus on “diving” methods and considered their combination with sub-MIPing, relaxation induced neighborhood search, and truncated backtracking using a Limited Discrepancy Search. These add-ons serves as local-search or diversification/intensification mechanisms. We also consider feasibility pump approaches. The methods are numerically tested on standard models such as Cutting Stock, Vertex Coloring, Generalized Assignment, Lot-Sizing, and Vehicle Routing problems.

6.1.3. Combining Bender’s and Dantzig-Wolfe Decomposition

In the follow-up of [56], [88], [89], [90], we are finalizing our work on the combination of Dantzig-Wolfe and Bender’s decomposition: Bender’s Master is solved by column generation [91]. The application we considered is a multi-layer network design model arising from a real-life telecommunication application where traffic routing decisions imply the installation of expensive nodal equipment. Customer requests come in the form of bandwidth reservations for a given origin destination pair. Bandwidth requirements are expressed as a multiple of nominal granularities. Each request must be single path routed. Grooming several requests on the same wavelength and multiplexing wavelengths in the same optical stream allow the packing of more traffic. However, each addition or withdrawal of a request from a wavelength requires optical to electrical conversion for which a specific portal equipment is needed. The objective is to minimize the number of such equipment. We deal with backbone optical networks, therefore with networks with a moderate number of nodes (14 to 20) but thousands of requests. Further difficulties arise from the symmetries in wavelength assignment and traffic loading. Traditional multi-commodity network flow approaches are not suited for this problem. Four alternative models relying on Dantzig-Wolfe and/or Benders’ decomposition are introduced and compared. The formulations are strengthened using symmetry breaking restrictions, variable domain reduction, zero-one decomposition of integer variables, and cutting planes. The resulting dual bounds are compared to the values of primal solutions obtained through hierarchical optimization and rounding procedures. For realistic size instances, our best approaches provide solutions with optimality gap of approximately 5% on average in around 2 hours of computing time.

6.1.4. Branching in Branch-and-Price: a generic scheme

Our innovative branching scheme, proposed for its compatible with the column generation procedure (it implies no structural modifications to the pricing problem) is now published in Mathematical Programming A [23]. The scheme proceeds by recursively partitioning the sub-problem solution set. Branching constraints are enforced in the pricing problem instead of being dualized in a Lagrangian way. The subproblem problem is solved by a limited number of calls to the provided solver. The scheme avoids the enumeration of symmetric solutions.
6.1.5. **Strong Branching Inequalities for Convex Mixed Integer Nonlinear Programs**

Strong branching is an effective branching technique that can significantly reduce the size of the branch-and-bound tree for solving Mixed Integer Nonlinear Programming (MINLP) problems. The focus of our paper [24] is to demonstrate how to effectively use discarded information from strong branching to strengthen relaxations of MINLP problems. Valid inequalities such as branching-based linearizations, various forms of disjunctive inequalities, and mixing-type inequalities are all discussed. The inequalities span a spectrum from those that require almost no extra effort to compute to those that require the solution of an additional linear program. In the end, we perform an extensive computational study to measure the impact of each of our proposed techniques. Computational results reveal that existing algorithms can be significantly improved by leveraging the information generated as a byproduct of strong branching in the form of valid inequalities.

6.1.6. **Linear and Nonlinear Inequalities for a Nonseparable Quadratic Set**

We described some integer-programming based approaches for finding strong inequalities for the convex hull of a quadratic mixed integer nonlinear set containing two integer variables that are linked by linear constraints. This study [31] was motivated by the fact that such sets appear can be defined by a convex quadratic program, and therefore strong inequalities for this set may help to strengthen the formulation of the original problem. Some of the inequalities we define for this set are linear, while others are nonlinear (specifically conic). The techniques used to define strong inequalities include not only ideas related to recent perspective reformulations of MINLPs, but also disjunctive and lifting arguments. Initial computational tests will be presented.

6.1.7. **On the composition of convex envelopes for quadrilinear terms**

Within the framework of the spatial Branch-and-Bound algorithm for solving Mixed-Integer Nonlinear Programs, different convex relaxations can be obtained for multilinear terms by applying associativity in different ways. The two groupings ((x1x2)x3)x4 and (x1x2x3)x4 of a quadrilinear term, for example, give rise to two different convex relaxations. In previous work, we proved that having fewer groupings of longer terms yields tighter convex relaxations. In this paper [35], we give an alternative proof of the same fact and perform a computational study to assess the impact of the tightened convex relaxation in a spatial Branch-and-Bound setting.

6.1.8. **Stable sets in claw-free graphs**

A **stable set** is a set of pairwise non adjacent vertices in a graph and a graph is **claw-free** when no vertex contains a stable set of size three in its neighborhood. Given weights on the vertices, the stable set problem (a NP-hard problem in general) consists in selecting a set of pairwise non adjacent vertices maximizing the sum of the selected weights. The stable set problem in claw-free graphs is a fundamental generalization of the classic matching problem that was shown to be polynomial by Minty in 1980 (G. Minty. On maximal independent sets of vertices in claw-free graphs. J. Combinatorial Theory B, 28:284-304 (1980)). However, in contrast with matching, the polyhedral structure (i.e. the integer hull of all stable sets in a claw-free graph) is not very well understood and thus providing a ‘decent’ linear description of this polytope has thus been a major open problem in our field.

We proposed a new algorithm to find a maximum weighted stable set in a claw-free graph [45] whose complexity is now drastically better than the original algorithm by Minty (n^3 versus n^6, where n is the number of vertices). We also provided a description of the polyhedra in an extended space (i.e. using additional artificial variables) and an **efficient procedure** to separate over the polytope in polynomial-time [27]. Beside those main contributions, we published another papers on the strongly minimal facets of the polytope [22].

We also published two survey papers on both the algorithmic and polyhedral aspects of the problem [32], [16].
6.1.9. Chvátal-Gomory rank of 0/1 polytopes

In [17], we study the Chvátal-Gomory rank of 0/1 polytopes. The Chvátal-Gomory procedure is a generic cutting plane procedure to derive the integer hull of polyhedra, and the rank is the number of iterations needed. We revisited a classic framework by Chvátal, Cook and Hartmann (V. Chvátal, W. Cook, and M. Hartmann. *On cutting-plane proofs in combinatorial optimization*. Linear Algebra and its Applications, 114/115:455-499 (1989)) to prove lower bounds on the CG-rank and we made it more accessible (the original framework was hard to apply). It allowed us to give a very simple construction and to improve the lower bound on the rank of general 0/1 polytopes (the previous weaker lower bound relied on a sophisticated existence theorem by Erdős). This result is important as it shed some new light on a supposedly well understood procedure.

6.1.10. The Circular-Chromatic number

Another central contribution of our team concerns the chromatic number of a graph (the minimum number of independent stable sets needed to cover the graph). We proved that the chromatic number and the clique number of some superclasses of perfect graphs is computable in polynomial time [19], [18]. We investigated the circular-chromatic number. It is a well-studied refinement of the chromatic number of a graph (designed for problems with periodic solutions): the chromatic number of a graph is the integer ceiling of its circular-chromatic number. Xuding Zhu noticed in 2000 that circular cliques are the relevant circular counterpart of cliques, with respect to the circular chromatic number, thereby introducing circular-perfect graphs, a super-class of perfect graphs.

We proved that the chromatic number of circular-perfect graphs is computable in polynomial time [73], thereby extending Grötschel, Lovász and Schrijver’s result to the whole family of circular-perfect graphs. We gave closed formulas for the Lovász Theta number of circular-cliques (previously, closed formulas were known for circular-cliques with clique number at most 3 only), and derived from them that the circular-chromatic number of circular-perfect graphs is computable in polynomial time [34].

6.2. Model Specific Developments and Applications

Participants: Cédric Joncour, Andrew Miller, Arnaud Pêcher, Pierre Pesneau, Ruslan Sadykov, Gautier Stauffer, Damien Trut, François Vanderbeck.

The models on which we made progress can be partitioned in three areas: “Packing and Covering Problems”, “Network Design and Routing”, and “Planning, Scheduling, and Logistic Problems”.

6.2.1. Bin-Packing and Knapsack with Conflicts

The bin-packing problem consists in finding the minimum number of bin of fixed size one needs to pack a set of items of different sizes. We studied a generalization of this problem where items can be in conflicts and thus cannot be put together in the same bin. We show in [21] that the instances of the literature with 120 to 1000 items can be solved to optimality with a generic Branch-and-Price algorithm, such as our prototype BaPCod, within competitive computing time. Moreover, we solved to optimality all the 37 open instances. The approach involves generic primal heuristics, generic branching, but a specific pricing procedure.

The knapsack variant encountered in our bin packing problem resolution considers conflicts between items. This problem is quite difficult to solve compared to the usual knapsack problem. The latter is already NP-hard, but can be usually efficiently solved by dynamic programming. We have shown that when the conflict graph (the graph defining the conflicts between the items) is an interval graph, this generalization of the knapsack can also be solved quite efficiently by dynamic programming with the same complexity than the one to solve the common knapsack problem. For the case, when the conflict graph is arbitrary, we proposed a very efficient enumeration algorithm which outperforms the approaches used in the literature.
6.2.2. Using graph theory for solving orthogonal knapsack problems

We investigated the orthogonal knapsack problem, with the help of graph theory. The multi-dimensional orthogonal packing problem (OPP) is defined as follows: given a set of items with rectangular shapes, the problem is to decide whether there is a non-overlapping packing of these items in a rectangular bin. The rotation of items is not allowed. A powerful characterization of packing configurations by means of interval graphs was introduced by Fekete and Schepers using an efficient representation of all geometrically symmetric solutions by a so called packing class involving one interval graph (whose complement admits a transitive orientation: each such orientation of the edges corresponds to a specific placement of the forms) for each dimension. Though Fekete & Schepers’ framework is very efficient, we have however identified several weaknesses in their algorithms: the most obvious one is that they do not take advantage of the different possibilities to represent interval graphs.

In [13], [14], [57], we give two new algorithms: the first one is based upon matrices with consecutive ones on each row as data structures and the second one uses so-called MPQ-trees, which were introduced by Korte and Möhring to recognize interval graphs. These two new algorithms are very efficient, as they outperform Fekete and Schepers’ on most standard benchmarks.

6.2.3. Inventory routing and pickup-and-delivery problems

Inventory routing problems combine the optimization of product deliveries (or pickups) with inventory control at customer sites. In [15], we considered the planning of single product pickups over time: each site accumulates stock at a deterministic rate; the stock is emptied on each visit. Our objective is to minimize a surrogate measure of routing cost while achieving some form of regional clustering by partitioning the sites between the vehicles. The fleet size is given but can potentially be reduced. Planning consists in assigning customers to vehicles in each time period, but the routing, i.e., the actual sequence in which vehicles visit customers, is considered as an “operational” decision. We developed a truncated branch-and-price algorithm. This exact optimization approach is combined with rounding and local search heuristics to yield both primal solutions and dual bounds that allow us to estimate the deviation from optimality of our solution. We were confronted with the issue of symmetry in time that naturally arises in building a cyclic schedule (cyclic permutations along the time axis define alternative solutions). Central to our approach is a state-space relaxation idea that allows us to avoid this drawback: the symmetry in time is eliminated by modelling an average behavior. Our algorithm provides solutions with reasonable deviation from optimality for large scale problems (260 customer sites, 60 time periods, 10 vehicles) coming from industry. The subproblem is interesting in its own right: it is a multiple-class integer knapsack problem with setups. Items are partitioned into classes whose use implies a setup cost and associated capacity consumption.

Through the internship of Damien Trut, we studied the optimization problem consisted in the planning of the pick-up of full waste container and delivery of empty container at customer sites by simple vehicles that can carry a single container, or vehicles with a trailer attached that have a total capacity of 2 containers but require more time when handling containers. The model is a multi-period, multi-vehicle, pickup and delivery problem, with “many-to-many” multi commodity transfer requirements and transhipment nodes. In its short term variant, urgent order are coming online. We developed the prototype of a branch-and-price approach for this problem. The prototype was used by Exeo to convince their customer of the potential benefit of decision aid tools to automatically generate vehicle routes. Next, we shall be considering the dimensioning of a vehicle fleet and their allocation to cluster of collect points in a periodic solution (for a PhD project).

In collaboration with the group of M-C Speranza of the university of Brescia (Italy), we study the Vehicle Routing Problem with Discrete Split Deliveries (a customer demand can be partition in integer lot assigned to different vehicles). The development is done within BaPCod with specialized pricing and branching scheme.

6.2.4. Time-Dependent Travelling Salesman Problem and Resource Constrained Shortest Path

In [12] we present a new formulation for the Time-Dependent Travelling Salesman Problem (TDTSP). The main feature of our formulation is that it uses, as a subproblem, an exact description of the n-circuit problem. We present a new extended formulation that is based on using, for each node, a stronger subproblem, namely a n-circuit subproblem with the additional constraint that the corresponding node is not repeated in the circuit.
Although the new model has more variables and constraints than the model of Picard and Queyranne (1978), the results given from our computational experiments show that the linear programming relaxation of the new model gives, for many of the instances tested, gaps that are close to zero. We also provided a complete characterization of the feasible set of the corresponding linear programming relaxation in the space of the variables of the PQ model.

Following this work, we proposed an extended formulation in terms of the Asymmetric Travelling Salesman Problem (ATSP) in [33]. A tightening the linear programming relaxation is obtained by i) enhancing the subproblem arising in the standard multicommodity flow (MCF) model for the ATSP and then ii) by using modelling enhancement techniques. We compare the linear programming relaxation of the new formulation with the linear programming relaxation of the three compact and non-dominated formulations presented in Oncan et al. (2009). As a result of this comparison we present an updated classification of formulations for the asymmetric traveling salesman problem (ATSP).

In the internship of André Linhares, we studied the Resource Constrained Shortest Path Problem (RCSPP): we presented some of the state-of-the-art dynamic programming methods for solving the RCSPP in a unified manner, and we proposed some variants of these algorithms. We assessed the effectiveness of these algorithms through computational experiments.

### 6.2.5. Machine scheduling

The column-and-row generation method presented in [28], [25] is quite effective for the general machine scheduling problem. In our work [29], we show indeed that one of the most efficient approaches to solve this problem is to use time-indexed Integer Programming formulation, and to deal with its huge size by generating variables and constraints dynamically. The numerical results of [29], highlight the significant reduction in computing times that results from applying the column-and-row generation approach.

In [20], we consider the scheduling jobs in parallel, i.e., jobs can be executed on more than one processor at the same time. With the emergence of new production, communication and parallel computing system, the usual scheduling requirement that a job is executed only on one processor has become, in many cases, obsolete and unfounded. In this work, we consider the NP-hard problem of scheduling malleable jobs to minimize the total weighted completion time (or mean weighted flow time). For this problem, we introduce the class of “ascending” schedules in which, for each job, the number of machines assigned to it cannot decrease over time while this job is being processed. We prove that, under a natural assumption on the processing time functions of jobs, the set of ascending schedules is dominant for the problem. This result can be used to reduce the search space while looking for an optimal solution.

### 6.2.6. One warehouse multi-retailer problem

The One-Warehouse Multi-retailer problem (OWMR) is a very important NP-hard inventory control problem arising in the distribution of goods when one central warehouse is supplying a set of final retailers facing demand from customers. In [30], we provide a simple and fast 2-approximation algorithm for this problem (i.e. an algorithm ensuring a deviation by a factor at most two from the optimal solution). This result is both important in practice and in theory as it allows to approximate large real-world instances of the problem (we implemented this algorithm at IBM and it is within 10% of optimality in practice) and the techniques we developed appear to apply to more general settings. We are extending our results to other inventory control problems.

### 6.3. Software prototypes, Generic Developments and Specific Tools

**Participants:** Cédric Joncour, Romain LeGuay, Pierre Pesneau, Ruslan Sadykov, François Vanderbeck.

#### 6.3.1. BaPCod - a generic branch-and-price code

The development of the prototype software platform is now supported by our junior engineer, Romanin Leguay, who started in September. He developed a new interface with the underlying MIP solver allowing multiple solvers to be called in the same run. The svn depository was re-organized in view of the increasing number of users to whom Romain offer precious support. Romain is currently redesigning parts of the code in the perspective of its parallelisation and is doing code profiling to identify bottlenecks.
The software platform BaPCod is continuously improved to include all the methodological features that arise from our research, in particular in our collaborative project with Brazil: SAMBA. BaPCod serves there as a proof-of-concept code and is useful for the transfer of knowledge between the parties, including the company GAPSO (a Brazilian spin-up launched by these academics). We have two new institutional Beta users: EDF and Paris 6.
6. New Results

6.1. White Noise-based Stochastic Calculus with respect to Multifractional Brownian Motion

Participants: Joachim Lebovits, Jacques Lévy Véhel.

The purpose of this work is to build a stochastic calculus with respect to (mBm) with a view to applications in finance and particularly to stochastic volatility models. We use an approach based on white noise theory.

6.1.1. White Noise-based Stochastic Calculus with respect to multifractional Brownian motion

The following results may be found in [28]. Integration with respect to mBm requires stochastic spaces, noted \( S^k(R), B(S^k(R)), \mu \) where \( \mu \) is probability measure given by Bôcher Minlos theorem, one can build to spaces, noted \( (S) \) and \( (S^*) \) which will play an analogous role to the spaces \( S(R) \) and \( S'(R) \) for tempered distributions. We recall that \( S(R) \) is the Schwartz space of rapidly decreasing functions which are infinitely differentiable and \( S'(R) \) is the space of tempered distributions. Let us moreover note \( (L^2) \) the space of random variables defined on the probability space \( (S(R), B(S(R)), \mu) \) which admit a second order moment. The mBm \( B^{(h)} \) has the following Wiener-Itô chaos decomposition in \( (L^2) \):

\[
B^{(h)}(t) = \sum_{k=0}^{+\infty} \int_{[0,t]} M_{h(t)}(e_k) >_{L^2(R)} \ldots M_{h(t)}(d_h^{(k)}) >_{e_k} = \sum_{k=0}^{+\infty} \left( \int_0^t M_{h(t)}(e_k)(s) ds \right) <., e_k > \quad (8)
\]

where \( (e_k)_{k \in \mathbb{N}} \) denotes the family of Hermite functions, defined for every integer \( k \) in \( \mathbb{N} \), by \( e_k(x) := \pi^{-1/4}(2^k k!)^{-1/2} e^{-x^2/2} h_k(x) \) and where \( (h_k)_{k \in \mathbb{N}} \) is the family of Hermite polynomial, defined for every integer \( k \) in \( \mathbb{N} \), by \( h_k(x) := (-1)^k e^{x^2} \frac{d^k}{dx^k}(e^{-x^2}) \). Note moreover that \( M_H \) is an operator from \( S(R) \) to \( L^2(R) \) for every real \( H \) in \( (0,1) \) and \( <., e_k > \) is a centered random Gaussian variable with variance equal to 1 for all \( k \) in \( \mathbb{N} \). We can now define a process, noted \( W^{(h)} \), from \( R \) to \( (S^*) \), which is the derivative of \( B^{(h)} \) in sense of \( (S^*) \) by

\[
W^{(h)}(t) = \sum_{k=0}^{+\infty} \left[ \frac{d}{dt} \left( \int_0^t M_{h(t)}(e_k)(s) ds \right) \right] <., e_k > . \quad (9)
\]

Hence we define integral with respect to mBm of any process \( \Phi : R \to (S^*) \) as being the element of \( (S^*) \) given by:

\[
\int_R \Phi(s, \omega) dB^{(h)}(s) = \int_R \Phi(s) \circ W^{(h)}(s) ds (\omega), \quad (10)
\]

where \( \circ \) denotes the Wick product on \( (S^*) \). It is then possible to get Itô-\( \tilde{A} \) formulas and Tanaka formula such as
\[
\int_0^T \frac{\partial f}{\partial x}(t, B^{(h)}(t)) \, dB^{(h)}(t) = f(T, B^{(h)}(T)) - f(0, 0) - \int_0^T \frac{\partial f}{\partial t}(t, B^{(h)}(t)) \, dt - \frac{1}{2} \int_0^T \left( \frac{d}{dt} [R_h(t, t)] \right) \frac{\partial^2 f}{\partial x^2}(t, B^{(h)}(t)) \, dt.
\]

(11)

for functions with sub exponential growth and where the last equality holds in \( L^2 \).

Once this stochastic calculus with respect to \( mBm \) is defined, we can solve differential equations arising in mathematical finance.

6.1.2. Multifractional stochastic volatility

Multifractional stochastic volatility

The results of this part may be found in [6]. We assume that, under the risk-neutral measure, the forward price of a risky asset is the solution of the S.D.E.

\[
\begin{align*}
\frac{dF_t}{F_t} & = \sigma_t dW_t, \\
\frac{d\ln (\sigma_t)}{\sigma_t} & = \theta (\mu - \ln (\sigma_t)) \, dt + \gamma h \, dB^{h}_t + \gamma \sigma dW^{\sigma}_t, \quad \sigma_0 > 0, \quad \theta > 0,
\end{align*}
\]

(12)

where \( W \) and \( W^{\sigma} \) are two standard Brownian motions and \( B^{h} \) is a multifractional Brownian motion independent of \( W \) and \( W^{\sigma} \) with functional parameter \( h \), which is assumed to be continuously differentiable.

We assume that \( W \) is decomposed into \( \rho dW^{\sigma}_t + \sqrt{1 - \rho^2} dW^{F}_t \), where \( W^{F} \) is a Brownian motion independent of \( W^{\sigma} \). Note that \( dB^{h}_t \) denotes differentiation in the sense of white Noise theory. The solution of the volatility process \( \sigma_t \) is

\[
\sigma_t = \exp \left( \ln (\sigma_0) e^{-\theta t} + \mu (1 - e^{-\theta t}) + \gamma \int_0^t e^{\theta(s-t)} dW^{\sigma}_s + \gamma h \int_0^t e^{-\theta t} I_t (B^{h}) \right),
\]

(13)

where \( I_t (B^{h}) := e^{\theta t} B^{h}_t - \theta \int_0^t e^{\theta s} B^{h}_s \, ds \).

Since the solution the previous S.D.E. is not explicit for \( (F_t)_{t \in [0,T]} \) we use preconditioning and then cubature methods in order to get an approximation of it. This model allows to take into account the well-known "smile" effect of volatility, as well as its evolution at various maturities.

6.1.3. Approximation of \( mBm \) by \( fBms \)

In [18], we establish that a sequence of well-chosen lumped fractional Brownian motions converges in law to a multifractional Brownian motion. This allows to define stochastic integrals with respect to \( mBm \) by "transporting" corresponding stochastic integrals with respect to \( fBm \).

6.2. Sample paths properties of the set-indexed Lévy process

Participant: Erick Herbin.

In collaboration with Prof. Ely Merzbach (Bar Ilan University, Israel).

In [24], the class of set-indexed Lévy processes is considered using the stationarity property defined for the set-indexed fractional Brownian motion in [23]. Following Ivanoff-Merzbach’s definitions of an indexing collection \( \mathcal{A} \) and its extensions \( \mathcal{C}_0 = \{ U \setminus V ; U, V \in \mathcal{A} \} \) and

\[
\mathcal{C} = \left\{ U \setminus \bigcup_{1 \leq i \leq n} V_i ; n \in \mathbb{N}; U, V_1, \ldots, V_n \in \mathcal{A} \right\},
\]


A set-indexed process $X = \{X_U; \ U \in \mathcal{A}\}$ is called a set-indexed Lévy process if the following conditions hold.

1. $X_{\emptyset'} = 0$ almost surely, where $\emptyset' = \bigcap_{U \in \mathcal{A}} U$.
2. The increments of $X$ are independent: for all pairwise disjoint $C_1, \ldots, C_n$ in $\mathcal{A}$, the random variables $\Delta X_{C_1}, \ldots, \Delta X_{C_n}$ are independent.
3. $X$ has $m$-stationary $C_0$-increments, i.e. for all integer $n$, all $V \in \mathcal{A}$ and for all increasing sequences $(U_i)_i$ and $(A_i)_i$ in $\mathcal{A}$, we have
   \[
   \forall i, m(U_i \setminus V) = m(A_i) \Rightarrow (\Delta X_{U_1 \setminus V}, \ldots, \Delta X_{U_n \setminus V}) \overset{(d)}{=} (\Delta X_{A_1}, \ldots, \Delta X_{A_n})
   \]
4. $X$ is continuous in probability.

On the contrary to previous works of Adler and Feigin (1984) on one hand, and Bass and Pyke (1984) on the other hand, the increment stationarity property allows to obtain explicit expressions for the finite-dimensional distributions of a set-indexed Lévy process. From these, we obtained a complete characterization in terms of Markov properties.

The question of continuity is more complex in the set-indexed setting than for real-parameter stochastic processes. For instance, the set-indexed Brownian motion can be not continuous for some indexing collection.

We consider a weaker form of continuity, which studies the possibility of point jumps.

The \textit{point mass jump} of a set-indexed function $x : \mathcal{A} \to \mathbb{R}$ at $t \in \mathcal{T}$ is defined by
\[
J_t(x) = \lim_{n \to \infty} \Delta x_{C_n(t)}, \quad \text{where} \quad C_n(t) = \bigcap_{C \in \mathcal{C}_n} C
\]
and for each $n \geq 1$, $\mathcal{C}_n$ denotes the collection of subsets $U \setminus V$ with $U \in \mathcal{A}_n$ (a finite sub-semilattice which generates $\mathcal{A}$ as $n \to \infty$) and $V \in \mathcal{A}_n(u)$. A set-indexed function $x : \mathcal{A} \to \mathbb{R}$ is said pointwise-continuous if $J_t(x) = 0$, for all $t \in \mathcal{T}$.

\textbf{Theorem} Let $\{X_U; \ U \in \mathcal{A}\}$ be a set-indexed Lévy process with Gaussian increments. Then for any $U_{\text{max}} \in \mathcal{A}$ such that $m(U_{\text{max}}) < +\infty$, the sample paths of $X$ are almost surely pointwise-continuous inside $U_{\text{max}}$, i.e.

\[
P(\forall t \in U_{\text{max}}, J_t(X) = 0) = 1.
\]

In the general case, for all $\epsilon > 0$, For all $U \in \mathcal{A}$ with $U \subset U_{\text{max}}$, we define
\[
N_U(B) = \# \{ t \in U : J_t(X) \in B \},
X_U^B = \int_B x.N_U(dx),
\]
for all $B \in \mathcal{B}_\epsilon$, the $\sigma$-field generated by the opened subsets of $\{x \in \mathbb{R} : |x| > \epsilon\}$. The sample paths of the set-indexed Lévy processes can be derived from the following Lévy-Ito decomposition proved in [24].

\textbf{Theorem} Let $(\sigma, \gamma, \nu)$ the generating triplet of the SI Lévy process $X$.

Then $X$ can be decomposed as

\[
\forall \omega \in \Omega, \forall U \in \mathcal{A}, \quad X_U(\omega) = X_U^{(0)}(\omega) + X_U^{(1)}(\omega),
\]
where
1. \( X^{(0)} = \{ X^{(0)}_U; U \in \mathcal{A} \} \) is a set-indexed Lévy process with Gaussian increments, with generating triplet \((\sigma, \gamma, 0)\),
2. \( X^{(1)} = \{ X^{(1)}_U; U \in \mathcal{A} \} \) is the set-indexed Lévy process with generating triplet \((0, 0, \sigma)\), defined for some \( \Omega_1 \in \mathcal{F} \) with \( P(\Omega_1) = 1 \) by

\[
\forall \omega \in \Omega_1, \quad \forall U \in \mathcal{A}, \quad X^{(1)}_U(\omega) = \int_{|x| > 1} x N_U(dx, \omega) + \lim_{\epsilon \downarrow 0} \int_{\epsilon < |x| \leq 1} x \left[ N_U(dx, \omega) - m(U) \right] \nu(dx), \tag{16}
\]

where \( N_U \) is defined in (13) and the last term of (14) converges uniformly in \( U \subset U_{\text{max}} \) (for any given \( U_{\text{max}} \in \mathcal{A} \)) as \( \epsilon \downarrow 0 \),
3. and the processes \( X^{(0)} \) and \( X^{(1)} \) are independent.

### 6.3. Hölder regularity of Set-Indexed processes

**Participants:** Erick Herbin, Alexandre Richard.

In collaboration with Prof. Ely Merzbach (Bar Ilan University, Israel).

In the set-indexed framework of Ivanoff and Merzbach ([54](#)), stochastic processes can be indexed not only by \( \mathbb{R} \) but by a collection \( \mathcal{A} \) of subsets of a measure and metric space \((\mathcal{T}, d, m)\), with some assumptions on \( \mathcal{A} \). In [25](#), we introduce and study some assumptions on the metric indexing collection \((\mathcal{A}, d_A)\) in order to obtain a Kolmogorov criterion for continuous modifications of SI stochastic processes. Under this assumption, the collection is totally bounded and a set-indexed process with good incremental moments will have a modification whose sample paths are almost surely Hölder continuous, for the distance \( d_A \).

Once this condition is established, we investigate the definition of Hölder coefficients for SI processes. From the real-parameter case, the most straightforward are the local (and pointwise) Hölder exponents around \( U_0 \in \mathcal{A} \): \[ \tilde{\alpha} X(U_0) = \sup \left\{ \alpha : \limsup_{\rho \to 0} \sup_{U, V \in B_{\rho}(U_0, \mathcal{A})} \frac{|X_U - X_V|}{d_{\mathcal{A}}(U, V)}^\alpha < \infty \right\}. \]

When the processes are Gaussian, a deterministic counterpart to this exponent is defined as it is in the real-parameter framework. For all \( U_0 \in \mathcal{A} \), we proved that almost surely, the random and the deterministic exponents are equal. Also, we proved that for the local exponents, this result holds almost surely, uniformly on \( \mathcal{A} \).

Given the particular structure of \( \mathcal{A} \), other coefficients of Hölder regularity were studied on \( \mathcal{C} \):

\[
\mathcal{C} = \left\{ A \setminus \bigcup_{k=1}^n B_k : A, B_1, \ldots, B_n \in \mathcal{A}, n \in \mathbb{N} \right\}. \tag{17}
\]

On specific subclasses \( \mathcal{C}^l \) of \( \mathcal{C} \) (satisfying \( \bigcup_{l \geq 1} \mathcal{C}^l = \mathcal{C} \)), the local (and pointwise) \( \mathcal{C}^l \)-Hölder exponents are defined:
\[ \tilde{\alpha}_{X, \mathcal{L}}(U_0) = \sup \left\{ \alpha : \limsup_{\rho \to 0} \sup_{U \in B_d(U_0, \rho)} \sup_{V \in B_l(U_0, \rho)} \frac{|\Delta X_{U \setminus V}|}{d_A(U, V)^\alpha} < \infty \right\}, \quad (18) \]

and this definition is proved to be independent of \( l \), leading to the definition of \( \tilde{\alpha}_{X, \mathcal{L}}(U_0) \). It is compared to \( \tilde{\alpha}_{X}(U_0) \) and related to the Hölder exponent of the process projected on flows (a flow is a continuous increasing path in \( \mathcal{A} \)). This last technique permits to show that the pointwise Hölder exponent of the SIfBm is almost surely uniformly equal to \( H \), the Hurst parameter of the SIfBm. This completes some previous results on the multiparameter fractional Brownian motion.

The last exponent which is studied is the exponent of pointwise continuity:

\[ \alpha_{pc,l}(t) = \sup \left\{ \alpha : \limsup_{n \to \infty} \frac{|\Delta X_{C_n(t)}|}{m(C_n(t))^{\sigma}} < \infty \right\} \quad (19) \]

for all \( t \in \mathcal{T} \), where \( C_n(t) \) is the smaller set of \( C_n \) containing \( t \). Almost sure results are also obtained in that case. For instance, the coefficient of pointwise continuity of a SI Brownian motion equals \( 1/2 \) a.s.

All these results are finally applied to the SIfBm and the SI Ornstein-Uhlenbeck process ([1]).

### 6.4. Stochastic 2-microlocal analysis

**Participants:** Erick Herbin, Paul Balança.

Stochastic 2-microlocal analysis has been introduced in [19] to study the local regularity of stochastic processes. If \( X = (X_t)_{t \in \mathbb{R}_+} \) is a stochastic process, then for all \( t_0 \in \mathbb{R}_+ \), a function \( s' \mapsto \sigma_{X,t_0}(s') \) called the 2-microlocal frontier is defined to characterize entirely the local regularity of \( X \) at \( t_0 \). In particular, for all \( s' \in \mathbb{R} \) such that \( \sigma_{X,t_0}(s') \in (0,1) \), it is defined as

\[ \sigma_{X,t_0}(s') = \sup \left\{ \sigma : \limsup_{\rho \to 0} \sup_{u,v \in B[t_0, \rho]} \frac{|X_u - X_v|}{|u - v|^{\sigma}} < \infty \right\}. \]

The 2-microlocal frontier gives a more complete picture of the regularity than classical pointwise and local Hölder exponents, which are widely used in the literature. Furthermore, it is stable under the action of (pseudo-)differential operators.

[19] mainly focused on Gaussian processes, and in particular obtained a characterization of the regularity for Wiener integrals \( X_t = \int_0^t \eta_u dW_u \), with \( \eta \in L^2(\mathbb{R}) \).

Our main goal was therefore to extend this result to any stochastic integral

\[ X_t = \int_0^t H_u dM_u, \]

where \( M \) is a local martingale and \( H \) an adapted continuous process.

In fact, in [15], we first reduced this problem to the study of local martingales, and we have shown that almost surely for all \( t \in \mathbb{R}_+ \), the 2-microlocal frontier of a local martingale \( M \), with quadratic variation \( \langle M \rangle \), satisfies
∀s′ ≥ −α; \[\sigma_M(t)(s') = \sum_{i} (s') = \frac{1}{2} \sum_{(M),t}(2s'),\]

where for any process X, \(\Sigma_{X,t}\) denotes the pseudo 2-microlocal frontier which is characterized as following

∀s′ ∈ R; \[\Sigma_{X,t}(s') = \sigma_{X,t}(s') \land (s' + p_{X,t}) \land 1,\]

where \(p_{X,t}\) corresponds to

\[p_{X,t} = \inf \left\{ n \geq 1 : X^{(n)}(t) \text{ exists and } X^{(n)}(t) \neq 0 \right\},\]

with the usual convention \(\inf\{\emptyset\} = +\infty.\)

As the previous result is based on Dubins-Schwarz representation theorem, it can be easily extended to characterize the regularity of time-changed multifractional Brownian motions. In this case, we obtain a similar equation where \(\frac{1}{2}\) is replaced by \(H(t)\), the value of the Hurst function at \(t.\)

Using this last equality, we can obtain the regularity of the stochastic integral \(X\) previously defined: almost surely for all \(t \in R_+\)

∀s′ ≥ −α; \[\sigma_{X,t}(s') = \sum_{H} (s') = \frac{1}{2} \sum_{(M),t} H^2 (2s').\]

In the particular case of an integration with respect to a Brownian motion \(B\), the result can be simplified using the stability under differential operators: for almost all \(\omega \in \Omega\) and for all \(t \in R_+\), the 2-microlocal frontier satisfies

1. if \(H(t) \neq 0\):

∀s′ ∈ R; \[\sigma_{X,t}(s') = \sigma_{B,t}(s') = \left(\frac{1}{2} + s'\right) \land \frac{1}{2};\]

2. if \(H(t) = 0\):

∀s′ ≥ −α; \[\sigma_{X,t}(s') = \left(\frac{1}{2} + \frac{\Sigma_{H^2,t}(2s')}{2}\right) \land \frac{1}{2},\]

unless \(H\) is locally equal to zero at \(t\), which induces in that case: \(\sigma_{X,t} = +\infty.\)

Based on this last characterization, we were able to study the regularity of stochastic diffusions. In particular, we illustrated our purpose with the square of \(\delta\)-dimensional Bessel processes which verify the following equation

\[Z_t = x + 2 \int_0^t \sqrt{Z_s} d\beta_s + \delta t.\]

6.5. Tempered multistable measures and processes

Participants: Jacques LÅ©vy VÅ©hel, Lining Liu.
This year, we concentrated on the following points:

- Define a new type of multistable processes called tempered multistable processes.
- Study the short time and long time behaviors of tempered multistable processes.
- Compare the multistable Lévy processes defined by finite-dimensional distributions (characteristic functions), Poisson representation and series representation.

The idea of the construction of tempered multistable measure and processes comes from the paper [63]. The interest of such processes is that they may be chosen to have moments of all orders. In addition, they are martingales. This will allow to construct stochastic (partial) differential equation driven by tempered multistable measures, which may be used to describe certain physical phenomena.

The characteristic function of a tempered multistable process \( X(t) \) is

\[
E(\exp iyX(t)) = \exp \left\{ \frac{1}{2} \int_0^t \Gamma(-\alpha(x)) \left[ \left( 1 - \frac{iy}{\theta} \right)^{\alpha(x)} + \left( 1 + \frac{iy}{\theta} \right)^{\alpha(x)} - 2 \right] \theta^{\alpha(x)} \, dx \right\} .
\]

We have investigated the long time and short time behaviors this process:

Short time behavior:
Let \( \alpha: \mathbb{R} \rightarrow [a, b] \subseteq (0, 2) \) be continuous. Let \( u \in \mathbb{R} \) and suppose that as \( v \rightarrow u, \)

\[
|\alpha(u) - \alpha(v)| = o \left( \frac{1}{\log |u - v|} \right) .
\]  

(20)

Then when \( h \rightarrow 0, \)

\[
h^{-1/\alpha(t)} [X(t + hu) - X(t)] \rightarrow Y_{\alpha(t)}(u) \]

(21)
in finite-dimensional-distributions, where

\[
Y_{\alpha(t)}(u) = \int 1_{[0,u]}(z) dM_{\alpha(t)}(z),
\]

and \( M_{\alpha(t)} \) is an \( \alpha(t) \) stable measure. In an other word, \( X(t) = M[0, t] \) is 1/\( \alpha(t) \)-localisable at \( t \) with local form \( Y_{\alpha(t)} \).

Long time behavior:
Let \( \alpha: \mathbb{R} \rightarrow [a, b] \subseteq (0, 2) \) be continuous and \( \lim_{s \rightarrow \infty} \alpha(s) \rightarrow \alpha. \) Then for \( h \rightarrow \infty \)

\[
h^{-1/2} [X(t + hu) - X(t)] \rightarrow \Gamma(2 - \alpha) B(u)
\]

(22)
in finite-dimensional-distributions, where \( B \) is standard Brownian motion.

Let us now describe our work on the multistable Lévy motion. For \( 0 < a \leq b < 2 \) and \( \alpha: \mathbb{R} \rightarrow [a, b], \) the multistable Lévy motion \( M_c \) defined by finite-dimensional distributions (characteristics function) is the process such that

\[
E(\exp \{i \sum_{j=1}^d \theta_j M_c(t_j)\}) = \exp \left( - \int \left| \sum_{j=1}^d \theta_j 1_{[0,t_j]}(s) \right|^{\alpha(s)} \, ds \right) ;
\]

(23)
There also exist a Poisson representation of multistable Lévy process $M_p$:

$$M_p(t) = \sum_{(X,Y) \in \Pi} C_{\alpha(X)} 1_{[0,t]}(X) Y^{< -1/\alpha(X)},$$  

(24)

where $(X,Y)$ be the random point of the Poisson process $\Pi$, $t > 0$, $Y^{< -1/\alpha(X)} = \text{sign}(Y)|Y|^{-1/\alpha(X)}$ and

$$C_{\alpha(X)} = \left( \frac{1}{\Gamma(1 - \alpha(X))} \cos \left( \frac{\pi}{2} \alpha(X) \right) \right)^{1/\alpha(X)};$$  

(25)

Finally, the series representation of multistable Lévy motion $M_s$ is

$$M_s(t) = \sum_{i=1}^{\infty} C_{\alpha(U_i)} \gamma_i^{1-1/\alpha(U_i)} 1_{(U_i \leq t)},$$  

(26)

where $\{\Gamma\}_{i \geq 1}$ is a sequence of arrival times of a Poisson process with unit arrival time, $\{U\}_{i \geq 1}$ is a sequence of i.i.d random variables with uniform distribution on $[0,t]$, $\{\gamma\}_{i \geq 1}$ is a sequence of i.i.d random variables with distribution $P(\gamma_i = 1) = P(\gamma_i = -1) = 1/2$. All three sequences $\{\Gamma\}_{i \geq 1}$, $\{U\}_{i \geq 1}$ and $\{\gamma\}_{i \geq 1}$ are independent, and

$$C_{\alpha(U_i)} = \left( \frac{1}{\Gamma(1 - \alpha(U_i))} \cos \left( \frac{\pi}{2} \alpha(U_i) \right) \right)^{1/\alpha(U_i)}.$$

(27)

We have proved that these three definitions yield the same process in law.

6.6. Local strings and the CH set

**Participant:** Jacques Lévy Véhel.

**In collaboration with Prof. Franklin Mendivil (Acadia University, Canada).**

We have extended the definition of fractal strings originally proposed in [59] and modified in [37] to deal with the local behaviour of fractal sets. This allows to analyze the pointwise oscillatory properties of locally self-similar sets ([38]).

We have also analyzed in details the structure of a set build by “stacking” Cantor sets with continuously varying dimensions (see figure 4). The resulting set, called “Christiane’s hair” set or CH set, displays a number of interesting properties. Each “strand of hair” is a $C^\infty$ curve. Its Hausdorff dimension is 2. Furthermore, it is Minkowski measurable in dimension 2 with vanishing Minkowski content.

6.7. General models for drug concentration in multi-dosing administration

**Participants:** Lisandro Fermin, Jacques Lévy Véhel.

**In collaboration with P.E Lévy Véhel (University of Nice-Sophia-Antipolis and Banque Postale).**
In the past two years, we have developed models for investigating the probability distribution of drug concentration in the case of non-compliance. We have focused on two aspects of practical relevance: the variability of the concentration and the regularity of its probability distribution. In a first article [29], in a series of three, is considered the case of multi-intravenous dosing using the simplest possible law to model random drug intake, i.e., a homogeneous Poisson distribution. In a second article [13], we consider the more realistic multi-oral model, and deal with the complications brought by the first-order kinetics, which are essentially technical. Finally, in [12], we put ourselves in a powerful mathematical frame, known as Piecewise Deterministic Markov process (PDMP), that allows us to deal with general drug intake schedules, going beyond the homogeneous Poisson case. We use a PDMP to model the drug concentration in the case of multiple intravenous doses. In this particular model, we consider that the doses administration regimen is modeled by a non-homogeneous Poisson process whose jump rate is controlled by mean of a Markov chain. In this sense our PDMP model is a generalization to the continuous-models studied in [29]. In the following we detail our PDMP model and the results obtained in the multi-IV case, see [12].

The model setting

Inspired by the PDMP model given in [47], [48], we consider a drug dosing stochastic regimen defined as follows.

Let us consider \((J_n)_{n\in\mathbb{N}}\) an irreducible Markov chain taking values in the state space \(K = \{1, ..., k\}\) with initial law \(\alpha_i = \mathbb{P}(J_0 = i)\) for all \(i \in K\) and transition probability matrix \(Q = (q_{ij})_{i,j\in K}\). We denote by \((T_n)_{n\in\mathbb{N}}\) the sequence of the random time doses and \((S_n)_{n\in\mathbb{N}}\) the time dose intervals; i.e. \(S_n = T_{n+1} - T_n\). We consider that the doses administration regimen is modeled by a non-homogeneous Poisson process whose jump rate is controlled by mean of a Markov chain. In this sense our PDMP model is a generalization to the continuous-models studied in [29]. In the following we detail our PDM model and the results obtained in the multi-IV case, see [12].

We consider that these doses translate into immediate increases of the concentration by the value \(d_i = D_i V_d\) if \(J_n = i\), where \(V_d\) is the apparent volume of distribution. After that, the effect of the dose taken at time \(T_n\) decreases exponentially fast with an exponential rate of elimination \(k_e\). We define \((\nu_t)_{t\in\mathbb{R}}\) by \(\nu_t = \sum_{n\geq 0} J_n \mathbb{1}_{[T_n, T_{n+1}[}(t)\). We denote by \((C_t)_{t\in\mathbb{R}}\) the drug concentration stochastic process which take values on \(\mathbb{R}_+^* = [0, \infty[\), we suppose that \(\mathbb{P}(C_0 = x) = 1\). Between the jumps, the dynamical evolution of the continuous time process \((C_t)\) is modeled by the flow \(\phi(t, x) = x \exp\{-k_e t\}.\)
Thus, the sample path of the stochastic process \( (C_t)_{t \in \mathbb{R}_+} \) with values in \( \mathbb{R}_+^* \), starting from a fixed point \( x \) is given by

\[
C_t = xe^{-k_x t} + \sum_{i \geq 1} d_i t e^{-k_i (t - T_i)} \mathbb{1}_{(t \geq T_i)}.
\]

(28)

The process \( (C_t, \nu_t)_{t \in \mathbb{R}_+} \) is a PDMP. From \([49]\), we have that the infinitesimal generator \( \mathbb{L} \) of \( (C_t, \nu_t)_{t \in \mathbb{R}_+} \) is given by

\[
\mathbb{L} f(x, i) = -k_x x \frac{d}{dx} f(x, i) + \lambda_i \sum_{j \in K} q_{ij} \left( f(x + d_j, j) - f(x, i) \right),
\]

(29)

with \( (x, i) \in E = \mathbb{R}_+^* \times K \) and \( f \in \mathbb{D}(\mathbb{L}) \) the set of measurable and differentiable on the first argument. 

The characteristic function of the concentration

The characteristic function \( \phi_\theta(t, x, i) \) of \( C_t \), given the starting point \( (x, i) \), is the unique solution of the following system

\[
\begin{align*}
\frac{\partial \phi_\theta}{\partial t} (t, x, i) &= -k_x x \frac{\partial \phi_\theta}{\partial x} (t, x, i) + \lambda_i \sum_{j \in K} q_{ij} \left( e^{\theta d_j} e^{-k_x t} \phi_\theta(t, x, j) - \phi_\theta(t, x, i) \right), \\
\phi_\theta(0, x, i) &= e^{i \theta x}.
\end{align*}
\]

(30)

Variability of the concentration

From (28) we have that the expectation \( m(t, x, i) = \mathbb{E}_{(x, i)} [C_t] \) of \( C_t \), given the starting point \( (x, i) \), is given by

\[
m(t, x, i) = xe^{-k_x t} + \sum_{j \in K} \lambda_j q_{ij} d_j \int_0^t e^{-k_j (t - s)} P_{i\nu}(s) ds,
\]

(31)

where \( P_{i\nu}(t) = \mathbb{P}(\nu_t = \nu | \nu_0 = i) \). The variance \( Var(t, i) \) of \( C_t \), given the initial state \( i \), is given by

\[
Var(t, i) = \sum_{\nu, j \in K} \lambda_\nu q_{ij} d_j^2 \int_0^t e^{-2k_j (t - s)} P_{i\nu}(s) ds - \left( \sum_{\nu, j \in K} \lambda_\nu q_{ij} d_j \int_0^t e^{-k_j (t - s)} P_{i\nu}(s) ds \right)^2
\]

\[
+ 2 \sum_{\nu, j \in K} \sum_{j' \in K} \lambda_\nu q_{ij} d_j \lambda_\nu q_{ij'} d_j' \int_0^t \int_0^{t - s} e^{-k_j (t - s)} P_{i\nu}(s) e^{-k_j (t - s - \tau)} P_{i\nu}(\tau) d\tau ds.
\]

(32)

The distribution of limit concentration

The characteristic function \( \varphi(\theta, i) \) of the limit concentration \( C \), given the starting state \( i \), satisfies

\[-k_x \theta \frac{d}{d\theta} \varphi(\theta, i) + \sum_{j \in K} \lambda_j q_{ij} e^{i \theta d_j} \varphi(\theta, j) - \lambda_i \varphi(\theta, i) = 0.\]

Thus, the random variables \( C(t) \) converge in distribution, when \( t \) tends to infinity, to a well defined random variable \( C \) whose characteristic function is
\[ \varphi(\theta) = \sum_{j \in K} \varphi(\theta, j). \]

Variability of the limit concentration

We denote by \( m_i \) the mean of the limit concentration \( C \) in the state \( \nu = i \) and \( m = \sum_{i \in K} m_i \) the mean of \( C \) and \( \text{Var} \) its variance. Then,

\[ m = \frac{1}{k_e} \sum_{i,j \in K} \pi_i \lambda_i q_{ij} d_j, \]

\[ m_i = \frac{1}{k_e} \sum_{j \in K} \pi_j \lambda_j q_{ji} d_i + \frac{1}{k_e} \left( \sum_{j \in K} \lambda_j q_{ji} m_j - \lambda_i m_i \right) \]

\[ \text{Var} = \frac{1}{2k_e} \sum_{i,j \in K} \pi_i \lambda_i q_{ij} d_j^2 + \frac{1}{k_e} \sum_{i,j \in K} \lambda_i q_{ij} d_j (m_i - \pi_i m). \]

Regularity of the limit concentration

The characteristic function \( \varphi \) satisfies

\[ |\varphi(\theta)| \sim K |\theta|^{-\mu_{\text{max}}}, \quad \theta \to \infty, \quad (33) \]

where \( K \) is a positive constant and \( \mu_{\text{max}} = \max_{i \in K} \frac{\lambda_i}{k_e}. \)

This result will allow us to describe in detail aspects of the limit distribution that are important for assessing the efficacy of therapy.

6.8. Complex systems design

Participant: Erick Herbin.

In collaboration with Dassault Aviation, EADS, EDF.

The preliminary design of complex systems can be described as an exploration process of a so-called design space, generated by the global parameters. An interactive exploration, with a decisional visualization goal, needs reduced-order models of the involved physical phenomena. We are convinced that the local regularity of phenomena is a relevant quantity to drive these approximated models. Roughly speaking, in order to be representative, a model needs more informations where the fluctuations are the more important (and consequently, where irregularity is the more important).

In collaboration with Dassault Aviation, EDF and EADS, we study how the local regularity can provide a good quantification of the concept of granularity of a model, in order to select the good level of fidelity adapted to a required precision.

Our works in that field can be expressed into:

- The definition and the study of stochastic partial differential equations driven by processes with prescribed regularity (that do not enter into the classical theory of stochastic integration).
- The study of the evolution of the local regularity inside stochastic partial differential equations (SPDE). Stochastic 2-microlocal analysis should provide informations about the local regularity of the solutions, in function of the coefficients of the equations. The knowledge of the fine behaviour of the solution of the SPDE will provide important informations in the view of numerical simulations.
SELECT Project-Team

6. New Results

6.1. Model selection in Regression and Classification

Participants: Gilles Celeux, Mohammed El Anbari, Clément Levrard, Robin Genuer, Erwan Le Pennec, Lucie Montuelle, Pascal Massart, Caroline Meynet, Jean-Michel Poggi.

Erwan Le Pennec continues his work with Serge Cohen (IPANEMA Soleil) on hyperspectral image segmentation based on a spatialized Gaussian Mixture Model. They derive, and implement within MIXMOD, an efficient minimization algorithm combining EM algorithm, dynamic programming and model selection [37]. They have applied this technique to analyze ancient materials [9]. This scheme is supported by a theoretical work on conditional density estimation [40]. In the framework of her PhD, Lucie Montuelle has studied some extension to this model to spatially varying logistic weights.

In collaboration with Marie-Laure Martin-Magniette (URGV et UMR AgroParisTech/INRA MIA 518) and Cathy Maugis (INSA Toulouse) has extended their variable selection procedure for model-based clustering and supervised classification to deal with high dimensional data sets with a backward selection procedure which is more efficient than the previous forward selection procedure in this context. [17]. Moreover they have shown the advantage of the model-based approach over a geometrical approach to select variable for clustering [13]. These variable selection procedures are in particular used for genomics applications which is the result of a collaboration with researchers of of URGV (Evry Genopole).

Caroline Meynet provided an $\ell_1$-oracle inequality satisfied by the Lasso estimator with the Kullback-Leibler loss in the framework of a finite mixture of Gaussian regressions model for high-dimensional heterogeneous data where the number of covariates may be much larger than the sample size. In particular, she has given a condition on the regularization parameter of the Lasso to obtain such an oracle inequality. This oracle inequality extends the $\ell_1$-oracle inequality established by Massart and Meynet [16] in the homogeneous Gaussian linear regression case. It is deduced from a finite mixture Gaussian regression model selection theorem for $\ell_1$-penalized maximum likelihood conditional density estimation, which is inspired from Vapnik’s method of structural risk minimization and from the theory on model selection for maximum likelihood estimators developed by Massart.

From an practical point of view, Caroline Meynet has introduced a procedure to select variables in model-based clustering in a high-dimensional context. In order to tackle with the problem of high-dimension, she has proposed to first use the Lasso in order to select different sets of variables and then estimate the density by a standard EM algorithm by reducing the inference to the linear space of the selected variables by the Lasso. Numerical experiments show that this method can outperform direct estimation by the Lasso.

In collaboration with Professor Abdallah Mkhadri (University of Marrakesh, Marocco), Gilles Celeux supervised the thesis of Mohammed El Anbari which concern regularisation methods in linear regression. In collaboration with Professor Abdallah Mkhadri (University of Marrakesh, Marocco), Mohammed El Anbari proposed a method to simultaneously select variables and favor a grouping effect where strongly correlated predictors tend to be in or out of the model together. Numerical experiments showed that their method can be preferred to Elastic-Net when the number of variables is less or equal to the sample size and remain competitive otherwise. Moreover, they have proposed AdaGril an extension of the the adaptive Elastic Net which incorporates information redundancy among correlated variables for model selection and estimation. Under weak conditions, They have established an oracle property of AdaGril. Numerical experiments show in some cases of AdaGril outperforms its competitors.

In collaboration with Jean-Michel Marin (Université de Montpellier) and Christian P. Robert (CEREMADE, Université Paris Dauphine) Gilles Celeux and Mohammed El Anbari highlight the interest of Bayesian regularization methods, using hierarchical non informative priors, compared with standard regularization methods in a poorly informative context through numerical experiments [47].
Clément Levrard worked on the obtention of fast rates of convergence for vector quantization. Using theoretical analogies between quantization seen as an unsupervised learning problem and the one of supervised learning by empirical contrast minimization, he has obtained a logarithmic improvement on the previously obtained bound. He has been furthermore able to define intelligible “margin type” condition under which fast rates can be obtained.

Since September 2008, Pascal Massart is the cosupervisor with Frédéric Chazal (GEOMETRICA) of the thesis of Claire Caillerie (GEOMETRICA). The project intends to explore and to develop new researches at the crossing of information geometry, computational geometry and statistics.

6.2. Statistical learning methodology and theory

**Participants:** Gilles Celeux, Christine Keribin, Erwan Le Pennec, Pascal Massart, Lucie Montuelle, Jean-Michel Poggi.

Unsupervised segmentation is an issue similar to unsupervised classification with an added spatial aspect. Functional data is acquired on points in a spatial domain and the goal is to segment the domain in homogeneous domain. The range of applications includes hyperspectral images in conservation sciences, fMRI data and all spatialized functional data. Erwan Le Pennec and Lucie Montuelle are focusing on the questions of the way to handle the spatial component from both the theoretical and the practical point of views as well as the choice of the number of clusters. Furthermore, as functional data require heavy computation, they are required to propose numerically efficient algorithms.

Gilles Celeux, Christine Keribin and the Ph.D. student Vincent Brault continue their work on the Latent Block Model. They have proposed an efficient algorithm coupling a Stochastic version of the EM algorithm including a Gibbs sampling step and the Variational EM algorithm. This SEM-VEM algorithm is insensitive to its initial position. On the other hand they got a closed formed expression of the Integrated Completed Likelihood for binary tables which allows for a reliable model selection criterion avoiding asymptotic approximation. 
Moreover, Christine Keribin derived sufficient conditions ensuring the identifiability of the Latent Block Model.

6.3. Reliability and Computer Experiments

**Participants:** Yves Auffray, Gilles Celeux, Rémi Fouchereau, Shuai Fu, Pascal Massart.

In the computer experiments field, the goal is to approximate an expensive black box function from a limited number of evaluations. The choice of these evaluations i.e. the choice of a design of (computer) experiments is a major issue.

This year Yves Auffray and Pierre Barbillon, in collaboration with Jean-Michel Marin (Université de Montpellier) have considered estimating the probability of rare events in the context of computer experiments. These rare events depends on the output of a physical model with random input variables. Since the model is only known through an expensive black box function, a crude Monte Carlo estimator does not perform well. Two strategies have been developed to cope with this difficulty: a Bayesian estimate and an importance sampling method. Both methods relies on Kriging metamodeling. They are able to achieve sharp upper confidence bounds on the rare event probability. These methods have been applied to a toy example and a real case study which consists of finding an upper bound of the probability that the trajectory of an airborne load collides the aircraft that has released it.

Following the previous work of the first year, Shuai Fu, under the direction of Gilles Celeux, focus on the design of experiments and its validation, which has become the main issues of the thesis. It leads both to theoretical and computational developments. An original DAC criterion has been proposed and leads to a Bayesian procedure of DAC-test to measure the quality of a design. For improving the design of experiments, an adaptive kriging procedure well adapted to the specific problem has been proposed. However, the algorithms require a too important computation time which should be reduced in future work.
In the framework of a CIFRE convention with Snecma-SAFRAN Rémy Fouchereau has started a thesis on
the modeling of fatigue damage for Inco718 supervised by Gilles Celeux. Inco718 is a Zinc-based alloy. To
determine its minimum lifetime, a lot of stress tests are made. The lloay lifetimes are reported as function of the
stress. The aim of this work is to analyse the resulting curves. A mixture model with a lognormal component
and a sum of two lognormals components is considered. Since the sum of two or more lognormal distribution
is not closed form. Inference on this model needs Monte Carlo integration within the EM algorithm. Despite
some unstability for small sample sizes, this model show encouraging and easily interpretable results.

6.4. Statistical analysis of genomic data

Participants: Gilles Celeux, Andrea Rau.

Andrea Rau and Gilles Celeux, in collaboration with Marie-Laure Martin-Magniette (URGV and UMR
AgroParisTech/INRA MIA 518) and Cathy Maugis-Rabusseau (IMT/INSA Toulouse) have developed a
method to cluster digital gene expression observations from high-throughput (HTS) data using Poisson mixture
models \[ 44 \]. The proposed model has the advantage of accounting for the particularities of HTS data and
providing straightforward procedures for parameter estimation and model selection. A series of simulation
experiments was done to compare the performance of the proposed model to that of previously proposed
clustering methods for similar sequence-based data, and the performance of the proposed approach was
examined on two real high-throughput sequencing data sets. The R package HTSCluster used to implement
the proposed Poisson mixture model has been made freely available on CRAN.

6.5. Curves classification, denoising and forecasting

Participant: Jean-Michel Poggi.

In collaboration with Farouk Mhamdi and Meriem Jaidane (ENIT, Tunis, Tunisia), Jean-Michel Poggi
proposed, in \[ 18 \], a method for trend extraction from seasonal time series through the Empirical Mode
Decomposition (EMD). Experimental comparison of trend extraction based on EMD, X11, X12 and Hodrick
Prescott filter are conducted. First results show the eligibility of the blind EMD trend extraction method.
Tunisian real peak load is also used to illustrate the extraction of the intrinsic trend.

In collaboration with Mina Aminghafari (Amirkabir University, Teheran), Jean-Michel Poggi made uses of
wavelets in a statistical forecasting purpose for time series. Recent approaches involve wavelet decompositions
in order to handle non stationary time series. They study and extended an approach proposed by Renaud et
al., to estimate the prediction equation by direct regression of the process on the Haar non-decimated wavelet
coefficients depending on its past values. The new variants are used first for stationary data and after for
stationary data contaminated by a deterministic trend \[ 3 \].

Jean-Michel Poggi was the supervisor (with A. Antoniadis) of the PhD Thesis of Jairo Cugliari-Duhalde which
takes place in a CIFRE convention with EDF. It is strongly related to the use of wavelets together with curves
clustering in order to perform accurate load consumption forecasting. The thesis develops methodological and
applied aspects linked to the electrical context as well as theoretical ones by introducing exogeneous variables
in the context of nonparametric forecasting time series (see \[ 27 \] and \[ 45 \]).

6.6. Neuroimaging, Statistical analysis of fMRI data

Participants: Gilles Celeux, Christine Keribin.

This research takes place as part of a collaboration with Neurospin on brain functional Magnetic Resonance
Imaging (fMRI) data. ( http://www.math.u-psud.fr/select/reunions/neurospin/Welcome.html ). This year it
concerned essentially regularisation in a supervised clustering methodology that includes spatial information
in the prediction framework, and yields clustered weighted maps.
6. New Results

6.1. Introduction

The new results are organized in the following sections:

1. decision under uncertainty,
2. foundations of machine learning,
3. supervised learning,
4. signal processing (sensor networks),
5. other results.

6.2. Decision Under Uncertainty


6.2.1. Reinforcement learning and approximate dynamic programming

In the domain of reinforcement learning and approximate dynamic programming, we identify two main lines of research.

6.2.1.1. Links between Approximate Dynamic Programming and Statistical Learning Theory

The main objective here is to use tools from statistical learning theory to derive finite-sample performance bounds for RL and ADP algorithms. The goal is to derive bounds on the performance of the policies induced by these algorithms in terms of the number of simulation data and the capacity and approximation power of the considered function and policy spaces. The results of this study allow us to have a better understanding of the functionality of these algorithms and help us to design them more efficiently. The main contributions to this research line in 2011 are:

- **Classification-based Policy Iteration with a Critic** [25], [51]. In collaboration with Bruno Scherrer (INRIA Nancy - Grand Est, Team MAIA) we extended last year work on classification-based policy iteration by adding a value function approximation component (critic) to rollout classification-based policy iteration (RCPI) algorithms. The idea is to use a critic to approximate the return after we truncate the rollout trajectories. This allows us to control the bias and variance of the rollout estimates of the action-value function. Therefore, the introduction of a critic can improve the accuracy of the rollout estimates, and as a result, enhance the performance of the RCPI algorithm. We presented a new RCPI algorithm, called direct policy iteration with critic (DPI-Critic), and provided its finite-sample analysis when the critic is based on the LSTD method. We also empirically evaluated the performance of DPI-Critic and compared it with DPI and LSPI in two benchmark reinforcement learning problems.

- **Finite-Sample Analysis of Least-Squares Policy Iteration** [10], [45]. We extended last year work on the finite-sample analysis of least-squares temporal-difference (LSTD) to the least-squares policy iteration (LSPI) algorithm. In particular, we analyzed how the error at each policy evaluation step is propagated through the iterations of a policy iteration method, and derive a performance bound for the LSPI algorithm.
• Speedy Q-Learning [16], [48]. We introduce a new convergent variant of Q-learning, called speedy Q-learning, to address the problem of slow convergence in the standard form of the Q-learning algorithm. We prove a PAC bound on the performance of SQL, which shows that for an MDP with $n$ state-action pairs and the discount factor $\gamma$ only $T = O\left(\log(n)/\left(\epsilon^2(1-\gamma)^4\right)\right)$ steps are required for the SQL algorithm to converge to an $\epsilon$-optimal action-value function with high probability. This bound has a better dependency on $1/\epsilon$ and $1/(1-\gamma)$, and thus, is tighter than the best available result for Q-learning. Our bound is also superior to the existing results for both model-free and model-based instances of batch Q-value iteration that are considered to be more efficient than the incremental methods like Q-learning.

• Selecting the State-Representation in Reinforcement Learning [34]. The problem of selecting the right state-representation in a reinforcement learning problem is considered. Several models (functions mapping past observations to a finite set) of the observations are given, and it is known that for at least one of these models the resulting state dynamics are indeed Markovian. Without knowing neither which of the models is the correct one, nor what are the probabilistic characteristics of the resulting MDP, it is required to obtain as much reward as the optimal policy for the correct model (or for the best of the correct models, if there are several). We propose an algorithm that achieves that, with a regret of order $T^{2/3}$ where $T$ is the horizon time.

• Transfer from Multiple MDPs [32]. Transfer reinforcement learning (RL) methods leverage on the experience collected on a set of source tasks to speed-up RL algorithms. A simple and effective approach is to transfer samples from source tasks and include them in the training set used to solve a target task. In this paper, we investigate the theoretical properties of this transfer method and we introduce novel algorithms adapting the transfer process on the basis of the similarity between source and target tasks. Finally, we report illustrative experimental results in a continuous chain problem.

6.2.1.2. RL in High-dimensional Spaces

The main objective here is to devise, analyze, implement, and experiment with RL algorithms whose sample and computational complexities do not grow rapidly with the dimension of the state space. We have tackled this problem from two different angles:

• Exploiting the Regularities of the Problem [57], [8], [27]. In order to solve RL in high dimensions, we should exploit all the regularities of the problem in hand. Smoothness is the most common regularity. We continued our collaboration with Amir massoud Farahmand and Csaba Szepesvári at the university of Alberta, Canada, and Shie Mannor at Technion, Israel, on using regularization methods for automatic model selection for value function approximation in RL. We have devised and analyzed the first $\ell_2$-regularized RL algorithms by adding $\ell_2$-regularization to three well-known ADP algorithms: fitted Q-iteration, modified Bellman residual minimization, and least-squares temporal-difference learning [57], [8]. The designed algorithms work in both linear and reproducing kernel Hilbert spaces. Sparsity is another form of regularity that clearly plays a central role in the emerging theory of learning in high dimensions. We have worked on using $\ell_1$-regularization in approximate dynamic programming and RL, which may also serve as a method for feature selection in value function approximation. We have derived finite-sample performance bounds for an algorithm resulting from adding $\ell_1$-penalty to the widely-used least-squares temporal-difference learning (LSTD) algorithm [27].

• Random Projections [28], [52]. We have looked into recent directions popularized in compressive sensing concerning the preservation of properties, such as norm or inner-product, of high dimensional objects when projected on possibly much lower dimensional random subspaces. We have studied the popular LSTD algorithm when a space of low dimension is generated with a random projection from the high-dimensional space, and derived performance bounds for the resulting algorithm [28], [52].

6.2.2. Planning and exploration vs. exploitation trade-off

In the domain of planning and exploration-exploitation algorithms, we identify two main lines of research.
6.2.2.1. Multi-arm Bandit, Online Learning and Optimization

- **Active Learning in Multi-Armed Bandit Problems** [18], [49], [24], [50]. This can be seen as an online allocation problem with several options and is closely related to the problem of *optimal experimental design* in statistics. The objective here is to allocate a fixed budget to a finite (or possibly infinite) number of options (arms) in order to achieve the best accuracy in estimating the quality of each option. In addition to having application in a number of different fields such as *online advertisement* and *personalizing treatment*, this problem is of specific importance in RL in which generating training data is usually expensive. In this framework, we have studied the following two problems: 1) estimating the mean values of all the arms uniformly well in a multi-armed bandit setting [18], [49], and 2) identifying the best arm in each of the bandits in a multi-bandit multi-armed setting [24], [50]. For each problem, we have developed algorithms with theoretical guarantees.

- **Finite Time Analysis of Stratified Sampling for Monte Carlo** [20]. We consider the problem of stratified sampling for Monte-Carlo integration. We model this problem in a multi-armed bandit setting, where the arms represent the strata (an interval in the input domain), and the goal is to estimate a weighted average of the mean values of the arms. We propose a strategy that samples the arms according to an upper bound on their standard deviations and compare its estimation quality to an ideal allocation that would know the standard deviations of the strata. We provide two regret analyses: a distribution-dependent bound $O(n^{-3/2})$ that depends on a measure of the disparity of the strata, and a distribution-free bound $O(n^{-4/3})$ that does not.

- **Optimistic Optimization of a Deterministic Function without the Knowledge of its Smoothness** [36]. We consider a global optimization problem of a deterministic function $f$ in a semi-metric space, given a finite budget of $n$ evaluations. The function $f$ is assumed to be locally smooth (around one of its global maxima) with respect to a semi-metric. We describe two algorithms based on optimistic exploration that use a hierarchical partitioning of the space at all scales. A first contribution is an algorithm, DOO, that requires the knowledge of . We report a finite-sample performance bound in terms of a measure of the quantity of near-optimal states. We then define a second algorithm, SOO, which does not require the knowledge of the semi-metric under which $f$ is smooth, and whose performance is almost as good as DOO optimally-fitted.

- **Finite-Time Analysis of Multi-armed Bandits Problems with Kullback-Leibler Divergences** [35]. We consider a Kullback-Leibler-based algorithm for the stochastic multi-armed bandit problem in the case of distributions with finite supports (not necessarily known beforehand), whose asymptotic regret matches the lower bound of Burnetas and Katehakis (1996). Our contribution is to provide a finite-time analysis of this algorithm; we get bounds whose main terms are smaller than the ones of previously known algorithms with finite-time analyses (like UCB-type algorithms).

- **Adaptive bandits: Towards the best history-dependent strategy** [33]. We consider multi-armed bandit games with possibly adaptive opponents. We introduce models $\Theta$ of constraints based on equivalence classes on the common history (information shared by the player and the opponent) which define two learning scenarios: (1) The opponent is constrained, i.e. he provides rewards that are stochastic functions of equivalence classes defined by some model. The regret is measured with respect to (w.r.t.) the best history-dependent strategy. (2) The opponent is arbitrary and we measure the regret w.r.t. the best strategy among all mappings from classes to actions (i.e. the best history-class-based strategy) for the best model. This allows to model opponents (case 1) or strategies (case 2) which handles finite memory, periodicity, standard stochastic bandits and other situations. When only one model is considered, we derive tractable algorithms achieving a tight regret (at time $T$) bounded by $O(\sqrt{TAC})$, where $C$ is the number of classes. Now, when many models are available, all known algorithms achieving a nice regret $O(\sqrt{T})$ are unfortunately not tractable and scale poorly with the number of models. Our contribution here is to provide tractable algorithms with regret bounded by $T^{2/3}C^{1/3} \log(|\Theta|)^{1/2}$. 

• Pure Exploration in Finitely-Armed and Continuous-Armed Bandits [5]. We consider the framework of stochastic multi-armed bandit problems and study the possibilities and limitations of forecasters that perform an on-line exploration of the arms. These forecasters are assessed in terms of their simple regret, a regret notion that captures the fact that exploration is only constrained by the number of available rounds (not necessarily known in advance), in contrast to the case when the cumulative regret is considered and when exploitation needs to be performed at the same time. We believe that this performance criterion is suited to situations when the cost of pulling an arm is expressed in terms of resources rather than rewards. We discuss the links between the simple and the cumulative regret. One of the main results in the case of a finite number of arms is a general lower bound on the simple regret of a forecaster in terms of its cumulative regret: the smaller the latter, the larger the former. Keeping this result in mind, we then exhibit upper bounds on the simple regret of some forecasters. The paper ends with a study devoted to continuous-armed bandit problems; we show that the simple regret can be minimized with respect to a family of probability distributions if and only if the cumulative regret can be minimized for it. Based on this equivalence, we are able to prove that the separable metric spaces are exactly the metric spaces on which these regrets can be minimized with respect to the family of all probability distributions with continuous mean-payoff functions.

• X-Armed Bandits [6]. We consider a generalization of stochastic bandits where the set of arms, \( X \), is allowed to be a generic measurable space and the mean-payoff function is locally Lipschitz with respect to a dissimilarity function that is known to the decision maker. Under this condition we construct an arm selection policy, called HOO (hierarchical optimistic optimization), with improved regret bounds compared to previous results for a large class of problems. In particular, our results imply that if \( X \) is the unit hypercube in a Euclidean space and the mean-payoff function has a finite number of global maxima around which the behavior of the function is locally continuous with a known smoothness degree, then the expected regret of HOO is bounded up to a logarithmic factor by \( \sqrt{n} \), that is, the rate of growth of the regret is independent of the dimension of the space. We also prove the minimax optimality of our algorithm when the dissimilarity is a metric. Our basic strategy has quadratic computational complexity as a function of the number of time steps and does not rely on the doubling trick. We also introduce a modified strategy, which relies on the doubling trick but runs in linearithmic time. Both results are improvements with respect to previous approaches.

• Learning with Stochastic Inputs and Adversarial Outputs [11]. Most of the research in online learning is focused either on the problem of adversarial classification (i.e., both inputs and labels are arbitrarily chosen by an adversary) or on the traditional supervised learning problem in which samples are independent and identically distributed according to a stationary probability distribution. Nonetheless, in a number of domains the relationship between inputs and outputs may be adversarial, whereas input instances are i.i.d. from a stationary distribution (e.g., user preferences). This scenario can be formalized as a learning problem with stochastic inputs and adversarial outputs. In this paper, we introduce this novel stochastic-adversarial learning setting and we analyze its learnability. In particular, we show that in binary classification, given a hypothesis space \( H \) with finite VC-dimension, it is possible to design an algorithm which incrementally builds a suitable finite set of hypotheses from \( H \) used as input for an exponentially weighted forecaster and achieves a cumulative regret of order \( \sqrt{n \text{VC}(H) \log n} \) with overwhelming probability. This result shows that whenever inputs are i.i.d., it is possible to solve any binary classification problem using a finite VC-dimension hypothesis space with a sub-linear regret independently from the way labels are generated (either stochastic or adversarial). We also discuss extensions to multi-label classification, regression, learning from experts and bandit settings with stochastic side information, and application to games.

• ICML Exploration-Exploitation Challenge [65], [63]. Olivier Nicol and Jérémie Mary won the ICML challenge on Exploration and Exploitation 2 organized by Cambridge on dataset provided by Adobe. The winning approach is based on ideas close to bayesian networks and Thomson sampling as Ad Predictor from Microsoft. These kind of succes emphases the need for better theoretical analysis of theses frameworks. The challenge was also a good occasion to think about the best way
to evaluate online politics (this part also attracts interest from Orange Labs). A publication to JMLR is submitted.

6.2.2. Planning

- **Optimistic Planning for Sparsely Stochastic Systems** [17]. We propose an online planning algorithm for finite action, sparsely stochastic Markov decision processes, in which the random state transitions can only end up in a small number of possible next states. The algorithm builds a planning tree by iteratively expanding states, where each expansion exploits sparsity to add all possible successor states. Each state to expand is actively chosen to improve the knowledge about action quality, and this allows the algorithm to return a good action after a strictly limited number of expansions. More specifically, the active selection method is optimistic in that it chooses the most promising states first, so the novel algorithm is called optimistic planning for sparsely stochastic systems. We note that the new algorithm can also be seen as model-predictive (receding-horizon) control. The algorithm obtains promising numerical results, including the successful online control of a simulated HIV infection with stochastic drug effectiveness.

- **Optimistic Planning in Markov decision processes** [46]. We review a class of online planning algorithms for deterministic and stochastic optimal control problems, modeled as Markov decision processes. At each discrete time step, these algorithms maximize the predicted value of planning policies from the current state, and apply the first action of the best policy found. An overall receding-horizon algorithm results, which can also be seen as a type of model-predictive control. The space of planning policies is explored optimistically, focusing on areas with largest upper bounds on the value or upper confidence bounds, in the stochastic case. The resulting optimistic planning framework integrates several types of optimism previously used in planning, optimization, and reinforcement learning, in order to obtain several intuitive algorithms with good performance guarantees. We describe in detail three recent such algorithms, outline the theoretical guarantees on their performance, and illustrate their behavior in a numerical example.

6.2.3. Applications

6.2.3.1. Management of ad campaigns on the web

More work has been dedicated to the topic aiming at optimizing ad campaigns on the web under real-time constraints, in a dynamic environment [9].

6.3. Foundations of Machine Learning

Participants: Daniil Ryabko, Azadeh Khaleghi, Romaric Gaudel.

6.3.1. Sequence prediction in the most general form

The problem of sequence prediction consists in forecasting, on each step of time $n$, the probabilities of the next outcome of the observed sequence of data $x_1, x_2, \ldots, x_n, \ldots$. In the most general formulation of the problem, we assume that we are given a set $\mathcal{C}$ of probability measures (on the space of infinite sequences). We can then assume that the sequence is generated by an unknown measure $\mu$ that belongs to $\mathcal{C}$, or that the measure $\mu$ is arbitrary, but we compare the performance of our predictor to that of the best predictor in $\mathcal{C}$.

6.3.1.1. Relation between the realizable and non-realizable cases of the sequence prediction problem

The realizable case of the sequence prediction problem is when the measure $\mu$ belongs to an arbitrary but known class $\mathcal{C}$ of process measures. The non-realizable case is when $\mu$ is completely arbitrary, but the prediction performance is measured with respect to a given set $\mathcal{C}$ of process measures. We are interested in the relations between these problems and between their solutions, as well as in characterizing the cases when a solution exists, and finding these solutions. In this work [13] we show that if the quality of prediction is measured by total variation distance, then these problems coincide, while if it is measured by expected average KL-divergence, then they are different. For some of the formalizations we also show that when a solution exists, it can be obtained as a Bayes mixture over a countable subset of $\mathcal{C}$. As an illustration to the
general results obtained, we show that a solution to the non-realizable case of the sequence prediction problem exists for the set of all finite-memory processes, but does not exist for the set of all stationary processes.

6.3.2. Statistical inference

We continue to obtain new results using the theoretical framework developed recently for studying time series generated by stationary ergodic time series. This year, new results obtained include a topological characterizing of composite hypotheses for which consistent tests exist, as well as new results on clustering.

6.3.2.1. A criterion for the existence of consistent tests

The most general result that we have obtained [14] on hypothesis testing provides a complete characterization (necessary and sufficient conditions) for the existence of a consistent test for membership to an arbitrary family \( H_0 \) of stationary ergodic discrete-valued processes, against \( H_1 \) which is the complement of \( H_0 \) to this class of processes. The criterion is that \( H_0 \) has to be closed in the topology of distributional distance, and closed under taking ergodic decompositions of its elements.

6.3.3. Clustering

6.3.3.1. Online clustering of time series

An asymptotically consistent algorithm has been proposed for the problem of online clustering of time series. There is a growing body of time series samples, each of which grows with time. On each time step, it is required to group these time series into \( k \) clusters. It is known that each of the time series is generated by one out of \( k \) stationary ergodic distributions. An algorithm is proposed that, for each fixed portion of samples, eventually (with probability 1) puts into the same group those and only those samples that were generated by the same distribution. Empirical performance of the algorithm is evaluated on synthetic and real data.

6.3.3.2. Clustering of ranked data

We introduced [47] a novel approach to clustering rank data on a set of possibly large cardinality \( n \in \mathbb{N}^+ \), relying upon Fourier representation of functions defined on the symmetric group \( S_n \). In the proposed setup, covering a wide variety of practical situations, rank data are viewed as distributions on \( S_n \). Cluster analysis aims at segmenting data into homogeneous subgroups, hopefully very dissimilar in a certain sense. Whereas considering dissimilarity measures/distances between distributions on the non commutative group \( S_n \), in a coordinate manner by viewing it as embedded in the set \( [0, 1]^n \) for instance, hardly yields interpretable results and leads to face obvious computational issues, evaluating the closeness of groups of permutations in the Fourier domain may be much easier in contrast. Indeed, in a wide variety of situations, a few well-chosen Fourier (matrix) coefficients may permit to approximate efficiently two distributions on \( S_n \) as well as their degree of dissimilarity, while describing global properties in an interpretable fashion. Following in the footsteps of recent advances in automatic feature selection in the context of unsupervised learning, we propose to cast the task of clustering rankings in terms of optimization of a criterion that can be expressed in the Fourier domain in a simple manner.

6.4. Supervised learning


6.4.1. Regression and classification

- Sparse Recovery with Brownian Sensing [19].

We consider the problem of recovering the parameter \( \alpha \) of a sparse function \( f \) (i.e. the number of non-zero entries of \( \alpha \) is small compared to the number \( K \) of features) given noisy evaluations of \( f \) at a set of well-chosen sampling points. We introduce an additional randomization process, called Brownian sensing, based on the computation of stochastic integrals, which produces a Gaussian sensing matrix, for which good recovery properties are proven, independently on the number of
sampling points $N$, even when the features are arbitrarily non-orthogonal. Under the assumption that $f$ is Hölder continuous with exponent at least $1/2$ we provide an estimate of the parameter with quadratic error $O(||\eta||/N)$, where $\eta$ is the observation noise. The method uses a set of sampling points uniformly distributed along a one-dimensional curve selected according to the features. We report numerical experiments illustrating our method.

- **Operator-valued Kernels for Nonlinear FDA** [31], [38], [30], [53] Following the extension of RKHS to functional setting [74], we further developed this work in [38] for functional supervised classification.

  We introduced a set of rigorously defined operator-valued kernels that can be valuably applied to nonparametric operator learning when input and output data are continuous smooth functions, and we have showed their use for solving the problem of minimizing a $L^2$-regularized functional in the case of functional outputs without the need to discretize covariate and target functions [53].

  The framework developed can also be applied when the input data are both discrete and continuous [30].

  Our fully functional approach has been successfully applied to the problems of speech inversion [31] and sound recognition [38], showing that the proposed framework is particularly relevant for audio signal processing applications where attributes are really functions and dependent of each other.

  This work is done in collaboration with Francis Bach (INRIA, Sierra), Alain Rakotomamonjy and Stéphane Canu (LITIS, Rouen).

- **Datum-wise representation** [44], [54]. We consider supervised classification. We introduce the concept of datum-wise representation for supervised classification [44]. While traditional approaches yield a “best” representation at the data space level, that is, the same representation is used for all the data, we proposed the idea, as well as an algorithm, that yields the “best” representation for each data. Among other appealing properties, this leads to sparse representation of each data, and an averaged sparser representation of each data in the data space. Along a classifier, the learning algorithm produces a “representer”, that is a function that yields a representation given a data.

  We further improved this approach to encompass various settings which are traditionally kept as different (cost-sensitive classification and different structured sparsity) [54].

- **Iso-regularization descent** [1]. Manuel Loth has defended his PhD dissertation [1] where he has provided a detailed presentation and analysis of his algorithm to solve the LASSO. This algorithm is very efficient. It is an active set algorithm that solves the LASSO by considering it a convex problem with linear constraints.

- **Learning with few examples** [41], [64]. Christophe Salperwyck has studied the performance of various classifiers when few examples are available. This is an important point in incremental learning, and few studies have been devoted to this particular setting. Performance we are accustomed to when the examples are quite numerous are severely disturbed in this setting. For more details, please see [41], [64].

- **Incremental discretization** [40]. In incremental learning, discretization should be adaptive in order to cope with the values of the attributes that are observed. This issue is currently under study by Christophe Salperwyck [40].

#### 6.5. Sensors Networks: Tracking, Localization and Communication

**Participants:** Emmanuel Delande, Emmanuel Duflos, Pierre Chainais, Philippe Vanheeghe.
6.5.1. The sensor management problem

The aim of this work is to manage a set of sensors to track vehicles or groups of people in land applications. Our work focuses on sensor management in the frame of the random finite sets where the Probability Hypothesis Density (PHD) is a well-known method for single-sensor multi-target tracking problems in a Bayesian framework, but the extension to the multi-sensor case seems to remain a challenge. We have proposed an extension of Mahler’s work to the multi-sensor case by providing an expression of the true PHD multi-sensor data update equation. Then, based on the configuration of the sensors fields of view (FOVs), a joint partitioning of both the sensors and the state space provides an equivalent yet more practical expression of the data update equation, allowing a more effective implementation in specific FOV configurations ([70]). This work is done in collaboration with Thales Communications. The multi-sensor / multi-target filtering problem by using PHD filtering methods are topics developed in the Th PhD thesis of Emmanuel Delande. This PhD thesis entitled "Multi-sensor PHD filtering with application to sensor management" will be defended in December 2011. In addition to the different questions described above, see also [22] and [23]. Then, a new approach using operational objectives, related to the type of application, for sensor manager is proposed.

6.5.2. Statistical signal processing: application to civil engineering

We have obtained a PICS (International Project for Scientific Cooperation) from the CNRS in 2008 for 3 years to work in cooperation with the Department of Civil and Environmental Engineering of the University of Waterloo (Canada). During this cooperation we have developed a belief functions based method to track the building materials on a construction site. ([71]). Based on this cooperation, during 2011 a new common research project with the same department of the University of Waterloo has been built, and is actually submitted for funding. The topic of this project is the using of nonparametric Bayesian models in the area of Non-destructive Testing.

6.5.3. Accurate Localization using Satellites in Urban Canyons

Today, Global Navigation Satellite Systems (GNSS) have penetrated the transport field through applications such as monitoring of containers. These applications do not necessarily request a high availability, integrity and accuracy of the positioning system. For safety applications (as complete guidance of autonomous vehicles), performances require to be more stringent. For, sensors may deliver very erroneous measurements because of such hard external conditions which reduce significantly the possibilities to receive direct signals. The consequences of environmental obstructions are unavailability of the service and reception of reflected signals that degrades in particular the accuracy of the positioning. Indeed, NLOS (Non Line Of Sight) signals, i.e. signals received after reflections on the surrounding obstacles, frequently occur in dense environments and degrade localization accuracy because of the delays observed on the propagation time measurement creating additional error on pseudorange estimation. In the previous years we have proposed new algorithms to improve the localization precision. This algorithm are based on two principles: a jump multimodel approach and a joint state - noise density estimation. We have focused this year on an approach using Dirichlet Process Mixture to track the noise density in urban canyon while estimating the position of the vehicle. Algorithm have been validated on real data collected in a French town : Belfort. Nicolas Viandier has defended his PhD on this subject on June 2011. ([76], [75], [84], [85], [4]). These results will be presented to the Workshop Non Parametric Bayes at the NIPS Conference en Decembre 2011 ([62]) and to the ICASSP 2011 Conference ([37]).

6.5.4. Internet of Things : Mitigation of Impulsive Noise Effects

The term "Internet of Things" has come to describe a number of technologies and research disciplines that enable the Internet to reach out into the real world of physical objects. Technologies like RFID, short-range wireless communications, real-time localization and sensor networks are now becoming increasingly common, bringing the Internet of Things into commercial use. In such applications the data sent by a thing to another may generate an impulse noise in the reception channel of objects in the neighbourhood. The noise appearing in such applications can be considered as α-stable. In this context, we’ve tackled the problem of interference mitigation in ad hoc networks. In such context, the multiple access interference (MAI) is known to be of an
impulsive nature. Therefore, the conventional Gaussian assumption cannot be considered to model this type of interference. Contrariwise, it can be accurately modeled by stable distributions. Here, this issue is addressed within an Orthogonal Frequency Division Multiplexing (OFDM) transmission link assuming a symmetric $\alpha$-stable model for the signal distortion due to MAI. We have proposed a method for the joint estimation of the transmitted multicarrier signal and the noise parameters. Based on sequential Monte Carlo (SMC) methods, the proposed scheme allows the online estimation using a Rao-blackwellized particle filter. These results have been presented to the International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2011) [29]. We are now focusing on Bayesian non-linear filtering with non-stationary $\alpha$-stable measurement noise. We have shown that a Dirichlet Process Mixture can improve the estimation by modelling the noise by both an infinite Cauchy mixture or an infinite $\alpha$-stable mixture. These first results will be presented to the Workshop Non-Parametric Bayes at the NIPS Conference in December 2011 [62].

6.5.5. Image processing and statistical image modeling

Pierre Chainais arrived in Sequel in September 2010 with the purpose of a thematic evolution toward non-parametric Bayesian approaches. This represents an important investment in very new directions on an emerging topic at the interface between machine learning and signal/image processing. Discussions have begun with Emmanuel Duflos and Philippe Vanheeghe on the use of non-parametric Bayesian approaches to blind deconvolution of noisy natural images. The main objective is to use together the typical structure and sparsity of space-scale representations of images.

Pierre Chainais has continued working on several older projects. One of them deals with the segmentation of nanotubes in microscopic imaging [12], [43]. B. Lebental at IFFSTAR works on the conception of new nano-sensors based on the use of carbon nanotubes to build a nano-membrane. P. Chainais has developed an image processing pipeline to analyse images of these nanomembranes so as to characterize their properties in a precise and objective manner. Among other properties, the histograms of orientations of the nanotubes is provided. This tool will be very useful since such nanosensors are becoming more and more common.

In solar astronomy [7], [21], we have proposed a tool for the virtual super-resolution of scale invariant textured images. The aim of this project was to provide astronomers with plausible high-resolution images to calibrate next generation spatial telescopes. In particular, our images can be used to optimize the compression algorithm to be embedded in a spatial telescope. In collaboration with M. Chevaldonné and J-M. Favreau (Université Clermont-Ferrand I), we work a software for texture synthesis on 3D surfaces [42] based on multifractal processes. A first version of the software is under current development. More marginal is our work on the use of stochastic processes for the simulation of turbulent pressure fields in collaboration with M. Pachebat (Laboratoire de Mécanique et d’Acoustique de Marseille) and Nicolas Totaro (LVA, INSA Lyon).
6. New Results

6.1. Local Component Analysis

Participants: Francis Bach, Nicolas Le Roux.

Kernel density estimation, a.k.a. Parzen windows, is a popular density estimation method, which can be used for outlier detection or clustering. With multivariate data, its performance is heavily reliant on the metric used within the kernel. Most earlier work has focused on learning only the bandwidth of the kernel (i.e., a scalar multiplicative factor). In this paper, we propose to learn a full Euclidean metric through an expectation-minimization (EM) procedure, which can be seen as an unsupervised counterpart to neighbourhood component analysis (NCA). In order to avoid overfitting with a fully nonparametric density estimator in high dimensions, we also consider a semi-parametric Gaussian-Parzen density model, where some of the variables are modelled through a jointly Gaussian density, while others are modelled through Parzen windows. For these two models, EM leads to simple closed-form updates based on matrix inversions and eigenvalue decompositions. We show empirically that our method leads to density estimators with higher test-likelihoods than natural competing methods, and that the metrics may be used within most unsupervised learning techniques that rely on such metrics, such as spectral clustering or manifold learning methods. Finally, we present a stochastic approximation scheme which allows for the use of this method in a large-scale setting [30].

6.2. Weakly Supervised Learning of Foreground-Background Segmentation using Masked RBMs

Participant: Nicolas Le Roux.

Collaboration with: Nicolas Heess (School of Informatics, University of Edinburgh) and John Winn (Machine Learning and Perception, Microsoft research Cambridge).

We propose an extension of the Restricted Boltzmann Machine (RBM) that allows the joint shape and appearance of foreground objects in cluttered images to be modeled independently of the background. We present a learning scheme that learns this representation directly from cluttered images with only very weak supervision. The model generates plausible samples and performs foreground-background segmentation. We demonstrate that representing foreground objects independently of the background can be beneficial in recognition tasks [15].

6.3. Clusterpath: an algorithm for clustering using convex fusion penalties

Participants: Toby Hocking, Francis Bach, Armand Joulin.


We present a new clustering algorithm by proposing a convex relaxation of hierarchical clustering, which results in a family of objective functions with a natural geometric interpretation. We give efficient algorithms for calculating the continuous regularization path of solutions, and discuss relative advantages of the parameters. Our method experimentally gives state-of-the-art results similar to spectral clustering for non-convex clusters, and has the added benefit of learning a tree structure from the data [16].

6.4. Convex and Network Flow Optimization for Structured Sparsity

Participants: Rodolphe Jenatton, Guillaume Obozinski, Francis Bach.

Collaboration with: Julien Mairal (Department of Statistics, University of California, Berkeley).
In [5], we consider a class of learning problems regularized by a structured sparsity-inducing norm defined as the sum of $\ell_2$- or $\ell_\infty$-norms over groups of variables. Whereas much effort has been put in developing fast optimization techniques when the groups are disjoint or embedded in a hierarchy, we address here the case of general overlapping groups. To this end, we present two different strategies: On the one hand, we show that the proximal operator associated with a sum of $\ell_\infty$-norms can be computed exactly in polynomial time by solving a quadratic min-cost flow problem, allowing the use of accelerated proximal gradient methods. On the other hand, we use proximal splitting techniques, and address an equivalent formulation with non-overlapping groups, but in higher dimension and with additional constraints. We propose efficient and scalable algorithms exploiting these two strategies, which are significantly faster than alternative approaches. We illustrate these methods with several problems such as CUR matrix factorization, multi-task learning of tree-structured dictionaries, background subtraction in video sequences, image denoising with wavelets, and topographic dictionary learning of natural image patches.

6.5. Multi-scale Mining of fMRI data with Hierarchical Structured Sparsity

**Participants:** Rodolphe Jenatton, Guillaume Obozinski, Francis Bach.

Collaboration with: Alexandre Gramfort, Vincent Michel, Evelyn Eger and Bertrand Thirion (Laboratoire de Neuroimagerie Assistée par Ordinateur (LNAO), CEA: DSV/I2BM/NEUROSPIN, PARIETAL (INRIA Saclay - Ile de France) and Neuroimagerie cognitive, INSERM: U992 – Université Paris Sud – CEA).

Inverse inference, or "brain reading", is a recent paradigm for analyzing functional magnetic resonance imaging (fMRI) data, based on pattern recognition and statistical learning. By predicting some cognitive variables related to brain activation maps, this approach aims at decoding brain activity. Inverse inference takes into account the multivariate information between voxels and is currently the only way to assess how precisely some cognitive information is encoded by the activity of neural populations within the whole brain. However, it relies on a prediction function that is plagued by the curse of dimensionality, since there are far more features than samples, i.e., more voxels than fMRI volumes. To address this problem, different methods have been proposed, such as, among others, univariate feature selection, feature agglomeration and regularization techniques. In this paper, we consider a sparse hierarchical structured regularization. Specifically, the penalization we use is constructed from a tree that is obtained by spatially-constrained agglomerative clustering. This approach encodes the spatial structure of the data at different scales into the regularization, which makes the overall prediction procedure more robust to inter-subject variability. The regularization used induces the selection of spatially coherent predictive brain regions simultaneously at different scales. We test our algorithm on real data acquired to study the mental representation of objects, and we show that the proposed algorithm non only delineates meaningful brain regions but yields as well better prediction accuracy than reference methods [24], [17].

6.6. Convergence Rates of Inexact Proximal-Gradient Methods for Convex Optimization

**Participants:** Mark Schmidt, Nicolas Le Roux, Francis Bach.

In [21], we consider the problem of optimizing the sum of a smooth convex function and a non-smooth convex function using proximal-gradient methods, where an error is present in the calculation of the gradient of the smooth term or in the proximity operator with respect to the non-smooth term. We show that both the basic proximal-gradient method and the accelerated proximal-gradient method achieve the same convergence rate as in the error-free case, provided that the errors decrease at appropriate rates. Using these rates, we perform as well as or better than a carefully chosen fixed error level on a set of structured sparsity problems.

6.7. Generalized Fast Approximate Energy Minimization via Graph Cuts: Alpha-Expansion Beta-Shrink Moves

**Participant:** Mark Schmidt.
In [20], we present alpha-expansion beta-shrink moves, a simple generalization of the widely-used alpha-beta swap and alpha-expansion algorithms for approximate energy minimization. We show that in a certain sense, these moves dominate both alpha-beta-swap and alpha-expansion moves, but unlike previous generalizations the new moves require no additional assumptions and are still solvable in polynomial-time. We show promising experimental results with the new moves, which we believe could be used in any context where alpha-expansions are currently employed.

6.8. Hybrid Deterministic-Stochastic Methods for Data Fitting

**Participant:** Mark Schmidt.

Collaboration with: Michael P. Friedlander (University of British Columbia).

Many structured data-fitting applications require the solution of an optimization problem involving a sum over a potentially large number of measurements. Incremental gradient algorithms offer inexpensive iterations by sampling only subsets of the terms in the sum. These methods can make great progress initially, but often slow as they approach a solution. In contrast, full gradient methods achieve steady convergence at the expense of evaluating the full objective and gradient on each iteration. We explore hybrid methods that exhibit the benefits of both approaches. Rate of convergence analysis shows that by controlling the size of the subsets in an incremental gradient algorithm, it is possible to maintain the steady convergence rates of full gradient methods. We detail a practical quasi-Newton implementation based on this approach, and numerical experiments illustrate its potential benefits [29].

6.9. Multi-task regression using minimal penalties

**Participants:** Matthieu Solnon, Sylvain Arlot, Francis Bach.

In [33] we study the kernel multiple ridge regression framework, which we refer to as multi-task regression, using penalization techniques. The theoretical analysis of this problem shows that the key element appearing for an optimal calibration is the covariance matrix of the noise between the different tasks. We present a new algorithm to estimate this covariance matrix, based on the concept of minimal penalty, which was previously used in the single-task regression framework to estimate the variance of the noise. We show, in a non-asymptotic setting and under mild assumptions on the target function, that this estimator converges towards the covariance matrix. Then plugging this estimator into the corresponding ideal penalty leads to an oracle inequality. We illustrate the behavior of our algorithm on synthetic examples.

6.10. Ask the locals: multi-way local pooling for image recognition

**Participants:** Nicolas Le Roux, Francis Bach.

Collaboration with: Y-Lan Boureau and Jean Ponce (Willow project-team, INRIA Paris-Rocquencourt) and Yann LeCun (Courant Institute of Mathematical Science (CIMS), New York University).

Invariant representations in object recognition systems are generally obtained by pooling feature vectors over spatially local neighborhoods. But pooling is not local in the feature vector space, so that widely dissimilar features may be pooled together if they are in nearby locations. Recent approaches rely on sophisticated encoding methods and more specialized codebooks (or dictionaries), e.g., learned on subsets of descriptors which are close in feature space, to circumvent this problem. In this work, we argue that a common trait found in much recent work in image recognition or retrieval is that it leverages locality in feature space on top of purely spatial locality. We propose to apply this idea in its simplest form to an object recognition system based on the spatial pyramid framework, to increase the performance of small dictionaries with very little added engineering. State-of-the-art results on several object recognition benchmarks show the promise of this approach [12].
6.11. Trace Lasso: a trace norm regularization for correlated designs
Participants: Edouard Grave, Guillaume Obozinski, Francis Bach.

Using the $\ell_1$-norm to regularize the estimation of the parameter vector of a linear model leads to an unstable estimator when covariates are highly correlated. In this paper, we introduce a new penalty function which takes into account the correlation of the design matrix to stabilize the estimation. This norm, called the trace Lasso, uses the trace norm of the selected covariates, which is a convex surrogate of their rank, as the criterion of model complexity. We analyze the properties of our norm, describe an optimization algorithm based on reweighted least-squares, and illustrate the behavior of this norm on synthetic data, showing that it is more adapted to strong correlations than competing methods such as the elastic net [14].

6.12. Shaping level sets through submodular functions
Participant: Francis Bach.

The concept of parsimony is central in many scientific domains. In the context of statistics, signal processing or machine learning, it may take several forms. Classically, in a variable or feature selection problem, a sparse solution with many zeros is sought so that the model is either more interpretable, cheaper to use, or simply matches available prior knowledge. In this work, we instead consider sparsity-inducing regularization terms that will lead to solutions with many equal values. A classical example is the total variation in one or two dimensions, which leads to piecewise constant solutions and can be applied to various image labelling problems, or change point detection tasks. In this work [9], we follow our earlier approach which consisted in designing sparsity-inducing norms based on non-decreasing submodular functions, as a convex approximation to imposing a specific prior on the supports of the predictors. Here, we show that a similar parallel holds for some other class of submodular functions, namely non-negative setfunctions which are equal to zero for the full and empty set. Our main instance of such functions are symmetric submodular functions and we show that the Lovász extension may be seen as the convex envelope of a function that depends on level sets (i.e., the set of indices whose corresponding components of the underlying predictor are greater than a given constant). By selecting specific submodular functions, we give a new interpretation to known norms, such as the total variation; we also define new norms, in particular ones that are based on order statistics with application to clustering and outlier detection, and on noisy cuts in graphs with application to change point detection in the presence of outliers.

6.13. Itakura-Saito Nonnegative Matrix Factorization with group sparsity
Participants: Augustin Lefèvre, Francis Bach.

Collaboration with: Cédric Févotte (Laboratoire traitement et communication de l’information (LTCI), CNRS: UMR5141 – Institut Télécom – Télécom ParisTech).

In [18], we propose an unsupervised inference procedure for audio source separation. Components in nonnegative matrix factorization (NMF) are grouped automatically in audio sources via a penalized maximum likelihood approach. The penalty term we introduce favors sparsity at the group level, and is motivated by the assumption that the local amplitude of the sources are independent. Our algorithm extends multiplicative updates for NMF; moreover we propose a test statistic to tune hyperparameters in our model, and illustrate its adequacy on synthetic data. Results on real audio tracks show that our sparsity prior allows to identify audio sources without knowledge on their spectral properties.

6.14. Online algorithms for nonnegative matrix factorization with the Itakura-Saito divergence
Participants: Augustin Lefèvre, Francis Bach.

Collaboration with: Cédric Févotte (Laboratoire traitement et communication de l’information (LTCI), CNRS: UMR5141 – Institut Télécom – Télécom ParisTech).
Nonnegative matrix factorization (NMF) is now a common tool for audio source separation. When learning NMF on large audio databases, one major drawback is that the complexity in time is $O(FKN)$ when updating the dictionary (where $(F;N)$ is the dimension of the input power spectrograms, and $K$ the number of basis spectra), thus forbidding its application on signals longer than an hour. We provide an online algorithm with a complexity of $O(FK)$ in time and memory for updates in the dictionary. We show on audio simulations that the online approach is faster for short audio signals and allows to analyze audio signals of several hours [31].

6.15. A Graph-matching Kernel for Object Categorization

**Participant:** Armand Joulin.

Collaboration with: Olivier Duchenne and Jean Ponce (Willow project-team, INRIA Paris-Rocquencourt).

In [13], we address the problem of category-level image classification. The underlying image model is a graph whose nodes correspond to a dense set of regions, and edges reflect the underlying grid structure of the image and act as springs to guarantee the geometric consistency of nearby regions during matching. A fast approximate algorithm for matching the graphs associated with two images is presented. This algorithm is used to construct a kernel appropriate for SVM-based image classification, and experiments with the Caltech 101, Caltech 256, and Scenes datasets demonstrate performance that matches or exceeds the state of the art for methods using a single type of features.


**Participant:** Guillaume Obozinski.

Collaboration with: Laurent Jacob (Department of Statistics, University of California at Berkeley) and Jean-Philippe Vert (INSERM U900, Mines ParisTech, Institut Curie).

We study in [25] a norm for structured sparsity which leads to sparse linear predictors whose supports are unions of predefined overlapping groups of variables. We call the obtained formulation latent group Lasso, since it is based on applying the usual group Lasso penalty on a set of latent variables. A detailed analysis of the norm and its properties is presented and we characterize conditions under which the set of groups associated with latent variables are correctly identified. We motivate and discuss the delicate choice of weights associated to each group, and illustrate this approach on simulated data and on the problem of breast cancer prognosis from gene expression data.

6.17. Sparse Image Representation with Epitomes

**Participants:** Louise Benoît, Francis Bach.

Collaboration with: Julien Mairal (Department of Statistics, University of California, Berkeley) and Jean Ponce (Willow project-team, INRIA Paris-Rocquencourt).

Sparse coding, which is the decomposition of a vector using only a few basis elements, is widely used in machine learning and image processing. The basis set, also called dictionary, is learned to adapt to specific data. This approach has proven to be very effective in many image processing tasks. Traditionally, the dictionary is an unstructured "flat" set of atoms. In this work, we study structured dictionaries which are obtained from an epitome, or a set of epitomes. The epitome is itself a small image, and the atoms are all the patches of a chosen size inside this image. This considerably reduces the number of parameters to learn and provides sparse image decompositions with shift invariance properties. We propose a new formulation and an algorithm for learning the structured dictionaries associated with epitomes, and illustrate their use in image denoising tasks. This work has resulted in a CVPR’11 publication [11].

6.18. Dictionary Learning for Deblurring and Digital Zoom

**Participant:** Florent Couzinié-Devy.
The paper [23] proposes a novel approach to image deblurring and digital zooming using sparse local models of image appearance. These models, where small image patches are represented as linear combinations of a few elements drawn from some large set (dictionary) of candidates, have proven well adapted to several image restoration tasks. A key to their success has been to learn dictionaries adapted to the reconstruction of small image patches [35]. In contrast, recent works have proposed instead to learn dictionaries which are not only adapted to data reconstruction, but also tuned for a specific task [34]. We introduce here such an approach to deblurring and digital zoom, using pairs of blurry/sharp (or low-/high-resolution) images for training, as well as an effective stochastic gradient algorithm for solving the corresponding optimization task. Although this learning problem is not convex, once the dictionaries have been learned, the sharp/high-resolution image can be recovered via convex optimization at test time. Experiments with synthetic and real data demonstrate the effectiveness of the proposed approach, leading to state-of-the-art performance for non-blind image deblurring and digital zoom.

6.19. Robust linear least squares regression

Participant: Jean-Yves Audibert.

Collaboration with: Olivier Catoni (École Normale Supérieure, CNRS and INRIA Paris-Rocquencourt, Classic project-team)

In [26], we consider the problem of robustly predicting as well as the best linear combination of d given functions in least squares regression, and variants of this problem including constraints on the parameters of the linear combination. For the ridge estimator and the ordinary least squares estimator, and their variants, we provide new risk bounds of order d/n without logarithmic factor unlike some standard results, where n is the size of the training data. We also provide a new estimator with better deviations in presence of heavy-tailed noise. It is based on truncating differences of losses in a min-max framework and satisfies a d/n risk bound both in expectation and in deviations. The key common surprising factor of these results is the absence of exponential moment condition on the output distribution while achieving exponential deviations. All risk bounds are obtained through a PAC-Bayesian analysis on truncated differences of losses. Experimental results strongly back up our truncated min-max estimator. This work is to appear in the Annals of Statistics in 2012.

6.20. Semantic hierarchies for image annotation

Participant: Jean-Yves Audibert.

Collaboration with: Anne-Marie Tousch (École des Ponts and ONERA) and Stéphane Herbin (ONERA)

In the survey [6], we argue that using structured vocabularies is capital to the success of image annotation. We analyze literature on image annotation uses and user needs, and we stress the need for automatic annotation. We briefly expose the difficulties posed to machines for this task and how it relates to controlled vocabularies. We survey contributions in the field showing how structures are introduced. First we present studies that use unstructured vocabulary, focusing on those introducing links between categories or between features. Then we review work using structured vocabularies as an input and we analyze how the structure is exploited.

6.21. Deviations of Stochastic Bandit Regret

Participant: Jean-Yves Audibert.

Collaboration with: Antoine Salomon (École des Ponts)

The work [19] studies the deviations of the regret in a stochastic multi-armed bandit problem. When the total number of plays n is known beforehand by the agent, previous works exhibit a policy such that with probability at least $1 - 1/n$, the regret of the policy is of order $\log n$. They have also shown that such a property is not shared by the popular UCB1 policy. This work first answers an open question; it extends this negative result to any anytime policy. The second contribution of this paper is to design anytime robust policies for specific multi-armed bandit problems in which some restrictions are put on the set of possible distributions of the different arms.
6.22. Minimax Policies for Combinatorial Prediction Games

Participant: Jean-Yves Audibert.

Collaboration with: Sébastien Bubeck (Centre de Recerca Matematica of Barcelona) and Gabor Lugosi (ICREA and Pompeu Fabra University)

In [8], we address the online linear optimization problem when the actions of the forecaster are represented by binary vectors. Our goal is to understand the magnitude of the minimax regret for the worst possible set of actions. We study the problem under three different assumptions for the feedback: full information, and the partial information models of the so-called "semi-bandit", and "bandit" problems. We consider both $L_\infty$-, and $L_2$-type of restrictions for the losses assigned by the adversary. We formulate a general strategy using Bregman projections on top of a potential-based gradient descent, which generalizes the ones studied in numerous recent works. We provide simple proofs that recover most of the previous results. We propose new upper bounds for the semi-bandit game. Moreover we derive lower bounds for all three feedback assumptions. With the only exception of the bandit game, the upper and lower bounds are tight, up to a constant factor. Finally, we answer an open question raised by W. M. Koolen, M. K. Warmuth, and J. Kivinen by showing that the exponentially weighted average forecaster is suboptimal against $L_\infty$ adversaries.


Participant: Francis Bach.

Collaboration with: Eric Moulines (Telecom ParisTech)

In [10], we consider the minimization of a convex objective function defined on a Hilbert space, which is only available through unbiased estimates of its gradients. This problem includes standard machine learning algorithms such as kernel logistic regression and least-squares regression, and is commonly referred to as a stochastic approximation problem in the operations research community. We provide a non-asymptotic analysis of the convergence of two well-known algorithms, stochastic gradient descent (a.k.a. Robbins-Monro algorithm) as well as a simple modification where iterates are averaged (a.k.a. Polyak-Ruppert averaging). Our analysis suggests that a learning rate proportional to the inverse of the number of iterations, while leading to the optimal convergence rate in the strongly convex case, is not robust to the lack of strong convexity or the setting of the proportionality constant. This situation is remedied when using slower decays together with averaging, robustly leading to the optimal rate of convergence. We illustrate our theoretical results with simulations on synthetic and standard datasets.
5. New Results

5.1. Statistical Physics

The analysis of multi-scale phenomena and asymptotic problems aiming at identifying the influence of microscopic scales on the macroscopic observations is a hot topic in the team. Results have been obtained concerning the derivation of effective law describing the behavior of a particle interacting with a thermal bath or a set of oscillators. This work, which combines modeling efforts, analysis and large computations, is the object of a longstanding collaboration with P. Parris (Missouri-Rolla) and is the heart of the PhD thesis of B. Aguer [37], [1]. More recently, S. De Bièvre and P. Parris have described the approach to equilibrium in such systems [11]. Transport in a related quantum mechanical model was studied in [6]. Some long time effective behavior of related models has been obtained in [58].

At the same time, M. Rousset is working on the numerical simulation of stochastically perturbed Molecular Dynamics. The main goal is to handle in the same simulation the fastest time scales (the oscillations of molecular bindings), and the slowest time scales (the so-called reaction coordinates). Recently, a monograph [60] has been published which summarized standard and state-of-the-art free energy calculations, that are used to accelerate slow variables in MD simulations. In [61] analysis of constrained dynamics is proposed, with associated numerical schemes. In [62], a new method has been proposed which drastically slows down the fast frequencies with a penalty and accelerates simulations, while conserving the statistical behavior of molecular systems.

Recently, in [64] M. Rousset has initiated some new research on variance reduction in hybrid methods, where a “fine-grained” model, typically a kinetic model, is simulated with particle/Monte-Carlo method, and the variance of the latter is reduced using the information of a “coarse-grained” model, a PDE computed with a grid method.

5.2. Hyperbolic Problems, Conservation laws and Gas Dynamics

The convergence analysis of numerical schemes for conservation laws with unstructured meshes with an original proof based on probabilistic argument is a striking result due to F. Lagoutière with F. Delarue, [56]. More generally, we refer to [59] for an overview of F. Lagoutière’s works.

J.-F. Coulombel has studied the stability of finite difference approximations of hyperbolic systems with boundary conditions. This series of works, part of which is a collaboration with A. Gloria, generalizes and simplifies previous results by Gustafsson, Kreiss, Tadmor, Wu and others. In collaboration with O. Guès and M. Williams, J.-F. Coulombel has also studied the justification of geometric optics for hyperbolic boundary value problems. The results describe the reflection of highly oscillating wave trains on a boundary. Eventually, J.-F. Coulombel has studied with S. Benzoni and N. Tzvetkov well-posedness issues for some nonlocal versions of Burgers equation.

5.3. Control in Fluid Mechanics

Results on active and passive control on the 3D Ahmed body in the 25° rear-window configuration were obtained in [4], [5], [21]. Moreover, results on the pressure wave generated by high-speed trains entering tunnels were published in [20].

5.4. Fluid/Particles Flows

We are interested in two-phase flows involving a dense and a disperse phase. These models lead to interesting mathematical questions, [38]. We develop new asymptotic preserving methods for fluid/particles flows [50]. This approach follows the scheme we developed for radiative transfer equations [51].
5.5. Plasma Physics

We obtained several results of asymptotic analysis concerning either kinetic or macroscopic models for charge transport, see [45], [44] and [52] (and we also refer to the related work [53]). Through the collaboration with Thales we proposed in [3] a method to compute numerically the boundary conditions for the Euler system derived from the BGK equation in the hydrodynamic limit. In particular this method is based on an analysis of boundary layers. In [7] we extend these techniques to the coupled Euler Poisson system.

5.6. Analysis and numerical simulation of the Schrödinger equation

The linear or nonlinear Schrödinger equation with potential is one of the basic equations of quantum mechanics and it arises in many areas of physical and technological interest, e.g. in quantum semiconductors, in electromagnetic wave propagation, and in seismic migration. The Schrödinger equation is the lowest order one-way approximation (paraxial wave equation) to the Helmholtz equation and is called Fresnel equation in optics, or standard parabolic equation in underwater acoustics. The solution of the equation is defined on an unbounded domain. If one wants to solve such a whole space evolution problem numerically, one has to restrict the computational domain by introducing artificial boundary conditions. So, the objective is to approximate the exact solution of the whole-space problem, restricted to a finite computational domain. A review article [39] was written this year to describe and compare the different current approaches of constructing and discretizing the transparent boundary conditions in one and two dimensions. However, these approaches are limited to the linear case (or nonlinear with the classical cubic nonlinearity: an article written was dedicated to this case this year [41]) and constant potentials. Therefore, in collaboration with X. Antoine (IECN Nancy and Inria Lorraine), we proposed to P. Klein to study, in her PhD thesis, the case of the Schrödinger equation with variable potentials. The study of the non-stationary one-dimensional case has already led to one publication [40] and some preliminary results in the stationary case are really promising. These cases are relevant since for example the equations appear in the Bose Einstein condensate with a quadratic potential.

This problem is obviously not limited to the Schrödinger equation and new developments are in progress on the Korteweg de Vries equation with M. Ehrhardt. This equation is more difficult to study due to its third order derivative in space.

Dispersive equations, such as the Schrödinger equation are also considered as boundary-value problems. For example, in [57], G.D. studies the long time asymptotics of the solutions of linear Schrodinger equations considered as initial-boundary value problems on the half-line and on bounded intervals when the boundary data are periodic functions of time. G.D. obtains theoretical results using a transformation method introduced by T. Fokas and provides several numerical experiments to support them.

The Diffusion Monte-Carlo method is a powerful strategy used by chemists to estimate the groundstate energy of a N-body Schrödinger Hamiltonian with high accuracy. However, the method suffers from two major limitations:

- The quantity of physical interest is more the energy variation with respect to a parameter than the energy itself.
- The case of Fermions (as electrons of atoms) relies on constraining the random walkers in some nodal pocket (which imposes skew-symmetry on the distribution of walkers) which is only approximately known (the Fixed Node Approximation).

In the article [63], Mathias Rousset proposed a new strategy for the case of Fermions to compute the energy variation with respect to a variation of the nodal pocket. The ultimate goal is to design a Monte-Carlo strategy able to optimize the Fixed Node Approximation.

5.7. Homogenization

We have three types of results regarding the homogenization theory and its applications.
The first series of results is related to nonlinear elasticity. In [2], A. Gloria has proved the convergence of a discrete model for rubber towards a nonlinear elasticity theory in collaboration with R. Alicandro and M. Cicalese. The numerical simulation of the model has been addressed within the ARC Disco by A. Gloria, P. La Tallec and M. Vidrascu (project team MACS). Comparisons with mechanical experiments are promising, and related inverse problems have been addressed in the post-doc of M. de Buhan. Two publications are in preparation. Related theoretical results in homogenization of nonlinear elasticity models have been obtained by A. Gloria and S. Neukamm (MPI Leipzig) in [15].

A second type of results concerns a quantitative theory of stochastic homogenization of discrete linear elliptic equations. A breakthrough has been obtained by A. Gloria and F. Otto (MPI Leipzig) in [16] and [17], who gave the first optimal variance estimate of the energy density of the corrector field for stochastic discrete elliptic equations. The proof makes extensive use of a spectral gap estimate and of deep elliptic regularity theory, bringing in fact the probabilistic arguments to a minimum. This analysis has enabled A. Gloria to propose efficient numerical homogenization methods, both in the discrete and continuous settings [13], [12]. In [14], A. Gloria and J. C. Mourrat has pushed the approach forward and introduced new approximation formulas for the homogenized coefficient. In [26] they have considered a more probabilistic approach and given a complete error analysis of a Monte-Carlo approximation of the homogenized coefficients in the discrete case. Work in progress concerns the generalization of the results on discrete elliptic equations to the continuous case.

The third direction of research concerns the periodic homogenization of a coupled elliptic/parabolic system arising in the modeling of nuclear waste storage. This work is in collaboration with the French agency ANDRA. A. Gloria, T. Goudon, and S. Krell have made a complete theoretical analysis of the problem, derived effective equations, and devised an efficient method to solve the effective problem numerically, based on the reduced basis approach. A publication is in preparation. This subject will be continued with the arrival of Z. Habibi.

5.8. Radiative Transfert

The interest of the team in developing efficient numerical methods preserving the asymptotic behavior of kinetic equations modelling the radiative transfer phenomena in the diffusive regime has recently brought two new innovative schemes. On the one hand, the moment closure equations proposed in [43] involve non local terms that lead to the introduction of specific numerical approximations. On the other hand, a new scheme [19] based on the projective integration procedure due to Gear and Kevrekidis was proved very efficient theoretically and numerically.

The team has also obtained a rigorous justification of the formation of so-called Zeldovich spikes in radiative shock profiles [32]. As long as the amplitude of the wave is sufficiently large, the temperature profile displays an overshoot. Such phenomenon does not occur for small amplitude waves, nor for viscous shock profiles.

5.9. Complex fluid flows

A numerical treatment [18] of phase transitions arising in the modelling of the behavior of polymers near glass transition by a non linear diffusion equation was conducted in the continuation of the theoretical results by Evans and Portilheiro and Mascia, Terracina and Tesei.

5.10. A posteriori error estimators for finite element methods

First, the generalization of [55] in the case of the full convection-reaction-diffusion case lead to robust estimators, for which the dependance of the constant in the data are explicitly given [8]. Then, we succeeded in obtaining some “asymptotic exact estimators” for the Reissner-Mindlin system in order to provide an accurate bound of the error, while keeping a reasonable computational cost [10], [24]. Finally, we obtained some residual-based a posteriori error estimators for the Maxwell system in its vectorial and scalar potential formulation [9], [23].
5. New Results

5.1. Mathematical Modelling

5.1.1. Mathematical and numerical model for nonlinear viscoplasticity

Participants: Nicolas Favrie, Sergey Gavrilyuk.

A macroscopic model describing elastic plastic solids is derived in a special case of the internal specific energy taken in separable form: it is the sum of a hydrodynamic part depending only on the density and entropy, and a shear part depending on other invariants of the Finger tensor. In particular, the relaxation terms are constructed compatible with the von Mises yield criteria. In addition, Maxwell-type material behaviour is shown up: the deviatoric part of the stress tensor decays during plastic deformations. Numerical examples show the ability of this model to deal with real physical phenomena [15].

5.1.2. A discrete model for compressible flows in heterogeneous media

Participants: Olivier Le Métayer, Alexandre Massol, Nicolas Favrie, Sarah Hank.

This work deals with the building of a discrete model able to describe and to predict the evolution of complex gas flows in heterogeneous media. In many physical applications, large scales numerical simulation is no longer possible because of a lack of computing resources. Indeed the medium topology may be complex due to the presence of many obstacles (walls, pipes, equipments, geometric singularities etc.). Aircraft powerplant compartments are examples where topology is complex due to the presence of pipes, ducts, coolers and other equipment. Other important examples are gas explosions and large scale dispersion of hazardous materials in urban places, cities or underground involving obstacles such as buildings and various infrastructures. In all cases efficient safety responses are required. Then a new discrete model is built and solved in reasonable execution times for large cell volumes including such obstacles. Quantitative comparisons between experimental and numerical results are shown for different significant test cases, showing excellent agreement [18].

5.1.3. A hyperbolic Eulerian model for dilute two-phase suspensions

Participants: Sarah Hank, Richard Saurel, Olivier Le Métayer.

Conventional modeling of two-phase dilute suspensions is achieved with the Euler equations for the gas phase and gas dynamics pressureless equations for the dispersed phase, the two systems being coupled by various relaxation terms. The gas phase equations form a hyperbolic system but the particle phase corresponds to a hyperbolic degenerated one. Numerical difficulties are thus present when dealing with the dilute phase system. In the present work, we consider the addition of turbulent effects in both phases in a thermodynamically consistent way. It results in two strictly hyperbolic systems describing phase’s dynamics. Another important feature is that the new model has improved physical capabilities. It is able, for example, to predict particle dispersion, while the conventional approach fails. These features are highlighted on several test problems involving particles jets dispersion and are compared against experimental data. With the help of a single parameter (a turbulent viscosity), excellent agreement is obtained for various experimental configurations studied by different authors [17].

5.1.4. Diffuse interface model for compressible fluid - Compressible elastic-plastic solid interaction

Participants: Nicolas Favrie, Sergey Gavrilyuk.
An Eulerian hyperbolic diffuse interface model for elastic plastic solid fluid interaction is constructed. The system of governing equations couples Euler equations of compressible fluids and a visco-plastic model of Maxwell type materials (the deviatoric part of the stress tensor decreases during plastic deformations) in the same manner as models of multicomponent fluids. In particular, the model is able to create interfaces which were not present initially.

The model is thermodynamically compatible: it verifies the entropy inequality. However, a numerical treatment of the model is particularly challenging. Indeed, the model is non-conservative, so a special numerical splitting is proposed to overcome this difficulty. The numerical algorithm contains two relaxation procedures. One of them is physical and is related to the plastic relaxation mechanism (relaxation toward the yield surface). The second one is numerical. It consists in replacing the algebraic equation expressing a mechanical equilibrium between components by a partial differential equation with a short relaxation time. The numerical method was tested in one dimensional case (Wilkins’ flying plate problem), two-dimensional plane case (impact of a projectile on a plate) and axisymmetrical case (Taylor test problem, impact with penetration effects, etc.). Numerical examples show the ability of the model to deal with real physical phenomena [13].

5.1.5. Criterion of hyperbolicity for non-conservative quasilinear systems admitting a partially convex conservation law

Participants: Alain Forestier, Sergey Gavrilyuk.

A system of conservation laws admitting an additional convex conservation law can be written as a symmetric $t$-hyperbolic in the sense of Friedrichs system. However, in mathematical modeling of complex physical phenomena, it is customary to use non-conservative hyperbolic models. We generalize the Godunov Friedrichs Lax approach to this new class of models [16].

5.1.6. A new model of roll waves: comparison with Brock’s experiments

Participants: Gaël Richard, Sergey Gavrilyuk.

We derive a mathematical model of shear flows of shallow water down an inclined plane. Periodic stationary solutions to this model describing roll waves were obtained. The solutions are in good agreement with experimental profiles of roll waves measured in Brock’s experiments (1967). In particular, the height of the vertical front of the waves, the shock thickness and the wave amplitude are well captured by the model [21].

5.1.7. Modelling gas dynamics in one-dimensional ducts with abrupt area change

Participants: R Menina, Richard Saurel, M Zereg, Lazhar Houas.

Most gas dynamic computations in industrial ducts are done in one dimension with cross-section-averaged Euler equations. This poses a fundamental difficulty as soon as geometrical discontinuities are present. The momentum equation contains a non-conservative term involving a surface pressure integral, responsible for momentum loss. Definition of this integral is very difficult from a mathematical standpoint as the flow may contain other discontinuities (shocks, contact discontinuities). From a physical standpoint, geometrical discontinuities induce multidimensional vortices that modify the surface pressure integral. In the present paper, an improved one-dimensional flow model is proposed. An extra energy (or entropy) equation is added to the Euler equations expressing the energy and turbulent pressure stored in the vortices generated by the abrupt area variation. The turbulent energy created by the flow area change interaction is determined by a specific estimate of the surface pressure integral. Model’s predictions are compared with two-dimensional averaged results from numerical solution of the Euler equations. Comparison with shock tube experiments is also presented. The new one dimensional averaged model improves the conventional cross-section-averaged Euler equations and is able to reproduce the main flow features [19].

5.2. Applications for specific flow problems

5.2.1. Modelling cavitating flow around underwater missiles

The diffuse interface model of Saurel et al. [9] is used for the computation of compressible cavitating flows around underwater missiles. Such systems use gas injection and natural cavitation to reduce drag effects. Consequently material interfaces appear, separating liquid and gas. These interfaces may have a really complex dynamics such that only a few formulations are able to predict their evolution. Contrarily to front tracking or interface reconstruction method the interfaces are computed as diffused numerical zones, that are captured in a routinely manner, as is done usually with gas dynamics solvers for shocks and contact discontinuity. With the present approach, a single set of partial differential equations is solved everywhere, with a single numerical scheme. This leads to very efficient solvers. The algorithm derived in Saurel et al. [43] is used to compute cavitation pockets around solid bodies. It is first validated against experiments done in cavitation tunnel at CNU. Then it is used to compute flows around high speed underwater systems (Shkval-like missile). Performance data are then computed showing method ability to predict forces acting on the system [20].

5.2.2. Propagation of a planar shock wave through a two-phase gas-liquid medium

Participants: Alain Chauvin, Georges Jourdan, Éric Daniel, Lazhar Houas, R Tosello.

We conducted a series of shock tube experiments to study the influence of a cloud of water droplets on the propagation of a planar shock wave. In a vertically oriented shock tube, the cloud of droplets was released downwards into the air at atmospheric pressure while the shock wave propagated upwards. Two shock wave Mach numbers, 1.3 and 1.5, and three different heights of clouds, 150 mm, 400 mm, and 700 mm, were tested with an air-water volume fraction and a droplet diameter fixed at 1.2 % and 500 µm, respectively. From high-speed visualization and pressure measurements, we analyzed the effect of water clouds on the propagation of the shock wave. It was shown that the pressure histories recorded in the two-phase gas-liquid mixture are different from those previously obtained in the gas-solid case. This different behavior is attributed to the process of atomization of the droplets, which is absent in the gas-solid medium. Finally, it was observed that the shock wave attenuation was dependent on the exchange surface crossed by the shock combined with the breakup criterion [12].
6. New Results

6.1. Autonomous Computing

Participants: Cécile Germain-Renaud, Michèle Sebag, Balázs Kégl, Yusik Kim, Julien Nauroy, Dawei Feng.

Within the classical objectives of Autonomics (self-*), two transversal lines of research have emerged.

Coping with non-stationarity Most existing work on modeling the dynamics of grid behavior assumes a steady-state system and concludes to some form of long-range dependence (slowly decaying correlation) in the associated time-series. But the physical (economic and sociologic) processes governing the grid behavior dispel the stationarity hypothesis. [68] proposes a categorization of the methods integrate non-stationarity into grid modeling. [9] considers a specific class of models: a sequence of stationary processes separated by breakpoints. The model selection question is now defined as identifying the breakpoints and fitting the processes in each segment, together with a validation methodology that empirically addresses the current lack of theoretical results concerning the quality of the estimated model parameters. Even when stationarity is acceptable, the markovian assumption might be too bold. [54] integrate Echo State Network-based regression into a reinforcement learning in continuous state space for fitting the Q function, with application to reactive grid scheduling.

Traces mining In order for an autonomic system to continuously infer knowledge from its monitoring (the so-called MAPE-K, Monitor-Analyze-Plan-Execute-Knowledge) loop, heterogeneous sources of data have to be integrated. [37] exemplifies two use cases of the Grid Observatory data on evaluating the performance of the major EGI scheduler, and blackhole detection.

The Green Computing Observatory [38] data include the detailed monitoring of the processors and motherboards, as well as the global site information, such as overall consumption and overall cooling. The data schema for publication is grounded in an ontology of measurements developed in collaboration with the MIS (Modélisation, Information et Systèmes) laboratory of University Picardie Jules Verne.

[42] proposes a new approach for analyzing behavioral traces: as most of them are indeed text documents, state of the art techniques in text mining, including Latent Dirichlet Allocation, can be exploited. The advantages are twofold: providing some level of explanation inferred from the data; and a relatively scalable way to capture the temporal variability of the behavior of interest, while retaining the full dimensionality of the problem at hand. A promising perspective for combining this approach and inferred segmentation has been identified and is currently explored.

6.2. Complex Systems

Participants: Jamal Atif, Nicolas Bredèche, Matthias Brendel, Cyril Furtlehner, Philippe Caillou, Jean-Marc Montanier, Hélène Paugam-Moisy, Marc Schoenauer, Michèle Sebag.

Evolutionary AI Planning: Divide And Evolve (DAE) DAE solves AI-planning problems by using an Evolutionary Algorithm to sequentially divide them into hopefully simpler problems that are handled by some embedded “classical” planner. Within the ANR project DESCARWIN, work has continued in collaboration with Thalès Research & Technology and ONERA Toulouse. A large part of the work this year has been devoted to writing a brand new version of the DAE software, facing difficulties of parallelization [90]. The resulting program entered the 7th International Planning Competition (IPC 2011) at the 21st International Conference on Planning and Scheduling (ICAPS 2011) and won the Gold Medal in the Temporal Track. Note that the Silver Medal was won by Vincent Vidal,
also member of the DESCARWIN team, using his planner YAHSP2 – the one that won the Gold Medal while embedded in DAE, thus demonstrating one more the added value of the DAE approach. Meanwhile, because DAE has many parameters (like most Evolutionary Algorithms), parameter tuning within DAE remains a difficult task, and an original approach has been proposed to learn the parameters based on some instance features [49], [50], [51]. Note that this method is however relevant of the “Crossing the Chasm” SIG (see Section 6.3), as it can be applied to any optimization algorithm that handles several instances of the same class.

Distributed Autonomous Robotics Resuming work done in 2010, we investigated further the issue of robotic swarm control whenever the environment is partially or completely unknown. This research is at the cross-road of Evolutionary Computation, Machine Learning and Robotics, and a light influence from Evolutionary Ecology, but with a strong focus on engineering (i.e. the goal remains to design algorithms). The topic we are interested in is the design of environment-driven self-adaptive distributed algorithms to enable survival at the level of a population of independent robotic units. The population is limited in size, and hardware implementation within real robots has already been achieved [7]. We have also focused our attention on specific aspects of swarm evolutionary dynamics under specific constraints, including the evolution of cooperative and/or altruistic behaviours [53], [52]. This research yielded interesting results, such as the emergence of altruistic behavior under simple, but specific, algorithmic constraint, as well as tuning mechanism to control the level of altruistic behavior in a population of robots. Perspectives of this work is currently under investigation.

The work done in 2010 about the division of labor among asynchronous and decentralized agents, where each agent is modelled from the competition between two spiking neurons, was further analyzed within a spatio-temporal (simulated) frame. The phase transitions between the asynchronous, the aperiodic and periodic synchronous regimes (depending on the sociability and excitability of the agents) was confirmed, with some counter-intuitive results about the overall merits and efficiency of synchronous behaviors [23].

We have also explored objective-driven online learning within real robotic hardware, both for single robot online behavior learning [78] as well as small group of robots for pattern formation learning [43]. Our activity in Evolutionary Robotics has also been strengthened by the publication of book which gather several contributions from major actors in the field [75], including an introduction paper on current trends and challenges in this domain [76].

From a slightly different perspective, our work on evolving generative and developmental representations has been continued, with an extensive study of robustness within developmental systems [8] and an investigation of the temporal dynamics at work within genetic regulatory networks for design [32]. While not strictly related to robotics, these contributions share the distributed nature of computation and ultimately aim at providing an efficient representation for designing and controlling large scale passive or active assembly of units (e.g. robots with complex morphologies).

Additionally, at the crossroad of Machine Learning and Evolutionary Computation, a new Reinforcement Learning approach based on modelling the user’s preferences was proposed [12], [84]; in the so-called Preference-based Policy Learning, the robot demonstrates some behaviors, is informed of the user’s preferences, builds a model of the user’s preferences and self-trains to build a new behavior hopefully more satisfactory according to the conjectured user’s preferences.

Statistical Physics Perspective Basic tools from statistical physics (scaling, mean-field techniques and associated distributed algorithms, exactly-solvable models) and probability have been used to model and optimize complex systems, either standalone or combined with MABS approaches. Results are

- In the context of the ANR TRAVESTI project dealing with spatial and temporal modelling of traffic congestion we have studied in [83] some specific properties of the Belief Propagation algorithm used for inference; we have proposed a way to encode dependencies between real variables with a latent binary MRF [89]; we have analyzed macro-states on traffic data and how these relate to belief propagation fixed points [35].
• Also in the ANR Travesti context [36], [88] for modelling congestion at the microscopic level we have proposed a new family of queueing processes where the service rate is coupled stochastically to the number of clients. With this formulation we have been able to relate an asymmetry between acceleration and braking to some condensation mechanism. In this framework we have also proposed a large deviation formulation of the fundamental diagram of traffic flow.

• In the design of multi-objective message passing algorithms we have shown how the state-of-the-art MAXSAT solver SP-Y can be used directly by an endogeneous clause elimination procedure, to sample Pareto Front of multi-objective 3-SAT problems [87].

Multi-agent and games • Within the InnovNation serious game project, in collaboration with Paraschool and BlueNove, we developed an e-brainstorming prototype for collaborative ideation. The prototype was tested on 150 students and is being improved to be commercialized. The realist network generator used to analyze the prototype games was also used to study the labor market, and to study the relative importance of friends and colleagues while seeking a job without ([44] or with [64] variable information transmission speed. We have shown that friends were the most useful when the labor market was at the equilibrium (approximately the same number of jobs and applicants).

• To analyze the logs of multi-agent based simulations (for example for the InnovNation project), we developed a tool to describe homogeneous agent clusters and their evolution ([73]). We use the cluster description to build agent models and generate new simulations with this model to validate the results. • Also in the complex social system analysis context, we applied LSA text mining tools on research projects and patent category descriptions to associate research clusters to their main research fields [70]. This was used to build a classification of the French research clusters based on their context, and especially the adequation of the cluster research specialization with the regional specialization [72].

Image understanding • Within the context of image understanding, a new sequential recognition framework has been proposed in [10]. Sequential image understanding refers to the decision making paradigm where objects in an image are successively segmented/recognized following a predefined strategy. Such an approach generally raises some questions about the “best” segmentation sequence to follow and/or how to avoid error propagation. In [10], we propose original approaches to answer these questions in the case where the objects to segment/recognize are represented by a model describing the spatial relations between objects. The process is guided by a criterion derived from visual attention, and more precisely from a saliency map, along with some spatial information to focus the attention. This criterion is used to optimize the segmentation sequence. Spatial knowledge is also used to ensure the consistency of the results and to allow backtracking on the segmentation order if needed. The proposed approach was applied for the segmentation of internal brain structures in magnetic resonance images. The results show the relevance of the optimization criteria and the interest of the backtracking procedure to guarantee good and consistent results. From a logical standpoint, sequential object recognition is formulated as an abduction process in [14], [66]. A scene is viewed as an observation and the task of interpretation is considered as the “best” explanation considering the prior knowledge about the scene context. Towards this aim, we introduce an algebraic-based framework unifying mathematical morphology, description logics and formal concept analysis. We propose to compute the best explanations of an observation through algebraic erosion over the Concept Lattice of a background theory which is efficiently constructed using tools from Formal Concept Analysis. We show that the defined operators are sound and complete and satisfy important rationality postulates of abductive reasoning.

6.3. Crossing the Chasm

Participants: Alejandro Arbelaez, Anne Auger, Robert Busa-Fekete, Nikolaus Hansen, Balázs Kégl, Manuel Loth, Nadjib Lazaar, Marc Schoenauer, Michèle Sebag.
Due to the departure of both PhD students funded within the Microsoft-INRIA joint lab after their successful defenses (Alvaro Fialho and Alejandro Arbelaez), some of the activities of this SIG have been slightly redefined this year, with the one-month visit of Prof. Th. Runarsson (University of Iceland) in October, and the arrival in November of two new post-docs, also funded by the joint lab (Nadjib Lazaar and Manuel Loth). A new direction of research has appeared, in line with both Adaptive Operator Selection (Alvaro Fialho’s PhD) and Continuous Search (Alejandro Arbelaez’ PhD).

Bandit-based choice of heuristics in combinatorial optimization This new direction of research deals with heuristic choice within an existing combinatorial solver using bandit-like algorithms, and the very first results deal with scheduling problems and will be published in early 2012 [57].

Adaptive Operator Selection In line with his PhD work, Alvaro Fialho has successfully used his Adaptive Operator Selection method to the on-line tuning of Differential Evolution in the multi-objective case [96].

Learn and Optimize (LaO), an instance-based parameter-tuning method. Though originally designed for Divide-And-Evolve framework (see Section 6.2), LaO is a generic method that learns the relationship between some instance features and the optimal parameters of the optimizer. The current version [49], [50], [51] uses Neural Network to directly learn the optimal parameters, and average performance increase compared to the default parameter set (that has won the temporal track in the IPC7 competition) is of more than 10%. On-going work uses rankSVM to learn a partial order on the features × parameter space.

Adaptive Constraint Programming Alejandro Arbelaez defended his PhD on May 31., led under the supervision of Youssef Hamadi and Michèle Sebag [1]. A survey of his PhD work has been published as a book chapter [74] and some of his last work was more concerned with optimizing the collaboration in distributed SAT solving in highly parallel environments [13].

Ranking by calibrated AdaBoost In [22], [21] we describe a learning-to-rank technique based on calibrated multi-class classification. We train a set of multi-class classifiers using AdaBoost.MH, we calibrate them using various techniques to obtain diverse class probability estimates, and, finally, we approximate the Bayes-scoring function (which optimizes the popular Information Retrieval performance measure NDCG), through mixing these estimates into an ultimate scoring function. Our method outperforms many standard ranking algorithms on the LETOR benchmark datasets, most of which are based on significantly more complex learning to rank algorithms than ours.

6.4. Continuous Optimization

Participants: Yohei Akimoto, Anne Auger, Zayed Bouzarkouna, Alexandre Chotard, Nikolaus Hansen, Ilya Loshchilov, Verena Heidrich-Meisner, Raymond Ros, Marc Schoenauer, Olivier Teytaud, Fabien Teynaud.

Our main expertise in continuous optimization is on stochastic search algorithms. We address theory, algorithm design and applications. The methods we investigate are adaptive techniques able to learn iteratively parameters of the distribution used to sample solutions. The Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is nowadays one of the most powerful method for continuous optimization without derivatives. We work on different variants of the CMA-ES to improve it in various contexts as described below. In addition we have contributed to give an information geometry perspective to stochastic optimization unifying both continuous and discrete algorithms using a family of probability distribution parametrized by continuous parameters. The framework proposed in this context allow to retrieve many existing stochastic optimization algorithms when instantiated with different family of probability distributions including the CMA-ES when using gaussian distributions [85]. We have moreover clarified important design principles based on invariances [85], [11].

New algorithms based on derandomization and for mixed-integer optimization A new variant of CMA-ES to address problem with mixed-integer variables (vectors with both discrete and continuous variables) has been designed [81]. New algorithms using new selection schemes combined with
derandomization have been designed and thoroughly theoretically and empirically investigated [16], [17]. A local search algorithm using an adaptive coordinate descent has been proposed [45]. We have as well investigated how to inject solutions in CMA-ES so as to improve performances if an oracle provide good solutions [82].

Distributed optimization We have proposed simple modifications of evolutionary algorithms so that they reach asymptotically the optimal \( \log(\lambda) \) speed-up with \( \lambda \) processors and the linear speed-up \( \Theta(\lambda) \) for a number of processors of order at most the dimensionality of the problem (for a pointwise solution); in particular bounding the selected population size in the Self-Adaptation algorithm and variants of this idea for fastening the decrease of the step-size in a relevant manner. All these works and more are part of Olivier Teytaud’s HDR thesis [3], defended on April 22., and also build the first part of Fabien Teytaud’s PhD [2], defended on December 8. We wrote a chapter on lower bounds for distributed derivative free optimization in [79].

Benchmarking We have continued our effort for improving standards in benchmarking and pursued the development of the COCO - COmparing Continuous Optimizers platform [80]. We are organizing for the GECCO 2012 conference the Black-Box-Optimization Workshop4.

Optimization with meta-models and surrogate We have investigated optimization using a coupling of CMA-ES and surrogates and applied it for the optimization of well placement [6]. We have proposed a new meta-model CMA-ES for the optimization of partially separable functions [19] and shown that it improves performances for solving the well placement problem [19].

Hyperparameter optimization In [18] we present hyper-parameter optimization results on tasks of training neural networks and deep belief networks (DBNs). We optimize hyper-parameters using random search and two new greedy sequential methods based on the expected improvement criterion. The sequential algorithms are applied to the most difficult DBN learning problems and find significantly better results than the best previously reported.

Multi-objective optimization We have investigated theoretically multi-objective algorithms based on the hypervolume [4] and proposed new selection operators based on tournament and multi-armed bandit framework [46].

Multimodal optimization We have shown in [58] a simple algorithm (a \((1 + 1)\)-ES with quasi-random restart and murder operator), which, at least in the sequential case, performs as efficiently as much more tricky algorithms.

Mathematical bounds for noisy optimization The paper [27] shows upper and lower confidence bounds and/or experiment algorithms in the noisy optimization setting; in particular we compared an optimization algorithm based on bandits and an surrogate-model version; whereas the bandit approach is much faster if the noise decreases quickly to zero around the optimum, the surrogate-model version is faster if the noise does not decrease to zero.

6.5. Optimal Decision Making

Participants: Olivier Teytaud, David Auger, Michèle Sebag, Cyril Furtlehner, Jean-Baptiste Hoock, Nataliya Sokolovska, Fabien Teytaud, Hassen Doghmen, Jean-Joseph Christophe, Jérémie Decock.

- Monte-Carlo Tree Search (MCTS) and Upper Confidence Trees (UCT) are main areas of the team. In particular, we ultra-weakly solved 7x7 Go by winning 20 games out of 20 against professional players in 7x7 Go, thanks to a Meta-Monte-Carlo-Tree Search[24]. The wins were with komi 9.5 as white, and 8.5 as black, suggesting that the ideal komi in 7x7 is 9. We also applied this algorithm to the recent “NoGo” framework, aimed at challenging MCTS for a game which looks like Go but with very different goals; our paper [26] was the first one applying MCTS to NoGo and now all strong programs use the MCTS approach for NoGo. We extended RAVE (Rapid Action Value Estimates) to the continuous settings [29]. In his PhD [2], Fabien Teytaud proposed several generic

4see http://coco.gforge.inria.fr/doku.php?id=bbob-2012
improvements of MCTS, including the use of (fast) decisive and anti-decisive moves for games, and applied it to the game of Havannah. An industrial application (to energy management) is proposed in [71]. A MCTS version for partially observable problems with bounded horizon was proposed in [86]. This version is proposed for the two-player case, but for simulations starting at the root; a version in the one-player case, starting from an arbitrary state (and therefore much more efficient for large horizon) is proposed in [30]. This work is extended by a belief state estimation by constraint satisfaction problems in [62]. Other developments and research around MCTS/UCT are described in the MoGo module.

- A related important algorithm is Nested Monte-Carlo; we got state of the art results for some traveling salesman variants with a very simple algorithm in [56].
- Fundamental analysis of partially observable games: we proved in [5] that partially observable games are undecidable (result also presented in the BIRS 2010 workshop and the Bielefeld seminar on Search Methodologies), even in the case of finite state spaces and deterministic transitions. This unexpected result is a priori a contradiction with known decidability results; this emphasizes the subtle difference between the classical decision problem (the existence of a strategy winning certainly, whatever may be the strategy of the opponent), which is used is most analysis, and the choice of the move with optimal winning probability. We pointed out that the relevant decision problem is, with no doubt, the latter; that the other decision problem has just been used because it is equivalent to choosing optimal play in the case of fully observable games; and, most importantly, that partially observable games are in fact undecidable, even in the finite deterministic case. On the other hand, on restricted settings, we have shown by some simple lemmas lower and upper bounds on the value of some partially observable games [63]. We extended Monte-Carlo Tree Search to the case of short-term partial information in [61]; this was successfully applied to the Urban Rival game, a widely played internet card game (now 17 millions of registered users) from a French company.
- Tuning of strategies: tuning strategies is a noisy optimization problem in which the convenient “variance of noise decreasing to zero around the optimum” usually does not hold. We have shown that in such a setting, the local bandit-style algorithms are slower than surrogate models; this is detailed in the continuous optimization part.
- We organized various computer-Go events, as due to the fame of our program MoGo we are often invited for such events; reports can be found in [95].
- We developed the “double progressive widening” trick, which is aimed at making consistent an algorithm from the finite case to the continuous stochastic case; we got good results in [60] on Q-Learning (with no mathematical proof) and on MCTS [28] (mathematical proof to be submitted soon).
- We have also worked on Nash equilibria of Matrix Games, where we proposed an algorithm for finding Nash equilibria faster when the Nash equilibrium is sparse [48], [47]; a mathematical proof is ready and will be submitted soon.
- Some works are in progress around applications of previous tools to active learning; active learning has also been investigated through conditional random fields in [59].
- Another related work, with motivations from autonomous robotics, combines the exploration of the search space through UCT, with an explicit model of the safe regions explored so far, called Deja-Vu. The Deja-Vu is used to constrain the exploration, mostly in the random phase, and is updated from the current explorations [67].
- The Ilab “Metis” just started; it’s an Ilab between Tao, the Inria-Saclay team MaxPlus, and the SME Artelys http://www.artelys.com for a joint work on numerical libraries in Energy Management.

6.6. Large and Deep Networks

Participants: Ludovic Arnold, Sylvain Chevallier, Anthony Mouraud, Hélène Paugam-Moisy, Sébastien Rebecchi, Michèle Sebag.
Deep networks: RBM or AA  The two main families of deep networks are implemented and studied by TAO: stacked RBM (Restricted Boltzmann Machines) and stacked AA (Auto-Associators).

Learning sparse features for deep networks  Inspired by the theory of compressed sensing and beyond the common methods based on dictionary learning, we have proposed to learn sparsity and accuracy simultaneously by alternating two constraints on the weights of an Auto-Associator [55].

Spiking neuron networks for swarm robotics  The model "SpikeAnts" [91] has been applied to a spatial robotic environment [23], in collaboration with Nicolas Bredeche (see Section 6.2), and has demonstrated even more its interest in the context of swarm robotics.
6. New Results

6.1. Probabilistic numerical methods, stochastic modelling and applications

Participants: Mireille Bossy, Nicolas Champagnat, Julia Charrier, Julien Clairisse, Madalina Deaconu, Samuel Herrmann, James Inglis, Pierre-Emmanuel Jabin, Antoine Lejay, Sylvain Maire, Sebastian Niklitschek Soto, Nicolas Perrin, Denis Talay, Etienne Tanré, Laurent Violeau.

6.1.1. Published works and preprints

- M. Bossy in collaboration with J.-F. Jabir (Univ. Chile) proved the well posedness of the confined Lagrangian models, in association with no-permeability boundary conditions. When the confining domain is a hyperplane, they proved the strong existence of the trace of the density of particles following the kinetic stochastic equation of a simplified McKean Vlasov Lagrangian model in [12], http://hal.inria.fr/inria-00515481/en. When the confining domain $D$ is bounded with smooth boundary, they constructed a confined primitive of Brownian motion in $D$ and characterized the solution to the corresponding martingale problem by showing that the time marginal density is the unique solution to a mild equation with specular condition. This key step allowed them to finish the construction in the non linear case, using previous work on Vlasov-Fokker-Plank PDE with specular boundary condition. Two papers are being written.

- In collaboration with J.-F. Jabir and J. Fontbona (CMM and Universidad de Chile, Santiago de Chile), M. Bossy and P.-E. Jabin have studied the link between the Lagrangian version of divergence free constraint (and the uniform density constraint), with an additional potential term, in the Lagrangian equation, having some similarity with the role of the Eulerian pressure term. They obtained the local existence of analytical solutions for an incompressible Lagrangian stochastic model in periodic domain. An article is currently being written.

- N. Champagnat worked with A. Lambert (Univ. Paris 6) on splitting trees with Poissonian mutations. Assuming that each mutation is neutral and gives a new type in the population, they obtained in [15] explicit expressions for the expected number of types carried by a fixed number of individuals living in the population at time $t$. In [31], they also obtained large time convergence results on the sizes of the largest families and the ages of the oldest families in the population. http://hal.inria.fr/inria-00515481/en, http://hal.inria.fr/inria-00616765/en.

- N. Champagnat and P.-E. Jabin studied the limit of some population dynamics models under the assumption that the time scale for mutations is much larger than the time scale for reproduction. They are able to provide the first full characterization of the corresponding limit equation [14], http://hal.inria.fr/inria-00488979/en.

- M. Deaconu and S. Herrmann developed a new method for the simulation of the hitting time of nonlinear boundaries for Bessel processes. This method is based on a walk on moving spheres algorithm and can be applied for the hitting time of a given level for the Cox-Ingersoll-Ross process [32], http://hal.inria.fr/hal-00636056/en. This work is part of the ANR MANDy project.

- S. Herrmann and E. Tanré worked on a scheme to construct an efficient algorithm to simulate the first hitting time of curves by a one dimensional Brownian motion. They apply the result to estimate the spiking time of leaky integrate fire models in neurosciences. This work is part of the ANR MANDy project.
P.-E. Jabin and F. Ben Belgacem (Univ. of Monastir, Tunisia) have studied a new class of models which have seen considerable development in applications for biosciences (flocking, chemotaxis, pedestrian flows...). These models include some non linear corrections to classical linear continuity equations. In [30], they introduce new, critical regularity estimates to obtain well posedness. http://www2.cscamm.umd.edu/~jabin/transportlcs2.pdf.

P.-E. Jabin and M. Hauray (Aix-Marseille Université) have studied the mean field limit for systems of many interacting particles. It is the only result able to deal with singular forces and physically realistic initial configurations [33]. http://hal.inria.fr/hal-00609453/en.

P.-E. Jabin and A. Nouri (Aix-Marseille Université) studied a highly singular kinetic equation in dimension 1. This equation is obtained as a quasi-neutral limit in plasma physics. In [18], they were able to prove well posedness in short time of analytic solutions. http://dx.doi.org/10.1016/j.crma.2011.03.024.

P.-E. Jabin and G. Raoul (Cambridge University) prove the convergence to a unique stable equilibrium for a wide class of competitive models in population dynamics [19]. http://dx.doi.org/10.1007/s00285-010-0370-8.

P.-E. Jabin and J. Calvo (Universidad de Granada) investigate the long time asymptotics of a new class of models for interacting particles inspired from various phenomena in the biosciences. In this model, when two particles collide they may coalesce and then completely stop moving [13], http://hal.inria.fr/hal-00601969/en.

In collaboration with G. Pichot (INRIA Rennes Bretagne Atlantique), A. Lejay has developed a new Monte Carlo methods for discontinuous media that relies on the simulation of the Skew Brownian motion [22], [35], http://hal.inria.fr/hal-00642194/en, http://hal.inria.fr/hal-00649170/en.

A. Lejay developed a new method for the simulation of a stochastic process in a layered media using the properties of the Brownian path [20], http://hal.inria.fr/inria-00583127/en.


S. Maire and E. Tanré have generalised the spectral methods for elliptic PDEs developed in [39], [40] to the case of pure Neumann boundary conditions. Some additional difficulties occur because the stochastic representation of the solutions is defined only up to an additive constant and as a limit involving local time approximations [38]. By taking into account these additional properties, they still obtained a spectral matrix having a condition number converging to one.

D. Talay and E. Tanré, in collaboration with F. Delarue and S. Rubenthaler (Univ. Nice – Sophia Antipolis), have given a precise approximation of the interspike intervals for the LIF model, describing the activity of a single neuron. This work is part of the ANR MANDy project (see Section 7.1.1).

D. Talay, in collaboration with M. Martinez (Univ. Paris-Est), achieved to develop their stochastic approach for one-dimensional transmission parabolic problems. Owing to their stochastic representation of the solutions, they obtained accurate pointwise estimates for the derivatives of these solutions, from which they got accurate convergence rate estimates in the weak sense for a numerically effective discretization scheme of stochastic differential equations with weighted local times which are related to elliptic partial differential operators under divergence form with a discontinuous coefficient [36], http://hal.inria.fr/inria-00607967/en.

6.1.2. Other works in progress

N. Champagnat studies in collaboration with S. Méléard (Ecole Polytechnique, Palaiseau) adaptive dynamics and evolutionary branching in individual-based models of populations competing for resources, similar to those involved in chemostat systems of ODEs.
- N. Champagnat studies in collaboration with A. Lambert the process of the time to the most recent common ancestor in a family of subcritical branching processes whose genealogy is given by splitting trees.

- J. Charrier joined the team in September as a post-doctoral researcher and began working with M. Bossy and D. Talay on the long time behaviour of stochastic particules systems in McKean-Vlasov interaction.

- J. Claissse continued his PhD. under the supervision of N. Champagnat and D. Talay on stochastic control of population dynamics. He completed a finite-horizon and an infinite-horizon optimal control problem on a birth-death process. He is currently working on a birth-death process whose parameters depend on a controlled ordinary differential equation. In addition, he is working on applications of branching processes in biology and optimal control theory, and more specifically in cancer therapy.

- M. Deaconu and S. Herrmann continue the study of the hitting time for Bessel processes in the situation of noninteger dimensions.

- J. Inglis joined the team in October 2011 as a post-doctoral researcher (ANR MANDy), and began working with E. Tanré, D. Talay, F. Delarue (University of Nice) and S. Rubenthaler (University of Nice) on problems related to the rigorous justification of mean field models used in neuroscience.

- J. Inglis, E. Tanré and M. Tejo (PUC, Chile) started a collaboration on the numerical simulation of spiking times of neurons described by some new stochastic models related to the Hodgkin-Huxley equation. This work is a part of Anestoch associated team.

- A. Lejay and S. Maire study some new Monte Carlo methods for multi-dimensional discontinuous media.

- In collaboration with J.-R. Li (INRIA Rocquencourt & Neurospin), A. Lejay studies some probabilistic representation for interface condition arising in diffusion Magnetic Resonance Imaging.

- In collaboration with G. Pichot and J. Erhel (INRIA Rennes Bretagne Atlantique), A. Lejay studies Monte Carlo methods for discontinuous media as well as benchmarks and test on existing methods.

- With L. Coutin (Univ. Toulouse), A. Lejay studies some perturbation results for solutions of Rough Differential Equations.

- S. Maire develops with C. de Luigi (Univ. du Sud – Toulon – Var) and Jerôme Lelong (IMAG, Grenoble) resolution algorithms for the price of various european options in high dimension by coupling an adaptive deterministic integration algorithm and Principal Component Analysis tools.

- S. Niklitschek continued his PhD. under the supervision of D. Talay on discretized stochastic differential equations related to one-dimensional partial differential equations of parabolic type involving a discontinuous drift coefficient. He obtained accurate pointwise estimates for the derivatives of these solutions, from which he gets convergence rate estimates in the weak sense of the stochastic discretization scheme. Now he is working on the extension of these results to the multi-dimensional setup.

- N. Perrin continued his PhD. on stochastic methods in molecular dynamics under the supervision of M. Bossy, N. Champagnat and D. Talay. He is studying a method due to P. Malliavin (French Academy of Science) based on the Fourier analysis of covariance matrices with delay in order to identify the fast and slow components of a molecular dynamics and to construct simplified projected dynamics. He also studied probabilistic interpretation of the nonlinear Poisson-Boltzmann equation in Molecular Dynamics with BSDEs [37], http://hal.inria.fr/hal-00648180/en.

- L. Violeau continued his PhD. on Stochastic Lagrangian Models and Applications to Downscaling in Fluid Dynamics under the supervision of M. Bossy and A. Rousseau (MER team, INRIA Sophia Antipolis – Méditerranée, Montpellier). He studied the convergence in law of a sequence of penalized processes to the so called reflected langevin process in a convex domain. He is currently working on the rate of convergence of the particle approximation of conditional McKean stochastic models.
• P-E. Jabin and D. Talay continue to develop their innovating approach, which combines stochastic analysis and PDE analysis, for the time varying Hamilton-Jacobi-Bellman-McKean-Vlasov equations of the Lasry and Lions mean-field stochastic control theory.

6.2. Financial Mathematics

Participants: Mireille Bossy, Paul Charton, El Hadj Aly Dia, Dalia Ibrahim, Denis Talay, Etienne Tanré.

6.2.1. Published works and preprints

• In collaboration with N. Maïzi (CMA – Mines ParisTech) and O. Pourtallier (COPRIN team, INRIA Sophia Antipolis – Méditerranée), M. Bossy, and E.H.A. Dia studied the indifference pricing for carbon emission allowances, as a short term model value of carbon (see Section 7.1.2 ). The indifference pricing methodology describes the way an industrial agent on the emission allowances market chooses his production strategy. An utility function represents the preferences of the producer and its risk aversion. The outputs of its production have stochastic prices on the market, so that the optimal production strategy arises as the solution of a stochastic control problem. We extended the model hypotheses under which we get the well-posedness of the stochastic control problem and the associated HJB equation. We exhibited a simple case (marginal costs constant in time) where we proved the regularity of the value function via the explicit solution of the stochastic control problem \([24]\), \(\text{http://hal.inria.fr/hal-00645033/en} \). This particular case now can serve as a benchmark for the numerical solver currently developed in the framework of the ADEME Convention. It will also serve as a demonstrator case, with the objective of a public diffusion of the simulator CarbonQuant.

• M. Cissé (ENSAE-Sénégal), P. Patie (Univ. libre de Bruxelles) and E. Tanré have solved explicitly the optimal stopping problem with random discounting and an additive functional as cost of observations for a regular linear diffusion \([17]\), \(\text{http://hal.inria.fr/inria-00458901/en} \).

6.2.2. Other works in progress

• P. Charton continued his PhD. under the supervision of M. Deaconu and A. Lejay. He studied some hedging strategies for day ahead markets of wind energy.

• Mathematical modelling for technical analysis techniques Since November 2009, D. Ibrahim has been working on her PhD. thesis on Mathematical modeling of technical analysis in finance, under supervision of D. Talay and E. Tanré. The aim of her work is to study the performances of a technical analysis tool designed to detect changes in the volatility term: the Bollinger Bands. First, she studied the performances of this indicator in a modified Black-Scholes model such that the rate of volatility changes at an unknown random time \(\tau\), independent of the Brownian motion governing the prices. She is interested to study whether this indicator can detect the changes in the volatility. So, she aims to study the tail probability of this indicator by using Karamata’s Tauberian Theorem for Laplace-Stieltjes transforms.

Secondly, she exhibited a mathematical optimal strategy by modifying usual techniques in both the dual and the classical PDE approaches in stochastic control theory, in order to circumvent the discontinuity of the filtration generated by the price process.

This work is part of the contract with FINRISK (see Section 8.3 ).

• P. Protter (Columbia University) and D. Talay started to develop a new bubble time evolution model.
6. New Results

6.1. Automatic Differentiation and parallel codes

Participants: Valérie Pascual, Laurent Hascoët, Hubert Alcin, Jean Utke [Argonne National Lab. (Illinois, USA)], Uwe Naumann [RWTH Aachen University (Germany)].

This research is an ongoing joint work between three teams working on AD. We study differentiation in reverse mode of programs that contain MPI communication calls. Instead of the commonly used approach that encapsulates the MPI calls into black-box subroutines that will be differentiated by hand, we are looking for a native differentiation of the MPI calls by the AD tool. The ultimate goal of this work is to generate the adjoint of MPI-parallel codes.

One issue is to reduce the variability of the available MPI procedures and parameters to a smaller number of elementary concepts. We then address the basic question of sends and recepts, that may be blocking or nonblocking, individual or collective, and so on. Essentially the adjoint of a send is a recv, and vice-versa, but the possibility of nonblocking isend's and irecv's requires more subtlety.

We continue adaptation of the tool’s static analysis to programs with message-passing communication. In the framework of flow-sensitive and context-sensitive data-flow analysis, we introduce new "channel" variables and modify the general static data propagation mechanism.

Implementation in TAPENADE is well advanced: the main MPI procedures are now correctly understood and all data-flow analyses are adapted. In 2011, we obtained a first valid tangent differentiated code of the team’s CFD code AIRONUM. Work is continuing to obtain the adjoint.

Our team focuses on AD based on program transformation. On the other hand, we closely follow the developments of operator-overloading AD tools towards message-passing communication. This requires a more complex definition of the overloaded communication primitives [33].

6.2. AD adjoints and Dynamic Memory

Participants: Laurent Hascoët, Jean Utke [Argonne National Lab. (Illinois, USA)].

Again in the adjoint mode, dynamic memory allocation and the associated pointer manipulations pose difficult problems. As the adjoint mode is bound to recover past values of variables, the need arises to recover past values of pointers too. However, these values are addresses often relative to some dynamically allocated memory. The original program often manages memory by allocating and deallocating memory on the run. Correspondingly, the adjoint program will have to reallocate memory as it walks backwards along the original allocation. If recovery of past values is implemented through storage, e.g. on a stack, the stored addresses refer to memory that has been deallocated, and do not correspond to the reallocated memory zones.

In 2011 we investigated this problem in two principal directions:

- We may consider not storing the address, but rather recompute it by repeating in the reverse sweep its calculation from the forward sweep. This is not always possible as pointer assignment and manipulation may be complex and distributed in the code. However, when possible, it is certainly a very efficient approach. We propose an data-flow algorithm to detect applicable cases.

- For the remaining cases where storage must be used, we study an address mapping mechanism, such that addresses inside some allocated memory can be dynamically converted into addresses inside the corresponding reallocated memory.
6.3. Resolution of linearised systems

Participants: Hubert Alcin, Olivier Allain [Lemma], Anca Belme, Marianna Braza [IMF-Toulouse], Alexandre Carabias, Alain Dervieux, Bruno Koobus [Université Montpellier 2], Carine Moussaed [Université Montpellier 2], Hilde Ouvrard [IMF-Toulouse], Stephen Wornom [Lemma].

The interaction between the sophisticated solution algorithm inside a program and the Automatic Differentiation of the program is a non-trivial issue. An iterative algorithm generally does not store the successive updates of the iterated solution vector. Furthermore, a modern iterative solution algorithm involves several nonlinear processes, like in:

- the evaluation of an optimal step, which results at least from a homographic function of the unknown,
- the orthonormalisation of the updates (Gram-Schmidt method, Hessenberg method).

Applying reverse AD to the iterative solution algorithm produces a linearised iterative algorithm which is transposed and therefore follows a reverse order, with exactly the same number of iterations, and needing exactly each of the iterated state solution vectors. This effect is considerably amplified in the case of the numerical simulation of unsteady phenomena with implicit numerical schemes. For example, the simulation of high Reynolds turbulent flows by a Large Eddy Simulation (LES) requires hundreds of thousands time steps, each of them involving a modern iterative solution algorithm.

In the 4-year ECINADS ANR project, we design more efficient solution algorithms and we examine the questions risen by their reverse differentiation. The application domain is the computation of high Reynolds turbulent flows with LES and hybrid RANS-LES models. The efficiency will be evaluated through the practical scalability on a large number of processors. This efficiency criterion also extends to the scalability of the reverse/adjoint algorithm. ECINADS also addresses the scalable solution of new approximations. ECINADS associates the university of Montpellier 2, the Institut de Mécanique des fluides de Toulouse and Lemma company. The kick off meeting of ECINADS was held at end of 2009.

In 2011, Hubert Alcin has performed a study of deflation and balancing coarse grid methods for a set of scalar models. The methods has been extended to the incompressible Navier-Stokes model in Lemma’s software ANANAS by Olivier Allain and to compressible Navier-Stokes by Bruno Koobus and Hubert Alcin. A collection of benchmark tests on these models has been performed and show a good scalability for the tested algorithms. An article is prepared on these results. Hubert Alcin has also studied a novel method for three-level preconditioning. The new method will be extended to compressible flows in cooperation with the Montpellier team (Carine Moussaed and Bruno Koobus).

6.4. Perturbation Methods

Participants: Alain Dervieux, Laurent Hascoët.

In the context of the European project NODESIM-CFD, the contribution of Tropics involved mainly the differentiation of perturbation methods and reduced order models for the management of uncertainties. These methods rely on Taylor series with second-order terms. The production of second derivative code is obtained through repeated application of Automatic Differentiation. Three strategies can be applied to obtain (elements of) the Hessian matrix, named Tangent-on-Tangent (ToT), Tangent-on-Reverse (ToR), and Reverse-on-Tangent (RoT). These new methods are disseminated through short courses, as those given by Alain Dervieux at ERCOFTAC sessions (Munich and Hampton). The application and extension of these methods have motivated a joint application from the Italian Aircraft company Alenia and Tropics in a FP7 proposal (Proposal CARDINA, nov. 2011).

6.5. Control of approximation errors

Participants: Frédéric Alauzet [GAMMA team, INRIA-Rocquencourt], Estelle Mbinky [GAMMA team, INRIA-Rocquencourt], Olivier Allain [Lemma], Anca Belme, Alexandre Carabias, Hubert Alcin, Alain Dervieux.
This is a joint research between three INRIA teams GAMMA (Rocquencourt), TROPICS, PUMAS and Lemma company. Roughly speaking, GAMMA brings mesh and approximation expertise, TROPICS contributes to adjoint methods, and CFD applications are developed in the context of PUMAS and Lemma.

The resolution of the optimum problem using the innovative approach of an AD-generated adjoint can be used in a slightly different context than the optimal shape design, namely the mesh adaptation. This will be possible if we can map the mesh adaptation problem into a differentiable optimal control problem. To this end, we have introduced a new methodology that consists in stating the mesh adaptation problem in a purely functional form: the mesh is reduced to a continuous property of the computational domain: the continuous metric. We minimize a continuous model of the error resulting from that metric. Thus the problem of searching an adapted mesh is transformed into the search of an optimal metric.

In 2011, a work on goal-oriented mesh adaptation for unsteady Euler flows has been extended, with further analysis, and a paper has been written and submitted to a journal. Its extension to the compressible Navier-Stokes model has been developed, [11] and a paper is being written. A further extension to Large Eddy Simulation is started. The method is being extended to a third-order approximation, the Vertex-CENO. This approximation was defined during a collaboration between university of Montpellier, IMM-Moscow and Tropics. A more accurate version has been studied by Alexandre Carabias and presented in Honom-Trento. A new theory involving error estimates and criteria has been developed by Gamma and Tropics. The extension of the multiscale adaptation method is considered by Estelle Mbinky at Rocquencourt. The extension of the goal-oriented method is considered by Alexandre Carabias at Sophia. Anisotropic mesh adaptation allows for better convergence to continuous solutions, and in particular more accurate a posteriori error estimates and correctors. The synergy between correctors and mesh adaptation is currently analysed and is the subject of a joint contribution (Gamma and Tropics) for the FP7 CARDINA proposal (nov. 2011).
5. New Results

5.1. Modeling Interfaces and Contacts

5.1.1. On the Morphology of Protein Binding Patches

Participants: Frédéric Cazals, Noël Malod-Dognin.

In collaboration with A. Bansal, former summer intern from IIT Bombay.

Understanding the specificity of protein interactions is a central question in structural biology, whence the importance of models for protein binding patches—a patch refers to the collection of atoms of a given partner accounting for the interaction. To improve our understanding of the relationship between the structure of binding patches and the biological function of protein complexes, we present a binding patch model decoupling the topological and geometric properties \[21\]. While the geometry is classically encoded by the 3D positions of the atoms, the topology is recorded in a graph encoding the relative position of concentric shells partitioning the interface atoms. The topological - geometric duality provides the basis of a generic dynamic programming based algorithm to compare patches, which is instantiated to respectively favor topological or geometric comparisons.

On the biological side, using a dataset of 92 co-crystallized structures organized in biological sub-families, we exploit our encoding and the two comparison algorithms in two directions. First, we show that Nature enjoyed the topological and geometric degrees of freedom independently while retaining a finite set of qualitatively distinct topological signatures, and show that topological similarity is a less stringent notion than the ubiquitously used geometric similarity. Second, we analyze the topological and geometric coherence of binding patches within sub-families and across the whole database, and show that complexes related to the same biological function can encompass geometrically distinct shapes. Previous work on binding patches focused on the investigation of correlations between structural parameters and biochemical properties on the one hand, and on structural comparison algorithms on the other hand. We believe that the abstraction coded by the topological - geometric duality paves the way to new classifications, in particular in the context of flexible docking.

The corresponding software is presented in section 4.1.1.

5.2. Modeling Macro-molecular Assemblies

5.2.1. Assessing the Reconstruction of Macro-molecular Assemblies with Toleranced Models

Participants: Frédéric Cazals, Tom Dreyfus.

In collaboration with Valérie Doye, Institut Jacques Monod, Paris.

In [20], we introduce TOleranced Models (TOM), a generic and versatile framework meant to handle models of macro-molecular assemblies featuring uncertainties on the shapes and the positions of proteins. A TOM is a continuum of nested shapes, the inner (resp. outer) ones representing high (low) confidence regions, we present statistics to assess features of this continuum at multiple scales. While selected statistics target topological aspects (pairwise contacts, complexes involving proteins of prescribed types), others are of geometric nature (geometric accuracy of complexes). We validate the TOM framework on recent average models of the Nuclear Pore Complex (NPC) obtained from reconstruction by data integration, and confront our statistics against experimental findings related to sub-complexes of the NPC. In a broader perspective, the TOM framework should prove instrumental to handle uncertainties of various kind, in particular in electron-microscopy and crystallography.
5.2.2. Probing a Continuum of Macro-molecular Assembly Models with Graph Templates of Sub-complexes

Participants: Frédéric Cazals, Tom Dreyfus.

Reconstruction by data integration is an emerging trend to reconstruct large protein assemblies, but uncertainties on the input data yield average models whose quantitative interpretation is challenging. This paper presents methods to probe fuzzy models of large assemblies against atomic resolution models of sub-systems.

Consider a Toleranced Model (TOM) of a macro-molecular assembly, namely a continuum of nested shapes representing the assembly at multiple scales. Also consider a template namely an atomic resolution 3D model of a sub-system of this assembly—also called a complex. We present algorithms performing a multi-scale assessment of the complexes of the TOM, by comparing the pairwise contacts which appear in the TOM against those of the template. These operations reduce to the comparison of graphs, which we perform by computing Maximal Common Induced Sub-graphs (MCIS) and Maximal Common Edge Sub-graphs (MCES).

We apply this machinery to recent average models of the NPC. First, we show how our contact analysis allows assessing the quality of probability density maps. Regarding particular sub-systems of the NPC, we focus on the Y-complex and on the T-complex. In particular for the latter, our analysis suggests a new 3D template of pairwise contacts.

We believe that these tools should become standard to assess the reconstruction of fuzzy assemblies.

The software associated to these developments is presented in section 4.1.2.

5.3. Protein Shape Matching and Family Identification

5.3.1. Using Dominances for Solving the Protein Family Identification Problem

Participant: Noël Malod-Dognin.

In collaboration with R. Andonov (IRISA), M. Le Boudic-Jamin (IRISA) and P. Kamath (former summer intern within the SYMBIOSE project at IRISA).

The 3D structure of macro-molecules underpins all biological functions. Similarities between protein structures may come from evolutionary relationships, and similar protein structures relate to similar functions.

The exponential growth of the number of known protein structures in the Protein Data Bank over the past decade led to the problem of protein classification. We mean here how to automatically insert new protein structures into an already existing classified database \( Q = \{ q_1, q_2, \cdots, q_m \} \) such as CATH or SCOP. The problem of determining in which classes new structures \( P = \{ p_1, p_2, \cdots, p_n \} \) belong, according to a similarity function \( S : Q \times P \rightarrow \mathbb{R}^+ \), is referred here as the Protein Family Identification Problem (FIP).

There are computational pitfalls in the FIP. The number of similarity scores \( S(q_i, p_j) \) that need to be computed is \( |Q| \times |P| \), where \( |P| \) can be very large (there are currently 152920 classified protein structures in the expert classification CATH). Moreover, computing a single similarity score is often equivalent to solving a NP-hard problem (ex: DALI, DAST, CMO, VAST, etc...).

In [17] and [18], we propose a notion of dominance between the protein structure comparison instances that allows the computation of optimal FIP without optimally solving all the comparison instances, and thus reduces the effect of the NP-Hardness of the similarity score.

5.4. Algorithmic Foundations

5.4.1. Shape Matching by Localized Calculations of Quasi-isometric Subsets

Participants: Frédéric Cazals, Noël Malod-Dognin.
Consider a protein complex involving two partners, the receptor and the ligand. In [16], we address the problem of comparing their binding patches, i.e. the sets of atoms accounting for their interaction. This problem has been classically addressed by searching quasi-isometric subsets of atoms within the patches, a task equivalent to a maximum clique problem, a NP-hard problem, so that practical binding patches involving up to 300 atoms cannot be handled. We extend previous work in two directions. First, we present a generic encoding of shapes represented as cell complexes. We partition a shape into concentric shells, based on the shelling order of the cells of the complex. The shelling order yields a shelling tree encoding the geometry and the topology of the shape. Second, for the particular case of cell complexes representing protein binding patches, we present three novel shape comparison algorithms. These algorithms combine a Tree Edit Distance calculation (TED) on shelling trees, together with Edit operations respectively favoring a topological or a geometric comparison of the patches. We show in particular that the geometric TED calculation strikes a balance, in terms of accuracy and running time between a purely geometric and topological comparisons, and we briefly comment on the biological findings reported in a companion paper [21].
5. New Results

5.1. RNA structures

5.1.1. RNA secondary structures: folding, design and evolution

In a collaboration with J. Waldispuhl (McGill, Canada) (Presented at the RECOMB’11 conference [32]), we used weighted grammatical models, introduced by members of the group [2], to perform an efficient exploration of the mutational landscape of RNA. We proposed an adaptive sampling algorithm, where weights were used to compensate an identified bias toward regions of higher GC-content within sampled sequences, thereby allowing for the exploration of more relevant portions of the evolutionary landscape. These adaptive sampling principles can be adapted into a method for the RNA design following similar principles. This constitutes a competitive alternative to local search strategies used by all existing tools for this problem. This work is ongoing as a collaboration with B. Berger group (MIT) and J. Waldispuhl (McGill, Canada), and a manuscript was recently submitted.

5.1.2. RNA knowledge-based potentials and 3D studies

We used the curated database of biologically interesting structures we have set up to perform a statistical analysis and developed knowledge-based potentials. The database server is available at http://csb.stanford.edu/rna.

We obtained RNA knowledge-based potentials that now performs well at different representation levels. They can be used in three well-known Molecular Dynamics (MD) and modeling software suites ENCAD [44], GROMACS (v3 and 4) [45] and MOSAICS [41] and are available for the community. The study we performed on a large number of new decoys showed that our potential outperforms Rosetta RNA scoring function [37] which is the gold standard. We show that not having correction terms for base-stacking and pairing can be of advantage when modelling loops at high resolution. The study was welcomed by the RNA community and published in RNA Journal (IF 6.5) [8].

We also refined the mixture model strategy we developed for building knowledge-based potentials. In collaboration with O. Schwander at LIX, we compared different mixture models: Dirichlet Process Mixture models (DPM), Kernel Density Estimation (KDE) models, Expectation Maximization models (MM) with different number of components (including a simplified version based on a post-processing step using K-Means). We showed that the Dirichlet Process Mixtures (DPM) is a good tradeoff despite its longer precomputation time as it provides a smooth potential having relatively few components. This study was presented at the MCMMB’11 conference and was submitted as a journal paper.

This work was done in collaboration with A. Sim, X. Huang an M. Levitt (Stanford University - GNAPI Associate team).

5.2. Proteins structures

5.2.1. Protein sequence alignment

In comparative protein modeling, the quality of a template model depends heavily on the quality of the initial alignment between a given protein with unknown structure to various template proteins, whose tertiary structure is available in the Protein Data Bank (PDB). Although pairwise sequence alignment has been solved for more than three decades, there remains a large discrepancy between the accuracy of the best sequence alignment between two amino acid sequences, as produced by the Needleman-Wunsch or Smith-Waterman algorithms, and that of the best structural alignment between two protein X-ray structures, as produced by the software DALI, CE, TOPOFIT, etc. To improve the quality of initial alignments in template modeling, one can integrate valuable information from an ensemble of generated suboptimal alignments, that is alignments whose score is below the best possible score. In a collaboration with P. Clote (Boston College/DIGITEO) [26], we presented a novel algorithm to produce suboptimal pairwise alignments.
5.2.2. Protein-protein interaction:

A protein-protein docking procedure traditionally consists in two successive tasks: a search algorithm generates a large number of candidate solutions, and then a scoring function is used to rank them in order to extract a native-like conformation. We have already demonstrated that using Voronoi constructions and a defined set of parameters, we could optimize an accurate scoring function. However, the precision of such a function is still not sufficient for large-scale exploration of the interactome.

Another geometric construction was also tested: the Laguerre tessellation. It also allows fast computation without losing the intrinsic properties of the biological objects. Related to the Voronoi construction, it was expected to better represent the physico-chemical properties of the partners. In , we present the comparison between both constructions.

We also worked on introducing a hierarchical analysis of the original complex three-dimensional structures used for learning, obtained by clustering. Using this clustering model we can optimize the scoring functions and get more accurate solutions. This scoring function has been tested on CAPRI scoring ensembles, and an at least acceptable conformation is found in the top 10 ranked solutions in all cases. This work was part of the thesis of Thomas Bourquard, defended in 2009.

A strong emphasis was recently made on the design of efficient complex filters. To achieve this goal, we focused on the use of collaborative filtering methods state of the art machine learning approaches combined with our genetic algorithm [ 9 ].

We have also proposed an approach that improves the predictions made by HEX, a state-of-the art docking tool developed by INRIA Nancy. We applied Voronoi fingerprint to the output of HEX and learn how to rank them, and we have tested new ranking strategies. The obtained ranking improve the initial ranking of HEX [ 33 ], [ 23 ].

We also decided to extend these techniques to the analysis of protein-nucleic acid complexes. The first preliminary developments and tests were performed by Adrien Guilhot during his M1 internship for two months.

5.2.3. Transmembrane β-barrels:

We have recently proposed an algorithm [ 31 ] that classifies Transmembrane β-Barrel Proteins (TMB) and predicts their structure. It first uses a simple probabilistic model to filter out the proteins and strands which are not beta-barrel. Then, we build a graph-theoretic model to fold into the super-secondary structure via dynamic programming. This step runs in \( O(n^3) \) time for the common up-down topology, and at most \( O(n^5) \) for the Greek key motifs, where \( n \) is the number of amino acids. Finally a predicted three-dimensional structure is built from the geometric criteria. If the pseudoenergy is insufficient, the protein is classified as a non-TMB protein.

We have tested this approach on TMB and non-TMB proteins for classification and structure prediction. We tested classification on a dataset of 14238 proteins including 48 TMB and 14190 non-TMB proteins. Our classification results are very accurate and comparable to other algorithms [ 21 ], [ 5 ]. Especially, our PPV, MCC and F-Scores are second only to a very recent algorithm by Freeman and Wimley [ 39 ], which relies heavily on training data. We also tested the structure prediction on 42 proteins from the TMB and compared to other existing algorithms. The results are comparable to existing algorithms, the accuracy ranges from 85-93%, depending upon the parameter used. This is very promising given that other algorithms rely heavily on homology and training datasets and may be overfitting. Our approach can be further improved by refining the energetic model, especially on turns and loops.

In addition, we have developed consensus methods to combine multiple secondary structures into one more reliable solution. Our results show that our technique can be used to combine multiple solutions to produce structures that are more than any of the input structures. These methods are based mainly on social choice theory and known properties of TMB proteins. In addition, we are working on methods for combining information on the super-secondary structures, and using them to augment the supersecondary structure provided by our approach.
5.3. Combinatorics and Annotation

5.3.1. Word counting and random generation

Cis-Regulatory modules (CRMs) of eukaryotic genes often contain multiple binding sites for transcription factors, or clusters. Formally, such sites can be viewed as words co-occurring in the DNA sequence. This gives rise to the problem of calculating the statistical significance of the event that multiple sites, recognized by different factors, would be found simultaneously in a text of a fixed length. A long-term research on word enumeration has been realized by the team. An extension to Hidden Markov Models has been realized recently in a collaboration with M. Roytberg (IMPB, Pushchino, Russia). It relies on a new concept of overlap graphs that efficiently overcomes the main difficulty - overlapping occurrences - in probabilities computation. This is part of E. Furletova’s thesis, to be defended soon. An implementation is available at http://server2.lpm.org.ru/bio. This algorithm provides a significant space improvement over a previous algorithm, AHO.PRO developed with our former associate team MIGEC. M. Régnier and S. Sheikh have addressed combinatorial problems on clumps that should allow further space decrease and large deviation results were presented at MCCMB’11.

An other application of word combinatorics has been started this year. During his internship, L. Pei (Paris-Sud 11 U.) provided a pipeline that simulates a random generation of reads and assembles them using MIRA software. This work will be pursued by D. Iakovishina in her thesis. It is a collaboration with MAGNONE at INRIA-BORDEAUX and IOGENE in Moscow.

A previous work [36], published in 2010, generalized Boltzmann samplers to multivariate objects, allowing for the efficient random generation achieving a fixed or approximate composition for context-free languages. However, the performances of such algorithms were only guaranteed in the case of strongly-connected context-free grammars. In a recent collaboration with O. Bodini, H. Tafat and C. Banderier (LIPN, Paris-XIII) we are working on characterizing the distributions arising from simply connected grammars. In a short paper accepted for presentation at the ANALCO’12 conference [24], we showed that: i) a large class of distributions can be reached for the number of occurrence of a single letter, arguably the simplest observable pattern; ii) simple grammars/regular expressions can be built that realize these distributions; iii) Classic Boltzmann samplers remain largely unaffected by this diversity.

Our work on random generation has applications in software testing and model-checking, in a collaboration with the Fortesse group at LRI [13], [29].

5.3.2. RNA combinatorics

Pseudoknots are usually ignored by popular software for RNA prediction. This means that, even under the daring assumptions of an unique and well-defined fold for RNA, coupled with a perfectly accurate energy model, the real structure of RNA will not be recovered perfectly. In a collaboration between AMIB members and S. Janssen (Universität Bielefeld), we investigated the practical implications of such a limitation. We used RNAFOLD, a popular software for the prediction of RNA structure on representative sequences of the RFAM database, which groups known RNA sequences into about 2000 functional families. We observed that 12% of RFAM families exhibited a total absence of overlap between predicted structures and manually-curated structures, derived from experimental or evolutionary data. Combination of RFAM annotations, a survey of literature, and a newly developed predictive method for the presence of a functional pseudoknots, we were able to validate that a large majority of the mispredicted families featured evidence of pseudoknots in the functional conformation. Preliminary results were presented by B. Raman at the Fifth Indo-French Bioinformatics Meeting [34].

In 2004, Condon and coauthors gave a hierarchical classification of exact RNA structure prediction algorithms according to the generality of structure classes that they handle. In [19], we completed this classification by adding two recent prediction algorithms. More importantly, we precisely quantified the hierarchy by giving closed or asymptotic formulas for the theoretical number of structures of given size n in all the classes but one. This allows to assess the tradeoff between the expressiveness and the computational complexity of RNA structure prediction algorithms.
Similar decompositions can be used for the design of algorithms that include tractable subclasses of pseudo-knots. In [30] Y. Ponty and C. Saule extended a unifying framework introduced by Roytberg and Finkelstein to design ensemble RNA algorithms. This framework uses a family of hypergraphs to describe the conformation space, allowing for a clear separation between the search space, i.e. the set of admissible conformations, and the intended application (Minimal Free-Energy folding, partition function, statistical sampling...). We illustrated the promises of such an approach by explicitly rephrasing three major search spaces within the framework, and introduced an algorithm for computing the moments of any additive feature in the Boltzmann distribution.

By comparing empirical observations with the expected behavior of a model, combinatorial methods can be used to identify an evolutionary pressure weighing on RNA. In a collaboration with P. Clote (Boston College/DIGITEO) [11], we used analytic combinatorics to study the expected distance between both ends of an RNA molecule, or 5'-3' distance. Postulating a Boltzmann distribution on all secondary structures, we showed that this parameter is bounded by a – typically small – constant value when the sequence length goes to the infinity. Computing this quantity on a database of experimentally-determined secondary structures, we observed that the 5'-3' distances take larger values than those predicted from the model. Furthermore, quite surprisingly, this quantity was shown to correlate positively with the length. We concluded by hypothesizing that the secondary structure of RNA may be under evolutionary pressure to fold in a modular way, creating independent domains on the exterior face.

5.3.3. Data integration

Recent years have seen a revitalization of Data Integration research in the Life Sciences. But the perception of the problem has changed: While early approaches concentrated on handling schema-dependent queries over heterogeneous and distributed databases, current research emphasizes instances rather than schemas, tries to place the human back into the loop, and intertwines data integration and data analysis. In this domain, the contribution of AMIB in 2011 has been three folds: First, we have followed our collaboration with Ulf Leser (invited in the AMIB group at LRI during 6 months in 2010) and have worked on the review of the past and current state of data integration for the Life Sciences and discussed recent trends in detail, which all pose various challenges for the database community in [28]. Additionally, we have worked on a vision of what should be done by workflow systems to make it possible to search, adapt, and reuse scientific workflows, the complete state-of-the-art on this domain has been provided [12]. Second, in close collaboration with oncologists from the Institut Curie and the Children’s Hospital of Philadelphia we have worked on the problem of ranking genes of interest associated to a given disease. The software GENEVALORIZATION has been designed and developed in this context and is able to provide a concise view of the literature available associated to a list of genes [10]. A second aspect of this research has been the design of a consensus ranking method, BioConsert, able to make the most (ie underline common points) of a set of established rankings [27]. This last point has been done in close collaboration with Sylvie Hamel invited professor in our group in 2010 (2 months). Third, we have presented a simple logical query language called RL for expressing different kinds of rules, especially well-suited to express association rules for transcriptomic data. In that context the challenge is to find out relationships between genes that reflect observations of how expression level of each gene affects those of others. The conjecture that association rules could be a model for the discovery of gene regulatory networks has already been partially validated. Nevertheless, several different kinds of rules between genes could be useful with respect to some biological objectives and we have designed a framework in which biologists may define their "own customized semantics" for rules with regard to their requirements. We have studied how the RL language behaves with respect to the well-known Armstrong’s axioms [22]. The main contribution of this paper is to exhibit a restricted form of RL-queries, yet with a good expressive power, for which Armstrong’s axioms are sound. From this result, this sublanguage turns out to have structural and computational properties which have been shown to be very useful in data mining, databases and formal concept analysis.

5.4. Systems Biology
In her thesis, M. Behzadi has developed a know-how on the behaviour of biological systems along a cooperation with an INSERM-INRA team based in Clermont-Ferrand. In the methodology that was developed, one computes the equations’ parameters from the experimental data in systems that can be considered at equilibrium. It was proved mathematically that some sub-domains are intrinsically stable and that their behaviour is not much affected by the initial conditions [4] for phospholipids biosynthesis. A review for carbone toxicity can also be found in [15]. Software Analyser software (MPSA) are currently under development by L. Paulevé.

Elementary flux mode is a fundamental concept as well as a useful tool in metabolic pathway analysis. However, when the networks are complex, the determination of elementary flux modes leads to combinatorial explosion of their number which prevents from drawing simple conclusions from their analysis. To deal with this problem, a biclustering method has developed [18] based on the Agglomeration of Common Motifs (ACoM). It was applied to the central carbon metabolism in Bacillus subtilis and to the yeast mitochondrial energy metabolism. It helped to give biological meaning to the different elementary flux modes and to the relatedness between reactions.

Once molecules and complexes participating in the signalling network have been identified, the relations between them (enzymatic reactions, activations, inhibitions, etc.) have to be deduced from experimental and literature data to build the influence graph. Partners INRA-BIOS and AMIB have started the development of a knowledge-based method, which uses the solver SOLAR, developed by NII (Tokyo), that allows automating this data integration task. We have already formalized the knowledge necessary for inferring the signalling network triggered by the FSH receptor, one famous GPCR. Preliminary results of this project A Sam are very encouraging.
5. New Results

5.1. New results in the theory of factorization of boundary value problems

Participants: Jacques Henry, Fadhel Jday, Maria Orey.

We are pursuing the development of the theory of factorization of boundary value problems as described in 3.1. Maria Orey who suspended her PhD thesis for a while due to health reason, has resumed her work on extending the method of factorization to the analogous in infinite dimension of the QR algorithm for matrices. This passes through the factorization of the normal equation for the least squares problem. This problem is solved and this allows a clear definition of the Q and R operators. She will defend her thesis in 2012.

F. Jday has obtained also a clear formulation for the factorization of the Stokes equation.

A progress has been made in the attempt to extend the factorization method to parabolic evolution equation. It appears that it is not the evolution problem that can be factorized with respect to space but the evolution operator $S(t)$ that transfers the solution from time 0 to $t$. The corresponding Riccati equation has been obtained but a full mathematical justification remains to be done.

5.2. Data completion problems for elliptic equations using the theory of factorization

Participants: Jacques Henry, Fadhel Jday.

F. Jday is continuing his thesis co-supervised by A. Ben Abda and J. Henry. The use of the method of factorization for the data completion problem has been presented in a paper in Inverse Problems when the domain is a cylinder, both Dirichlet and Neumann data are known on one face and are to be estimated on the other. Dirichlet boundary conditions were assumed on the lateral boundary. In relation to the inverse problem in electrocardiology one has to take into account more complex geometries. F. Jday considered the domain limited by two concentric spheres. The method still applies considering the family of concentric spheres deduced by homothety.

5.3. Modeling the activity of populations of neurons: study of synchronization

Participants: Jacques Henry, Gregory Dumont, Oana Tarniceriu.

During the second year of his PhD thesis G. Dumont continued to develop and refine a simulator of a population of leaky integrate and fire neurons, with a finite jump of potential response to a synaptic stimulation in order to compare with the results from another approach, the Fokker-Planck approach. It would be interesting to show that the Fokker Planck approach can be seen as a limiting case, for high frequency small jump of potential, of the one of B. Knight, L. Sirovich and L. Omurtag. At least it is suggested by the simulations. The second year focused on the study of mathematical aspects of the equation. The methods were inspired by the book of Benoit Perthame: Transport equation in biology. With these tools, the mathematical properties of the model of population density have been established: the existence and uniqueness of a solution. G. Dumont has also been able to establish that in special cases the equation has no solution, or more precisely that the solution blows up, assuming that there is no conduction delay of the spikes within the considered population. This blow up can be related to a synchronization. In case of a conduction delay the condition of existence of a solution are much larger. The situation of inhibiting synapses has also been studied. This has been submitted to the Journal of Mathematical Biosciences.

Using the principles of entropy, G. Dumont has shown that under certain circumstances the solution of the equation converges to a stationary solution and this can be interpreted as the desynchronization of the population of neurons.
Similar studies have been done on the thetaneuron model which accounts for self spiking neurons.

The question of the synchronization of a population of neurons is also studied in collaboration with O. Tarniceriu. We consider a population of identical self firing neurons that are weakly coupled and we study the long term evolution of the reaptrition of phases. We are now focusing to a population of 1D leaky integrate and fire neurons and we expect more precise results for this simple model.

5.4. Modeling in electrocardiology

Participants: Bedr’Eddine Ainseba, Jacques Henry, Yves Couëdirè, Simon Labarthe, Alejandro Lopez Rincon.

5.4.1. Modeling the electrical activity of the atria

The first year of the PhD thesis of Simon Labarthe has been dedicated to three main activities:

- a review of medical and mathematical literature about anatomy and physiology of atria and atrial fibrillation, numerical simulation of cardiac activity, mathematical analysis of cardiac models.
- the numerical implementation of simulation tools has been developed: the software is able to take into account a real 3D or 2D manifold geometry with realistic fibre orientation to simulate the bidomain or monodomain problem. Geometrical tools facilitating the fibre orientation construction have been added.
- numerical and theoretical studies of the modeling and the influence of fibre orientation in the atria and the pulmonary veins have been initiated.

5.4.2. Inverse problem in electrocardiology

The PhD thesis of Alejandro Lopez is devoted to improving the resolution of the inverse problem to recover the potential map on the heart from the measured potentials on the torso.

To have a best solution of the inverse problem many steps were taken. First, a software to solve the finite element method in 3d for the Poisson equation was created, and compared to an analytical solution. Also a software for the heat equation was created to prepare the the monodomain equation. The improvement of the forward solution helps to improve the inverse solution. An article explaining the software was included in the book "VEHÍCULOS AEROESPACIALES" (aerospace vehicles) published by the Sociedad Mexicana de Ciencia y Tecnología Aeroespacial (Mexican Society of Aerospace Science, and Technology) [19].

The second step was to develop a static solution of the inverse problem constructing a transfer matrix (heart to thorax), and solving with conjugate gradient. An article describing this software was accepted in the 22nd International Conference on Electronics, Communications and Computers.

As a third step the boundary element method was implemented to see the advantages and disadvantages in comparison with the finite element method.

The next step was to implement the monodomain model of the heart, to try to solve the problem of identification of parameters (in progress).

5.5. Invasion processes and modeling in epidemiology

Participants: Jean-Baptiste Burie, Arnaud Ducrot.

We have derived a homogenized version of our model of a fungal of disease of vine(see [7]) that takes into account the periodic row structure of vineyards. We used two-scale homogenization of Nguentseng and Allaire (see e.g. [20]): the macroscopic scale is the vineyard scale while the microscopic one is the plant scale. We have proved a result of convergence towards the homogenized model and numerical simulations have demonstrated a significant gain in computing time and stability.

These results have been presented at ECMTB2011 in Krakow (and also, at the Exploratory Workshop on Emerging Infectious Diseases and Mathematical Modelling, Barcelona 2011 and seminars at Franche-Comté University, Tamkang and National Taiwan University).

An article is to be submitted in SIAM J. Multiscale Modeling and Simulation in early 2012.
5. New Results

5.1. Medical Image Analysis

5.1.1. Spatial Decision Forests for MS Lesion Segmentation in Multi-Channel MR Images

Participants: Ezequiel Geremia, Nicholas Ayache, Olivier Clatz, Antonio Criminisi [MSR], Ender Konukoglu [MSR], Bjoern Menze [MIT].

- A new approach for MS lesions segmentation was proposed [33].
- Random forest for automatic segmentation of MS lesions in 3D MR images.
- Features: multi-channel MR intensities, priors, long-range spatial context, symmetry.
- Quantitative evaluation shows significant improvement over the MICCAI Grand Challenge 2008.
- The automatically learned decision sequence mimics the state-of-the-art pipeline.
- Independent validation carried out by the MICCAI Challenge website 2008.
- Exhaustive analysis of the discriminative power of channels and features.
- Analysis of the influence of random forest’s meta-parameters on the classification performance.

5.1.2. Left Ventricle Segmentation from Cardiac 4D Cine MRI Sequences

Participants: Jan Margeta [Correspondant], Ezequiel Geremia, Nicholas Ayache, Antonio Criminisi [MSR].

This work was performed in collaboration with Microsoft Research and was partly supported through its PhD Scholarship Programme.

- We extend the previous work for multiple sclerosis lesion segmentation of Geremia et al. [33] for spatio-temporal cardiac images.
- A fully automatic two layer left ventricle segmentation algorithm from 4D cardiac cine MRI sequences (See Fig. 2) was proposed for MICCAI STACOM LV segmentation challenge [63] using a random forest classification algorithm.
- Spatio-temporal features are used in the random forest framework to learn the segmentation task without explicitly defining the segmentation rules.
- Machine learning based MRI intensity standardization and pose normalization preprocessing pipeline was proposed to deal with diverse cardiac MRI datasets.

5.1.3. Design and use of anatomical atlases for automatic segmentation: application to radiotherapy of the head and neck region

Participants: Liliane Ramus [Correspondant], Grégoire Malandain, Vincent Grégoire [UCL], Juliette Thariat [CAL].

This work is done in collaboration with DOSIsoft S.A., Centre Antoine Lacassagne (CAL) and Université Catholique de Louvain.
Figure 1. **Segmenting Case CHB05 from the public MSGC dataset.** From left to right: preprocessed T1-weighted ($I_{T1}$), T2-weighted ($I_{T2}$) and FLAIR MR images ($I_{FLAIR}$) overlayed with the associated ground truth GT, the posterior map $\text{Posterior} = (P_{\text{lesion}}(v_k))_k$ displayed using an inverted grey scale and the FLAIR sequence overlayed with the segmentation $\text{Seg} = (\text{Posterior} \geq \tau_{\text{posterior}})$ with $\tau_{\text{posterior}} = 0.5$. Segmentation results show that most of lesions are detected. Although some lesions are not detected, e.g. peri-ventricular lesion in slice 38, they appear enhanced in the posterior map. Moreover the segmentations of slices 38 and 42 show peri-ventricular regions, visually very similar to MS lesions, but not delineated in the ground truth.
Figure 2. Two classification layers are used for left ventricle segmentation with random forests. (1) first layer posterior probability is used as a weight map for context aware MRI intensity standardization and cardiac pose normalization, (2) second layer is then used for a more confident left ventricle segmentation, (3) second level posterior probability isocontour overlay, (4) volumetric visualisation of the obtained segmentation.

In the context of radiotherapy of the head and neck, we propose different strategies to design anatomical atlases and we compare their performances for automatic segmentation [28]:

- We investigate average atlas construction, atlas stratification and patient-specific strategies based on the selection and fusion of the most appropriate atlases for each patient. We compared global, regional and local selection and fusion of the atlases.
- We show that the proposed patient-specific strategies enable to significantly improve the quality of the automatic segmentation in comparison with average atlas strategies. Visual results are presented on figure 3.
- We evaluated the proposed algorithms in two different contexts: segmentation of the lymph node levels and the organs at risk for radiotherapy planning, and segmentation of the teeth for post-irradiation dental care management. Automatic segmentations of the teeth are shown on figure 4.

5.2. Biological Image Analysis

5.2.1. Pre-clinical molecular imaging: reconstruction of tumors in rodents with SPECT imaging

Participants: Marine Breuilly [Correspondant], Grégoire Malandain, Nicholas Ayache, Jacques Darcourt [CAL], Philippe Franken [CAL], Thierry Pourcher [CEA].

This work is jointly conducted with the Transporter in Imagery and Oncologic Radiotherapy team (TIRO, CEA-CAL-UNSA) located in Nice.
Figure 3. Atlas-based segmentation of the lymph node level II using the average atlas (blue contours on left figure) and using the atlas that is locally adapted to the patient’s anatomy (yellow contours on right figure), compared with the manual contours (red contours on both figures).

Figure 4. Automatic segmentation of the teeth using a multi-atlas framework (in green) in comparison with the manual segmentation (in blue). Both segmentations are represented on the upper jaw.
The coupled CT and SPECT device allows to image both the anatomy (with the CT) and physiology information targeted by a dedicated radio-pharmaceutical tracer (here the tumors, with the SPECT). However, tumor quantification is impaired by the respiratory motion that induces an artificial enlargement of the moving structures. We propose then to select all the motion-less phases from a 4D SPECT images to reconstruct a motion-free 3D image. In addition, we also propose to correct for the heterogeneity of the respiratory cycles by re-tagging the SPECT raw data.

The resulting 3D motionless gated image shows improvement of volume accuracy compared to the non gated SPECT image; and noise reduction compared to the 4D SPECT image (see Figure 5).

![Figure 5. Coronal slices of fused SPECT and CT images from a NOD-SCID mouse (data acquired with GE eXplore speCZT CT 120): central hot spots reveal intraperitoneal metastasis from adenocarcinoma of the colon (PROb-mNIS). Left to right: non gated, phase at end of expiration, phase at end of inhalation, and motionless gated.](image)

### 5.3. Computational Anatomy

#### 5.3.1. The Kernel Bundle framework: Sparse Multiscale Diffeomorphic Deformations

**Participants:** Stefan Sommer [DIKU], Mads Nielsen [DIKU], François Lauze [DIKU], Xavier Pennec [Correspondant].

*This work is performed in collaboration between DIKU (University of Copenhagen) and Asclepios (Inria). It was initiated during a 5 month visit of Stefan Sommer in 2011.*

In the analysis and modeling of anatomical deformations, we expect deformations to have both large and small scale components. However, we expect these large and small scale deformation to occur at different places. Thus, one would like to represent anatomical deformations with a small number of deformation atoms are different scales that are sparsely distributed across space and scale.

- In [71], we propose a multi-scale kernel bundle framework (LDDKB) that extends the LDDMM framework by incorporating multiple kernels at multiple scales in the registration. Experiments show that the method automatically adapts to the right scales, and it therefore removes the need for classical scale selection methods.

- In [72], we derive the Kernel Bundle EPDiff evolution equations, which provide optimal warps in this new framework.
5.3.2. Longitudinal modeling of the structural changes of the brain in Alzheimer’s disease.

Participants: Marco Lorenzi [Correspondant], Xavier Pennec, Giovanni Frisoni [IRCCS Fatebenefratelli Brescia, Italy], Nicholas Ayache.

This work is done in collaboration with LENITEM, IRCCS San Giovanni di Dio Fatebenefratelli, Brescia, Italy.

This work addresses the analysis and the quantification of the longitudinal structural changes of the brain affected by Alzheimer’s disease (AD). We propose a framework based on the non-rigid registration of brain MRIs using stationary velocity fields. In 2011, the main scientific developments were:

- Unbiased detection of the structural changes in the brain [60]. The method robustifies the Demons diffeomorphic registration by estimating and removing the multiplicative and additive intensity biases affecting the images.
- Definition of a population-based atlas for the longitudinal brain structural changes. In this work we proposed the Schild’s Ladder as a general method for parallel transporting the subject-specific longitudinal deformation trajectories in a reference space [61]. The work was awarded with the runner-up prize at the “Information Processing in Medical Imaging (IPMI)” conference in Irsee, Germany. Other transport methods from the Lie group theory have been successively proposed and investigated [62].
- Group-wise statistical analysis of the brain longitudinal changes in multiple time points [60]. The framework has been applied for the analysis of the longitudinal brain changes in healthy subjects at risk of AD, and the results showed an accelerated progression of atrophy for the subjects positive to the marker of Alzheimer Aβ_{42} (see Figure 6).

Finally, the above framework has been promoted to the neuroscience community as diagnostic tool and support for clinical trials [75].

5.3.3. Statistical Analysis of White Matter Fiber Bundles

Participants: Stanley Durrleman [Correspondant], Pierre Fillard [Parietal, INRIA Saclay], Xavier Pennec, Alain Trouvé [CMLA, ENS Cachan], Nicholas Ayache.

This work is an application of the generic morphometric method developed in 2009 to the statistical analysis of the shape and texture of white matter fiber bundles extracted from Diffusion Tensor Images (DTI).

- Registration, Atlas Estimation and Variability Analysis of White Matter Fiber Bundles Modeled as Currents [32].

5.3.4. Comparison of endocranial ontogenies in chimpanzees and bonobos

Participants: Stanley Durrleman [Correspondant], Xavier Pennec, Alain Trouvé [CMLA, ENS Cachan], Nicholas Ayache, José Braga [AMIS, Univ. Toulouse 3].

This work has been performed in the context of the INRIA collaborative project ARC 3D-Morphine (PI: Sylvain Prima, IRISA) and a follow-up collaboration with José Braga at Université Paul Sabatier, Toulouse.

This work quantifies ontogenetic differences between bonobo and chimpanzee endocrania, using dental development as a timeline. Synthetic endocranial trajectories are estimated from time series cross-sectional data. Then, differences in morphology and in rate of shape changes is quantified using the spatiotemporal registration introduced in 2009.

- Comparison of the endocranial ontogenies between chimpanzees and bonobos via temporal regression and spatiotemporal registration.
Figure 6. Top. Average longitudinal brain atrophy for the healthy subjects at risk for Alzheimer’s disease (Aβ42 positive). Bottom. Annual percentage differential evolution modelled for the Aβ42 positive group with respect to the healthy aging. The analysis shows an increased ventricular expansion and the matter loss in the cortex and in the temporal areas.
Figure 7. The proposed method estimates a variability model from the white matter fiber bundles extracted from six subjects (a-original data, b- same data approximated at the current resolution). The model allows us to synthesize artificial bundles that reproduce the variability in shape and in fiber density estimated from the original data (c). Image taken from [32]

Figure 8. Typical endocranial trajectory of bonobos (top) and chimpanzees (bottom) estimated from time series cross-sectional surface data (59 chimpanzees and 60 bonobos)
Figure 9. Graph of time-wrap (magenta curve) putting into correspondence the developmental stages of the bonobos to that of the chimpanzees. The function measures the differences in rate of shape changes over time between the endocranial trajectories estimated in Fig. 8. This shows that the bonobos endocranial ontogeny is retarded by a factor 0.25 compared to that of the chimpanzees.

5.3.5. Statistical Modelling of Cardiac Growth and Deformation from Medical Images

Participants: Kristin McLeod [Correspondant], Tommaso Mansi, Adityo Prakosa, Maxime Sermesant, Xavier Pennec.

Parts of this work were performed within the framework of the EU project Care4me ITEA2, and the INRIA ARC Sirap, in collaboration with St Thomas Hospital, King’s College London, the REO team from INRIA Rocquencourt and Necker Paediatric Hospital in Paris.

This work builds on the statistical analysis framework for surfaces developed by Durrleman and Mansi in 2009.

- The iLogDemons motion tracking algorithm of Mansi et al. [37] was applied to a data-set of 15 subjects and 1 phantom each with a cine-MR, tagged-MR and echocardiography sequence as a part of the STACOM workshop challenge at MICCAI 2011 [64]. The paper received the Best Paper Award for the Motion Challenge.
- The work of Mansi et al for a statistical model of cardiac growth in the right ventricle [38] was extended to the left ventricle to obtain a bi-ventricular cardiac growth model for patients with repaired tetralogy of Fallot (see Fig. 10).
- The preliminary analysis of a statistical model for reduced blood flow simulations in the pulmonary artery proposed in 2010 is currently being extended to a journal version to further analyse the method on a larger data-set.

5.3.6. Statistical Modeling of Shapes Using Trees of Locally Affine Transformations

Participants: Christof Seiler [Correspondant], Xavier Pennec, Mauricio Reyes [Institute for Surgical Technology and Biomechanics, University of Bern, Switzerland].
Figure 10. Mean growth model computed from a population of 13 repaired tetralogy of Fallot patients. Both ventricles grow as body surface area (used as an index of growth) increases.

This work is performed in the context of the joint PhD of Christof Seiler at the Institute for Surgical Technology and Biomechanics, University of Bern, Switzerland and Asclepios INRIA.

The goal of this work is to analyze anatomical shapes through deformations defined with few but important and intelligible parameters. Advances towards this goal were the following in 2011.

- Fusion of the Log-demons registration and the Log-Euclidean polyaffine framework. The results of the new registration method applied to femur CT’s was presented at SPIE [69].
- Decomposition of diffeomorphic deformations into a tree of locally affine transformations applied to mandible CT’s. This work won the young scientist award at MICCAI 2011 in Toronto [68].

5.3.7. Atlas of Cardiac Fiber Architecture from DT-MRI

Participants: Hervé Lombaert, Hervé Delingette [Correspondant], Nicholas Ayache, Jean-Marc Peyrat, Pierre Croisille.

This work is a collaboration between Creatis, Lyon and INRIA Sophia Antipolis, including members of Ecole Polytechnique of Montreal. Financial support is partly from the National Science and Engineering Research Council of Canada, and the Michael Smith Foreign Study Program.

The variability of the cardiac fiber architecture has been investigated in a human population. An automatic method has been developed to construct an atlas of DTMRI images. A statistical analysis has been carried on using the Log-Euclidean metric.

- An automatic method has been developed to construct an atlas of DTMRI images. The first human atlas of DTMRI has been build [56]. This work received the Best Paper Award at the FIMH 2011 conference in New York City.
- Results have also been published in [59] where the variability of the cardiac architecture is measured using the Log-Euclidean metric.
The variability of the cardiac laminar sheets in human DTMRI has been studied [58].

Differences in a population of normal and abnormal hearts have been investigated [57].

5.4. Computational Physiology

5.4.1. Tumor Growth Modeling

Participants: Erin Stretton [Correspondant], Nicholas Ayache, Hervé Delingette, Bjoern Menze, Ender Konukoglu, Ezequiel Geremia, Emmanuel Mandonnet.

This work was funded by Care4me.

- Performing a sensitivity analysis of long time series of multi-modal data.
- Improving our current models and their inputs, including creating a white matter and brain mask template with the help of a neurosurgeon and integrating the Powell bound constrained optimization into the minimization routine.
- Performed comparison of using a patient DTI, an atlas DTI or no DTI at all showing the difference in accuracies in the simulation results since the patient DTI is not always available in a clinical setting or is of poor quality. We found that using an atlas DTI produced only slightly less accurate results than using a patient DTI, whereas using no DTI at all did not produce accurate results.

5.4.2. Synthetic Echocardiographic Image Sequences for Cardiac Inverse Electro-Kinematic Learning

Participants: Adityo Prakosa [Correspondant], Maxime Sermesant, Hervé Delingette, Eric Saloux [CHU Caen], Pascal Allain [Philips Healthcare], Pascal Cathier [Philips Healthcare], Patrick Etyngier [Philips Healthcare], Nicolas Villain [Philips Healthcare], Nicholas Ayache.
Figure 12. Tracked fibers from the human atlas of DTMRI of the myocardium.
This work is done in collaboration with Medisys, Philips Healthcare Suresnes, France, and Cardiology Department of CHU Caen, France.

- A database of 120 synthetic 3D echocardiography (US) sequences is created based on a cardiac electromechanical model (see Figure 13).
- Kinematic descriptors are extracted from the displacement field estimated from the synthetic 3D US sequence using the iLogDemons non-rigid registration method [37].
- Cardiac inverse electro-kinematic learning is done by using the database of synthetic 3D US sequences in order to estimate the cardiac depolarization times for the given kinematic descriptors [67]. First evaluation on two clinical sequences from patients with Left Bundle Branch Block shows encouraging results.

Figure 13. Synthetic 3D US. (1) original real image with (2) contour of the mesh at the corresponding time from the model simulation overlayed, (3) synthetic image generated with the model simulation with model contour overlay, (4) synthetic image.

5.4.3. Prediction of patient-specific Ventricular Tachycardia for radio-frequency ablation therapy planning

Participants: Jatin Relan [Correspondant], Maxime Sermesant, Hervé Delingette, Nicholas Ayache.

This work is funded by the FP7 European Project euHeart.

In this work, we build a patient-specific cardiac electrophysiology (EP) model derived from hybrid XMR imaging and non-contact electro-anatomical mapping procedure on a patient with heart failure. The model is then used to predict patient-specific arrhythmias, such as induced ischemic Ventricular Tachycardia (VT) (Fig. 14) and leads in generation and evaluation of patient-specific VT circuits, with critical exit points for Radio Frequency (RF) ablation. These predictions are now validated with some clinical cases, with electrophysiology mapping of induced VT in patients undergoing the clinical VT-Stim procedures (Fig. 14).

5.4.4. Real-time simulation of catheter ablation in the context of cardiac arrhythmia

Participants: Hugo Talbot [Correspondant], Federico Spadoni, Maxime Sermesant, Hervé Delingette, Stephane Cotin.
Figure 14. (a & b) Analysis of electroanatomical mapping data, to personalise the depolarisation and repolarisation wavefront dynamics. (c) Anatomical data personalisation along with ischemic regions (red). (d) A VT inducibility map showing regions with high probability of inducing VT with a VT-Stim procedure. (e) Critical exit point map showing the most eligible regions for RF ablation success. (f & g) An example of one of the various induced VT circuits, with VT-Stim model prediction. (h & i) Induced VT circuit in a clinical case (red = exit point).

This work is performed in the context of the euHeart project and the PhD of Hugo Talbot in collaboration with the Shacra (INRIA, Lille Nord Europe) team.

- This work aims at simulating in real-time the endovascular procedure of radiofrequency ablation of the left ventricle for patient suffering from Ventricular Tachycardia.
- Fast simulation of electrophysiology has been reached with a Eikonal model[40].
- Use the SOFA platform for simulating endovascular navigation and cardiac electrophysiology.

### 5.4.5. Personalized model of the heart for cardiac therapy planning

**Participants:** Stéphanie Marchesseau [Correspondant], Ken C.L. Wong, Hervé Delingette, Maxime Serum-sant, Nicholas Ayache.

This work has been performed in the context of the euHeart european project.

- We improved the existing electromechanical model of the heart (see Fig 16) to include mechanical non linearity, viscosity and strain rate dependent contractility. It was implemented in SOFA with a new four valves model to deal with the cardiac phases and enforce isovolumetric constraint.
- We have obtained first personalization of cardiac mechanics from 3 cine-MRI cases using a variational approach (adjoint method)[54].
Figure 15. (Left) Simulation of catheterization with a force-feedback device; Simulation of cardiac electrophysiology.

Figure 16. (Left) Electromechanical model of the heart that is coupled with medical images; (Right) Pressures and volumes curves of the left ventricle resulting from the simulation of one cardiac cycle.
6. New Results

6.1. Computational Diffusion MRI

This sub-theme is dedicated to describe our various contributions performed within the framework of Computational Diffusion MRI. In 6.1.1, we start by presenting our contributions to improving dMRI signal and optimize dMRI acquisition schemes. Then, we present our modeling contributions related to the problem of reconstructing and characterizing important Diffusion MRI features such as the Orientation Distribution Function (ODF) and the Ensemble Average Propagator (EAP) in 6.1.2, including contributions of the compressed sensing theory to dMRI and contributions to online motion detection. Finally, we end up, in 6.1.3, with some general applications such as tractography, clustering and microstructures recovery with pore size distribution estimation.

6.1.1. Improving dMRI Signal and Acquisitions

6.1.1.1. Optimal Design of Multiple Q-shells experiments for Diffusion MRI

Participants: Rachid Deriche, Emmanuel Caruyer, Iman Aganj [Department of Electrical and Computer Engineering, University of Minnesota], Christophe Lenglet [Department of Electrical and Computer Engineering, University of Minnesota], Guillermo Sapiro [Department of Electrical and Computer Engineering, University of Minnesota], Jian Cheng [ATHENA and LIAMA, China], Jiang Tianzi [LIAMA, China].

This work was partly supported by the CD-MRI Associate Team.

Recent advances in diffusion MRI make use of the diffusion signal sampled on the whole Q-space, rather than on a single sphere. While much effort has been done to design uniform sampling schemes for single shell experiment, it is yet not clear how to build a strategy to sample the diffusion signal in the whole Fourier domain. In this work, we proposed a method to generate acquisition schemes for multiple Q-shells experiment in diffusion MRI. The acquisition protocols we designed are incremental, which means they remain approximately optimal when truncated before the acquisition is complete. Our method is fast, incremental, and we can generate diffusion gradients schemes for any number of acquisitions, any number of shells, and any number of points per shell. The samples arranged on different shells do not share the same directions. The method has been tested for Spherical Polar Fourier reconstruction of the diffusion signal, and is based on Monte-Carlo simulations. Several preferred acquisition parameters are identified.

This work has been published in [20].

6.1.1.2. Incremental gradient table for multiple Q-shells diffusion MRI

Participants: Rachid Deriche, Emmanuel Caruyer, Iman Aganj [Department of Electrical and Computer Engineering, University of Minnesota], Christophe Lenglet [Department of Electrical and Computer Engineering, University of Minnesota], Guillermo Sapiro [Department of Electrical and Computer Engineering, University of Minnesota].

This work was partly supported by the CD-MRI Associate Team.

Most studies on sampling optimality for diffusion MRI deal with single Q-shell acquisition. For single Q-shell acquisition, incremental gradient table has proved useful in clinical setup, where the subject is likely to move, or for online reconstruction. In this work, we proposed a generalization of the electrostatic repulsion to generate gradient tables for multiple Q-shells acquisitions, designed for incremental reconstruction or processing of data prematurely aborted.

This work has been published in [21].

6.1.1.3. Impact of radial and angular sampling on multiple shells acquisition in diffusion MRI

Participants: Rachid Deriche, Sylvain Merlet, Emmanuel Caruyer.
In this work, we evaluated the impact of radial and angular sampling on multiple shells (MS) acquisition in diffusion MRI. The validation of our results is based on a new and efficient method to accurately reconstruct the Ensemble Average Propagator (EAP) in term of the Spherical Polar Fourier (SPF) basis from very few diffusion weighted magnetic resonance images (DW-MRI). This approach nicely exploited the duality between SPF and a closely related basis in which one can respectively represent the EAP and the diffusion signal using the same coefficients. We efficiently combined this relation to the recent acquisition and reconstruction technique called Compressed Sensing (CS). Based on results of multi-tensors models reconstruction, we showed how to construct a robust acquisition scheme for both neural fibre orientation detection and attenuation signal/EAP reconstruction.

This work has been published in [32].

6.1.1.4. Simultaneous Smoothing and Estimation of DTI via Robust Variational Non-local Means

Participants: Rachid Deriche, Meizhu Liu [Department of CISE, University of Florida, Gainesville, USA], Baba Vemuri [Department of CISE, University of Florida, Gainesville, USA].

Regularized diffusion tensor estimation is an essential step in DTI analysis. There are many methods proposed in literature for this task but most of them are neither statistically robust nor feature preserving denoising techniques that can simultaneously estimate symmetric positive definite (SPD) diffusion tensors from diffusion MRI. One of the most popular techniques in recent times for feature preserving scalar-valued image denoising is the non-local means filtering method that has recently been generalized to the case of diffusion MRI denoising. However, these techniques denoise the multi-gradient volumes first and then estimate the tensors rather than achieving it simultaneously in a unified approach. Moreover, some of them do not guarantee the positive definiteness of the estimated diffusion tensors. In this work, we proposed a novel and robust variational framework for the simultaneous smoothing and estimation of diffusion tensors from diffusion MRI. Our variational principle makes use of a recently introduced total Kullback-Leibler (tKL) divergence, which is a statistically robust similarity measure between diffusion tensors, weighted by a non-local factor adapted from the traditional non-local means filters. For the data fidelity, we use the nonlinear least-squares term derived from the Stejskal-Tanner model. We have performed experimental results depicting the positive performance of our method in comparison to competing methods on synthetic and real data examples.

This work has been published in [31].

6.1.1.5. Anisotropic LMMSE denoising of MRI based on statistical tissue models

Participants: Rachid Deriche, Gonzalo Vegas-Sánchez-Ferrero [Universidad de Valladolid, Spain], Santiago Aja Fernández [Universidad de Valladolid, Spain].

Linear Minimum Mean Squared Error Estimation (LMMSE) is a simple, yet powerful denoising technique within MRI. It is based on the computation of the mean and variance of the data being filtered according to a noise model assumed, which is usually accomplished by calculating local moments over squared neighborhoods. When these neighborhoods are centered in pixels corresponding to image contours, the estimation is not accurate due to the presence of two or more tissues with different statistical properties. In this work, we overcome this limitation by introducing an anisotropic LMMSE scheme: the grey levels of each tissue in the MRI volume are modeled as a Gamma-mixture, such that we can discriminate between the different matters to construct anisotropic neighborhoods containing only one kind of tissue. The potential of the Gamma distribution relies on its ability to fit both the Rician distribution traditionally used to model the noise in MRI and the non-central Chi noise found in modern parallel MRI systems.

This work is currently under submission.

6.1.2. Modeling in Diffusion MRI

6.1.2.1. Multiple q-Shell Diffusion Propagator Imaging

Participants: Rachid Deriche, Maxime Descoteaux [Sherbrooke University, Quebec], Denis Le Bihan [NeuroSpin, IFR 49 CEA Saclay], Jean-François Mangin [NeuroSpin, IFR 49 CEA Saclay], Cyril Poupon [NeuroSpin, IFR 49 CEA Saclay].
Many recent high angular resolution diffusion imaging (HARDI) reconstruction techniques have been introduced to infer an orientation distribution function (ODF) of the underlying tissue structure. These methods are more often based on a single-shell (one b-value) acquisition and can only recover angular structure information contained in the ensemble average propagator (EAP) describing the three-dimensional (3D) average diffusion process of water molecules. The EAP can thus provide richer information about complex tissue microstructure properties than the ODF by also considering the radial part of the diffusion signal. In this work, we presented a novel technique for analytical EAP reconstruction from multiple q-shell acquisitions. The solution is based on a Laplace equation by part estimation between the diffusion signal for each shell acquisition. This simplifies greatly the Fourier integral relating diffusion signal and EAP, which leads to an analytical, linear and compact EAP reconstruction. An important part of this work is dedicated to validate the diffusion signal estimation and EAP reconstruction on real datasets from ex vivo phantoms. We also illustrated multiple q-shell diffusion propagator imaging (mq-DPI) on a real in vivo human brain and performed a qualitative comparison against state-of-the-art diffusion spectrum imaging (DSI) on the same subject. mq-DPI is shown to reconstruct robust EAP from only several different b-value shells and less diffusion measurements than DSI. This opens interesting perspectives for new q-space sampling schemes and tissue microstructure investigation.

This work has been published in [13].

6.1.2.2. A Riemannian Framework for Ensemble Average Propagator Computing

Participants: Rachid Deriche, Jian Cheng [ATHENA and LIAMA, China], Aurobrata Ghosh, Jiang Tianzi [LIAMA, China].

This work was partly supported by the Association France Parkinson and the ANR NucleiPark project.

In Diffusion Tensor Imaging (DTI), Riemannian framework (RF) has been proposed for processing tensors, which is based on Information Geometry theory. Many papers have shown that RF is useful in tensor estimation, interpolation, smoothing, regularization, segmentation and so on. Recently RF also has been proposed for Orientation Distribution Function (ODF) computing and it is applicable to any Probability Density Function (PDF) based on any orthonormal basis representation. Spherical Polar Fourier Imaging (SPFI) was proposed recently to fast and robustly estimate the ODF and Ensemble Average Propagator (EAP) from arbitrary sampled DWI signals. In this work, we proposed the RF for EAPs and implemented it via SPFI. We proved that the RF for EAPs is diffeomorphism invariant, which is the natural extension of affine invariant RF for tensors. It could avoid the so-called swelling effect for interpolating EAPs, just like the RF for tensors. We also proposed the Log-Euclidean framework (LEF), Affine-Euclidean framework (AEF), for fast processing EAPs, and Geometric Anisotropy (GA) for measuring the anisotropy of EAPs, which are all the extensions of previous concepts in RM for tensors respectively.

This work has been published in [22].

6.1.2.3. Diffeomorphism Invariant Riemannian Framework for Ensemble Average Propagator Computing

Participants: Rachid Deriche, Jian Cheng [ATHENA and LIAMA, China], Aurobrata Ghosh, Jiang Tianzi [LIAMA, China].

This work was partly supported by the Association France Parkinson and the ANR NucleiPark project.

In Diffusion Tensor Imaging (DTI), Riemannian framework based on Information Geometry theory has been proposed for processing tensors on estimation, interpolation, smoothing, regularization, segmentation, statistical test and so on. Recently Riemannian framework has been generalized to Orientation Distribution Function (ODF) and it is applicable to any Probability Density Function (PDF) under orthonormal basis representation. Spherical Polar Fourier Imaging (SPFI) was proposed for ODF and Ensemble Average Propagator (EAP) estimation from arbitrary sampled signals without any assumption. Tensors only can represent Gaussian EAP and ODF is the radial integration of EAP, while EAP has full information for diffusion process. To our knowledge, so far there is no work on how to process EAP data. In this work, we presented a Riemannian framework as a mathematical tool for such task. We proposed a state-of-the-art Riemannian framework for EAPs by representing the square root of EAP, called wavefunction based
on quantum mechanics, with the Fourier dual Spherical Polar Fourier (dSPF) basis. In this framework, the exponential map, logarithmic map and geodesic have closed forms, and weighted Riemannian mean and median uniquely exist. We analyzed theoretically the similarities and differences between Riemannian frameworks for EAPs and for ODFs and tensors. The Riemannian metric for EAPs is diffeomorphism invariant, which is the natural extension of the affine-invariant metric for tensors. We proposed Log-Euclidean framework to fast process EAPs, and Geodesic Anisotropy (GA) to measure the anisotropy of EAPs. With this framework, many important data processing operations, such as interpolation, smoothing, atlas estimation, Principal Geodesic Analysis (PGA), can be performed on EAP data. The proposed Riemannian framework was validated in synthetic data for interpolation, smoothing, PGA and in real data for GA and atlas estimation. Riemannian median is much robust for atlas estimation.

This work has been published in [23].

6.1.2.4. Theoretical Analysis and Practical Insights on EAP Estimation via a Unified HARDI Framework

Participants: Rachid Deriche, Jian Cheng [ATHENA and LIAMA, China], Jiang Tianzi [LIAMA, China].

This work was partly supported by the Association France Parkinson and the ANR NucleiPark project.

Since Diffusion Tensor Imaging (DTI) cannot describe complex non-Gaussian diffusion process, many techniques, called as single shell High Angular Resolution Diffusion Imaging (sHARDI) methods, reconstruct the Ensemble Average Propagator (EAP) or its feature Orientation Distribution Function (ODF) from diffusion weighted signals only in single shell. Q-Ball Imaging (QBI) and Diffusion Orientation Transform (DOT) are two famous sHARDI methods. However, these sHARDI methods have some intrinsic modeling errors or need some unreal assumptions. Moreover they are hard to deal with signals from different q-shells. Most recently several novel multiple shell HARDI (mHARDI) methods, including Diffusion Propagator Imaging (DPI), Spherical Polar Fourier Imaging (SPFI) and Simple Harmonic Oscillator based Reconstruction and Estimation (SHORE), were proposed to analytically estimate EAP or ODF from multiple shell (or arbitrarily sampled) signals. These three methods all represent diffusion signal with some basis functions in spherical coordinate and use plane wave formula to analytically solve the Fourier transform. To our knowledge, there is no theoretical analysis and practical comparison among these sHARDI and mHARDI methods. In this work, we proposed a unified computational framework, named Analytical Fourier Transform in Spherical Coordinate (AFT-SC), to perform such theoretical analysis and practical comparison among all these five state-of-the-art diffusion MRI methods. We compared these five methods in both theoretical and experimental aspects. With respect to the theoretical aspect, some criteria are proposed for evaluation and some differences together with some similarities among the methods are highlighted. Regarding the experimental aspect, all the methods were compared in synthetic, phantom and real data. The shortcomings and advantages of each method were highlighted from which SPFI appears to be among the best because it uses an orthonormal basis that completely separates the spherical and radial information.

This work has been published in [24].

6.1.2.5. Compressive Sensing Ensemble Average Propagator Estimation via L1 Spherical Polar Fourier Imaging

Participants: Rachid Deriche, Sylvain Merlet, Emmanuel Caruyer, Jian Cheng [ATHENA and LIAMA, China], Jiang Tianzi [LIAMA, China].

In diffusion MRI (dMRI) domain, many High Angular Resolution Diffusion Imaging (HARDI) methods were proposed to estimate Ensemble Average Propagator (EAP) and Orientation Distribution Function (ODF). They normally need many samples, which limits their applications. Some Compressive Sensing (CS) based methods were proposed to estimate ODF in Q-Ball Imaging (QBI) from limited samples. However EAP estimation is much more difficult than ODF in QBI. Recently Spherical Polar Fourier Imaging (SPFI) was proposed to represent diffusion signal using Spherical Polar Fourier (SPF) basis without specific assumption on diffusion signals and analytically obtain EAP and ODF via the Fourier dual SPF (dSPF) basis from arbitrarily sampled signal. Normally the coefficients of SPF basis are estimated via Least Square with weighted L2 norm regularization (L2-SPFI). However, L2-SPFI needs a truncated basis to avoid overfitting, which brings some estimation errors. By considering the Fourier relationship between EAP and signal and the Fourier basis pair provided in SPFI, we proposed a novel EAP estimation method, named L1-SPFI, to estimate
EAP from limited samples using CS technique, and favorably compared it to the classical L2-SPFI method. L1-SPFI estimates the coefficients in SPFI using least square with weighted L1 norm regularization. The weights are designed to enhance the sparsity. L1-SPFI significantly accelerates the ordinary CS based Fourier reconstruction method. This is performed by using SPF basis pair in CS estimation process which avoids the numerical Fourier transform in each iteration step. By considering high order basis in L1 optimization, L1-SPFI improves EAP reconstruction especially for the angular resolution. The proposed L1-SPFI was validated by synthetic, phantom and real data. The CS EAP and ODF estimations are discussed in detail and we showed that recovering the angular information from CS EAP requires much less samples than exact CS EAP reconstruction. Various experiments on synthetic, phantom and real data validate the fact that SPF basis can sparsely represent DW-MRI signals and L1-SPFI largely improves the CS EAP reconstruction especially the angular resolution.

This work has been published in [25], [26].

6.1.2.6. Spherical Polar Fourier EAP and ODF Reconstruction via Compressed Sensing in Diffusion MRI

Participants: Rachid Deriche, Sylvain Merlet, Aurobrata Ghosh, Jian Cheng [ATHENA and LIAMA, China].

In diffusion magnetic resonance imaging (dMRI), the Ensemble Average Propagator (EAP), also known as the propagator, describes completely the water molecule diffusion in the brain white matter without any prior knowledge about the tissue shape. In this work, we described a new and efficient method to accurately reconstruct the EAP in terms of the Spherical Polar Fourier (SPF) basis from very few diffusion weighted magnetic resonance images (DW-MRI). This approach exploits the duality between SPF and a closely related basis in which one can respectively represent the EAP and the diffusion signal using the same coefficients, and efficiently combines it to the recent acquisition and reconstruction technique called Compressed Sensing. Our work provides an efficient analytical solution to estimate, from few measurements, the diffusion propagator at any radius. We also provide a new analytical solution to extract an important feature characterising the tissue microstructure: the Orientation Distribution Function (ODF). We illustrate and prove the effectiveness of our method in reconstructing the propagator and the ODF on both noisy multiple q-shell synthetic and phantom data.

This work has been published in [33].

6.1.2.7. On Line Reconstruction and Motion Detection in HARDI

Participants: Rachid Deriche, Emmanuel Caruyer, Iman Aganj [Department of Electrical and Computer Engineering, University of Minnesota], Christophe Lenglet [Department of Electrical and Computer Engineering, University of Minnesota], Guillermo Sapiro [Department of Electrical and Computer Engineering, University of Minnesota].

This work was partly supported by the CD-MRI Associate Team.

With acquisition protocols such as high angular resolution diffusion imaging, head motion can become an issue. Although the misalignment between diffusion-weighted images (DWIs) can be corrected in a post-processing step, this might increase partial volume effects, because of the relatively low spatial resolution of DWIs and interpolation in the registration procedure. If able to detect motion online, the scanner technician could be issued a warning and make a decision accordingly. Orientation distribution functions (ODF) can be reconstructed online using a Kalman filter (KF). In this work, we presented three contributions related to the problem of online ODF reconstruction and motion detection in HARDI. First, we developed a proper error propagation accounting for the non-linear transform on the diffusion signal. Next, we developed two motion detection algorithms, based on the monitoring of residuals, and compared them using synthetic data.

This work has been published in [18].

6.1.2.8. Online Motion Detection in High Angular Resolution Diffusion Imaging

Participants: Rachid Deriche, Emmanuel Caruyer, Iman Aganj [Department of Electrical and Computer Engineering, University of Minnesota], Christophe Lenglet [Department of Electrical and Computer Engineering, University of Minnesota], Guillermo Sapiro [Department of Electrical and Computer Engineering, University of Minnesota].
The orientation distribution function (ODF) can be reconstructed online incrementally from diffusion-weighted MRI with a Kalman filtering framework. This online reconstruction can provide real-time feedback to the practitioner, especially appreciated for long acquisition protocols typical in Q-ball imaging. On top of the Kalman filter, we proposed a method to evaluate online the reconstruction accuracy of the estimated ODF in constant solid angle. In addition, monitoring the residuals of the Kalman filter, we designed, based on statistical tests, two algorithms for online detection of subject motion. The proposed techniques, tested on real and synthetic data under various experimental conditions, can detect rotation by angle less than 3°.

This work has been published in [19].

6.1.3. From DW-MRI to Fiber Pathways and Microstructures Recovery

6.1.3.1. A Polynomial Approach for Maxima Extraction and Its Application to Tractography in HARDI

Participants: Rachid Deriche, Aurobrata Ghosh, Demian Wassermann [Harvard Medical School].

A number of non-parametrically represented High Angular Resolution Diffusion Imaging (HARDI) spherical diffusion functions have been proposed to infer more and more accurately the heterogeneous and complex tissue microarchitecture of the cerebral white-matter. These spherical functions overcome the limitation of Diffusion Tensor Imaging (DTI) at discerning crossing, merging and fanning axonal fiber bundle configurations inside a voxel. Tractography graphically reconstructs the axonal connectivity of the cerebral white-matter in vivo and non-invasively, by integrating along the direction indicated by the local geometry of the spherical diffusion functions. Tractography is acutely sensitive to the local geometry and its correct estimation. In this work, we first proposed a polynomial approach for analytically bracketing and numerically refining with high precision all the maxima, or fiber directions, of any spherical diffusion function represented non-parametrically. This permits an accurate inference of the fiber layout from the spherical diffusion function. Then we proposed an extension of the deterministic Streamline tractography to HARDI diffusion functions that clearly discern fiber crossings. We also extended the Tensorline algorithm to these HARDI functions, to improve on the extended Streamline tractography. We illustrated our proposed methods using the Solid Angle diffusion Orientation Distribution Function (ODF-SA). We presented results on multi-tensor synthetic data, and real in vivo data of the cerebral white-matter that show markedly improved tractography results.

This work has been published in [30].

6.1.3.2. Tract-based statistical analyzes in dMRI in autism spectrum disorder

Participants: Rachid Deriche, Anne-Charlotte Philippe, Demian Wassermann [Harvard Medical School, Boston, MA], Pablo Barttfeld [Integrative Neuroscience Laboratory, Physics Dept. University of Buenos Aires, Buenos Aires, Argentina], Jorge Calvar [Fundación parala Luchacontralas Enfermedades Neurológicasdel.Infancia, Buenos Aires, Argentina], Ramon Leiguarda [Fundación para la Lucha contra las Enfermedades Neurológicasdel.Infancia, Buenos Aires, Argentina], Bruno Wicker [INCM CNRS, Marseille, France], Mariano Sigman [Integrative Neuroscience Laboratory, Physics Dept. University of Buenos Aires, Buenos Aires, Argentina].

This work was partly supported by the ECOS-Sud grant.

Abnormal face processing is one of the hallmark features of social impairments in autism spectrum disorder (ASD). Previous neuroimaging studies showed that the fusiform gyrus is involved in face perception and is not or abnormally activated in autistic subjects. The aim of this study was to quantify potential anatomical differences in the white matter tracts that traverse the fusiform gyrus, the prefrontal cortex and the superior temporal gyrus, and correlate them with ADOS scores in ASD subjects. We used Diffusion Tensor MRI (DT) images to assess the integrity of automatically segmented white matter bundles connecting these brain areas. Then, we performed statistical analysis on these fiber bundles using diffusivity measures calculated from DT to characterize tissue microstructure changes and correlate these changes with ADOS scores. 7 adults with high functioning autism or Asperger syndrome and 11 typical adults participated in the study. We found several clusters with dissimilarities between ASD and control subjects in FA measures on tracts traversing the fusiform gyrus in both hemispheres of the brain. We observed a significant reduction of FA values in a cluster on a bundle joining the superior temporal gyrus to the prefrontal.
This work has been published in [37]. A related work on large-scale network analysis reflecting big-world characteristics in ASD has been published in [35].

6.1.3.3. Unsupervised automatic white matter fiber clustering using a Gaussian mixture model

**Participants:** Rachid Deriche, Meizhu Liu [Department of CISE, University of Florida, Gainesville, USA], Baba Vemuri [Department of CISE, University of Florida, Gainesville, USA].

Fiber tracking from diffusion tensor images is an essential step in numerous clinical applications. There is a growing demand for an accurate and efficient framework to perform quantitative analysis of white matter fiber bundles. In this work, we proposed a robust framework for fiber clustering. This framework is composed of two parts: accessible fiber representation, and a statistically robust divergence measure for comparing fibers. Each fiber is represented using a Gaussian mixture model (GMM), which is the linear combination of Gaussian distributions. The dissimilarity between two fibers is measured using the total square loss function between their corresponding GMMs (which is statistically robust). Finally, we performed the hierarchical total Bregman soft clustering algorithm on the GMMs, yielding clustered fiber bundles. Further, our method is able to determine the number of clusters automatically. We performed experimental results depicting favorable performance of our method on both synthetic and real data examples.

This work is currently under submission.

6.1.3.4. Riemannian geometry based brain white matter fiber clustering

**Participants:** Rachid Deriche, Ali Demir [Sabanci University, Istanbul, Turkey], Gozde Unal [Sabanci University, Istanbul, Turkey].

This work was partly supported by the PHC Bosphore grant.

Clustering of reconstructed brain white matter fibers into meaningful anatomical bundles becomes an important tool for detailed analysis of brain white matter diseases via diffusion MRI. In this work we developed a Riemannian geometry based geodesic distance measure between fiber pairs, which is then utilized in fiber clustering. A second contribution is a fiber selection algorithm, which compresses the dataset and speed up the computation time. We demonstrated methods on synthetic kissing fibers dataset, and validated on a brain white matter atlas.

This work is currently under submission.

6.1.3.5. White Matter Clustering Revisited

**Participants:** Rachid Deriche, Alvaro-Alejandro Sanchez-Moscosa.

This work was partly supported by the Association France Parkinson and the ANR NucleiPark project.

In [70], we have introduced an interesting hierarchical agglomerative based algorithm that represents and clusters white matter bundles under a Gaussian Process framework. In this work, a new implementation which drastically improves the performance of the clustering algorithm before mentioned is proposed and validated. This implementation notably improves the performance of the clustering algorithm [70] and has been validated on real data. This approach has the advantage of running in a lower abstraction level, which leads to lowered memory requirements, shorter running times and ultimately provides the possibility to process more densely seeded tractographies. The new implementation is then used to process densely seeded streamline tractographies which provide more localized fiber bundles. Additionally, as Parkinson's Disease is believed to induce changes in the axonal bundles, tract-based statistical analysis is performed on interesting fiber tracts to find significant differences in diffusion scalar measures.

6.1.3.6. Using Radial NMR Profiles to Characterize Pore Size Distributions

**Participants:** Rachid Deriche, John Treilhard [Queen’s University, CA].
Extracting information about axon diameter distributions in the brain is a challenging task which provides useful information for medical purposes; for example, the ability to characterize and monitor axon diameters would be useful in diagnosing and investigating diseases like amyotrophic lateral sclerosis (ALS) or autism. In [75], three families of operators are defined, whose action upon an NMR attenuation signal extracts the moments of the pore size distribution of the ensemble under consideration; also a numerical method is proposed to continuously reconstruct a discretely sampled attenuation profile using the eigenfunctions of the simple harmonic oscillator Hamiltonian – the SHORE basis. The work we have performed here extends this method to other bases that can offer a better description of attenuation signal behaviour – in particular, we proposed the use of the radial Spherical Polar Fourier (SPF) basis. Testing is performed to contrast the efficacy of the radial SPF basis and SHORE basis in practical attenuation signal reconstruction. The robustness of the method to additive noise is tested and analyzed. We demonstrated that a low-order attenuation signal reconstruction outperforms a higher-order reconstruction in subsequent moment estimation under noisy conditions. We proposed the simulated annealing algorithm for basis function scale parameter estimation. Finally, analytic expressions are derived and presented for the action of the operators on the radial SPF basis (obviating the need for numerical integration, thus avoiding a spectrum of possible sources of error).

This work is currently under submission.

6.2. Brain functional imaging using MEG/EEG

6.2.1. EEG forward problem

6.2.1.1. OpenMEEG software library

Participants: Maureen Clerc, Emmanuel Olivi, Alexandre Gramfort, Théodore Papadopoulo.

This work was partly supported by the Regional Council of Provence Alpes Cote d’Azur and the ANR ViMAGINE.

To recover the sources giving rise to electro- and magnetoencephalography in individual measurements, realistic physiological modeling is required, and accurate numerical solutions must be computed. The OpenMEEG software library solves the electromagnetic forward problem in the quasistatic regime, for head models with piecewise constant conductivity. The core of OpenMEEG consists of the symmetric Boundary Element Method, which is based on an extended Green Representation theorem. OpenMEEG is able to provide lead fields for four different electromagnetic forward problems: Electroencephalography (EEG), Magnetoencephalography (MEG), Electrical Impedance Tomography (EIT), and intracranial electric potentials (IPs). OpenMEEG is open source and multiplatform. It can be used from Python and Matlab in conjunction with toolboxes that solve the inverse problem; its integration within FieldTrip is operational since release 2.0.

Some new developments have concerned the organization of the computations to compute the lead fields. A lead-field matrix is obtained by concatenating the forward fields computed for thousands of sources characterized by their positions, orientations and strengths. A line of this lead-field matrix represents the physical quantity (potential for EEG, or some components of the magnetic field for MEG) at a sensor for each source. The number of sources largely exceeds the number of sensors (up to 256 electrodes for EEG, and less than 600 squids for MEG). When solving the forward problem with a BEM (Boundary Element Method), the lead-field matrix is generally computed column-by-column, i.e. source by source, which represents nsources resolutions of the forward problem. Using the adjoint operator of the forward problem, one can reduce the computations to sensors resolutions. Some previous works [72], [73] have used similar techniques for efficient computation of the lead-fields using finite element methods. The adjoint method [69] generalizes the Helmholtz reciprocity theorem and here is proposed its implementation using the BEM provided by the open-source software OpenMEEG

This work has been published in [15] and [36].

6.2.1.2. Conductivity calibration for the EEG forward problem

Participants: Maureen Clerc, Emmanuel Olivi, Alexandre Gramfort, Théodore Papadopoulo, Jean-Michel Badier [INSERM U751, La Timone, Marseille], Martine Gavaret [INSERM U751, La Timone, Marseille], Laurent Koessler [CRAN Nancy].
Bioelectric phenomena at low temporal frequency can be described by the electrostatic Poisson equation

$$\text{div} \, \sigma \nabla V = \text{div} \, J^p,$$

where \( J^p \) are primary current sources and \( \sigma \nabla V \) the Ohmic current. Appropriate boundary conditions (b.c.) must be set on the domain boundary, typically imposing the potential (Dirichlet b.c.) or the normal current (Neumann b.c.).

Solving this electrostatics equation for the electric potential is a problem common to different fields such as electroencephalography (EEG), electrocardiography (ECG), functional electrical stimulation. The main difficulty of the model concerns the conductivity field, which is not homogeneous, and whose values depend on the tissue type. Although the tissue structure can be revealed by imaging methods such as CT, Magnetic Resonance Imaging, diffusion Magnetic Resonance Imaging, conductivity values must nevertheless be assigned to the different tissues. In order to calibrate conductivity, injected current Electrical Impedance Tomography (EIT) can be used: it consists of injecting current on the outer boundary, and measuring the associated electric potential.

The thrust of OpenMEEG is to propose accurate forward problems, in several instances, including electro- and magneto-encephalography (EEG-MEG) and Electrical Impedance Tomography (EIT). OpenMEEG allows to compute the electric potential and magnetic fields due to boundary current injection or to sources within the compartments.

We apply this methodology to conductivity calibration in the context of high-density EEG pre-surgical epileptic exploration. High-density EEG measurements may be used to solve the inverse problem of source localization, in order to localize the foci of the epileptic activity within the brain. The solution of the inverse problem relies on a forward problem, linking sources to measurements. In turn, this forward problem is dependent on the conductivity of the head tissues.

Prior work has shown that the scalp-to-skull conductivity ratio is a sensitive parameter for source localization, because it has an influence on the depth of the estimated dipoles \([68]\). There have been several studies demonstrating the feasibility of injected-current Electrical Impedance Tomography to calibrate the scalp-to-skull conductivity ratio \([57]\,\,[49]\). We are currently collaborating with our partners from La Timone in Marseille in a clinical assessment of the use of injected current Electrical Impedance Tomography to calibrate the scalp-to-skull conductivity ratio.

This work has been presented at the international conference on Electrical Impedance Tomography, see \([27]\) and \([28]\).

### 6.2.1.3. Coupling numerical methods for the forward problem

**Participants:** Maureen Clerc, Emmanuel Olivi, Théodore Papadopoulo, Mariette Yvinec [Geometrica Project-Team, INRIA Sophia Antipolis Méditerranée].

**This work was partly supported by the ANR grant ViMAGINE.**

Electroencephalography (EEG) and magnetoencephalography (MEG) are two modalities that aim at measuring brain activity. Source localization from external data such EEG or MEG, requires a good understanding of the electromagnetic behavior of the patient head. Several models can be used, representing more or less complex geometrical shapes, and conductivity profiles. Different numerical methods allow to cope with different types of models: the Finite Element Method (FEM) can handle very general conductivity models, whereas the Boundary Element Method (BEM) is limited to piecewise constant conductivity. On the other hand, BEM is more capable than FEM to accurately represent sources in isotropic media.
Using a domain decomposition approach, we propose to independently use BEM or FEM in different sub-domains. In the EEG forward problem considered, the BEM is limited to the domain where the sources lie (the brain) while the other tissues are handled with the FEM. This leads to an accurate description of the sources while allowing for inhomogeneous and anisotropic conductivity. The proposed method is first validated against analytical solutions in multi-sphere models. Results of the forward problem are presented for a four-layer realistic head-model incorporating a burr-hole in the skull. Convergence of the iterative algorithm is analyzed numerically. The domain decomposition framework provides a way of taking the best advantage of both methods, thus significantly improving the accuracy in the resolution of the forward EEG problem, as well as time and memory consumption.

This work is part of Emmanuel Olivi’s PhD thesis [11], and a journal article is under submission.

6.2.1.4. White matter anisotropy

Participants: Maureen Clerc, Emmanuel Olivi, Alexandre Gramfort, Théodore Papadopoulo.

This work was partly supported by the ANR grant ViMAGINE.

Conductivity of tissues in the vicinity of the sources is especially influential on the MEG and EEG forward fields. Those tissues include white matter, whose conductivity is anisotropic because of its fiber structure. While white matter anisotropy can be measured thanks to Diffusion-Weighted MRI, it is rarely incorporated in MEG and EEG head models. Boundary Element Methods can only deal with piecewise constant conductivities, therefore ruling out white matter anisotropy that has a complex structure, and Finite Element Methods have been developed to deal with anisotropic conductivity, but require very fine meshes, thus huge linear systems. We have extended the BEM framework to incorporate white matter anisotropy by treating anisotropic conductivity as a perturbation of an isotropic one. With this extension we have been able to compute the influence of a fiber within the brain white matter on the electric potential, and to validate the results by comparison with a Finite Element Method.

This work has been published in [34].

6.2.2. Inverse problems in MEG and EEG

6.2.2.1. Rational Approximations

Participants: Maureen Clerc, Théodore Papadopoulo, Juliette Leblond [APICS Project Team, INRIA Sophia Antipolis Méditerranée], Jean-Paul Marmorat [Centre de Mathématiques Appliquées, Ecole des Mines].

In functional neuroimaging, a crucial problem is to localize active sources within the brain non-invasively, from the knowledge of the electromagnetic measurements outside the head. Identification of point sources from boundary measurements is an ill-posed inverse problem. In the case of electroencephalography (EEG), measurements are only available at electrode positions, the number of sources is not known in advance, and the medium within the head is inhomogeneous. We pursue our ongoing work on EEG source localization, based on rational approximation techniques in the complex plane. The method is used in the context of a nested sphere head model, in combination with a cortical mapping procedure. Results on simulated data prove the applicability of the method in the context of realistic measurement configurations.

This work has been submitted to a journal and published as a Research Report in [38].

6.2.3. Brain Computer Interfaces

6.2.3.1. New features for Motor Imagery

Participants: Maureen Clerc, Joan Fruitat, Théodore Papadopoulo, Eoin Thomas.

This work was partly supported by ANR grant CoAdapt.

Our goal is to build a training free BCI based on beta rebound detection and discrimination during the first stage of use, while the learning of the conventional sensorimotor rhythms is done. We show in this preliminary study that it is possible to use the beta rebound to discriminate, real and imagined, right versus left hand movement with either no or very little training.
6.2.3.2. A bandit algorithm for exploring mental imagery

**Participants:** Maureen Clerc, Joan Fruitet, Alexandra Carpentier [Sequel Project Team, INRIA Lille Nord Europe], Rémi Munos [Sequel Project Team, INRIA Lille Nord Europe].

_This work was partly supported by ANR grant Co-Adapt._

This study presents a new procedure to automatically select a discriminant motor task for an asynchronous brain-controlled button. This type of control pertains to Brain Computer Interfaces (BCI). When using sensorimotor rhythms in a BCI, several motor tasks, such as moving the right or left hand, the feet or the tongue, can be considered as candidates for the control. This report presents a method to select as fast as possible the most promising task. We develop for this purpose an adaptive algorithm UCB-classif based on the stochastic bandit theory and build an EEG experiment to test our method. By not wasting time on inefficient tasks, our algorithm can focus on the most promising ones, resulting in a faster task selection and a more efficient use of the BCI training session. This leads to better classification rates for a xed time budget, compared to a standard task selection.

This work has been published in [39] and is currently under submission to a journal.

6.3. Multi-Imaging Modalities

6.3.1. Coupling neuronal and haemodynamic models

6.3.1.1. Modeling of the neurovascular coupling in epileptic discharges

**Participants:** Maureen Clerc, Théodore Papadopoulo, Nicole Voges [former Athena postdoc], Christian Bénar [INSERM U751 Marseille], Solenna Blanchard [INSERM U642 Rennes], Fabrice Wendling [INSERM U642 Rennes], Habib Benali [INSERM U678 Paris], Olivier David [INSERM U594 Grenoble].

_This work was partly supported by ANR MultiModel._

Despite the interest in simultaneous EEG-fMRI studies of epileptic spikes, the link between epileptic discharges and their corresponding hemodynamic responses is poorly understood. In this context, biophysical models are promising tools for investigating the mechanisms underlying observed signals. We have applied a metabolic-hemodynamic model to simulated epileptic discharges, in part generated by a neural mass model. We analyzed the effect of features specific to epileptic neuronal activity on the blood oxygen level dependent (BOLD) response, focusing on the issues of linearity in neurovascular coupling and on the origin of negative BOLD signals. We found both sub- and supra-linearity in simulated BOLD signals, depending on whether one observes the early or the late part of the BOLD response. The size of these non-linear effects is determined by the spike frequency, as well as by the amplitude of the excitatory activity. Our results additionally indicate a minor deviation from linearity at the neuronal level. According to a phase space analysis, the possibility to obtain a negative BOLD response to an epileptic spike depends on the existence of a long and strong excitatory undershoot. Moreover, we strongly suggest that a combined EEG-fMRI modeling approach should include spatial assumptions. The present study is a step towards an increased understanding of the link between epileptic spikes and their BOLD responses, aiming to improve the interpretation of simultaneous EEG-fMRI recordings in epilepsy.

This work has been published in [17].

6.3.1.2. A nested cortex parcellation combining analysis of MEG forward problem and diffusion MRI tractography

**Participants:** Anne-Charlotte Philippe, Maureen Clerc, Théodore Papadopoulo, Rachid Deriche.
Understanding the relationship between structure and function is a major challenge in neuroscience. Diffusion MRI (dMRI) is the only non-invasive modality allowing to have access to the neural structure. Magnetoencephalography (MEG) is another non-invasive modality that allows a direct access to the temporal succession of cognitive processes. Functional cortex parcellation being one of the most important ways to understanding structure-function relationship, we propose an innovative method merging MEG and dMRI to parcellate the cortex. The combination of MEG forward problem and connectivity information reveals cortical areas generating a similar magnetic field at sensors while having a similar connectivity. Results show suitable clusters that forecast interesting studies for inter- and intra-subject comparisons of the cortex parcellations. The automatic nested cortex parcellation we propose could be a first step to analyse sources that are seeds of long or short range connectivity and to differentiate these connectivities in the white matter.

This work is currently under submission.

6.3.1.3. Improved computer-aided detection of small polyps in CT colonography using interpolation for curvature estimation

Participants: Rachid Deriche, Jiamin Liu [Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, Bethesda], Suraj Kabadi [Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, Bethesda], Robert Van Uitert [Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, Bethesda], Nicholas Petrick [Center for Devices and Radiological Health, U.S. Food and Drug Administration, Maryland], Ronald M. Summers [Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, Bethesda].

Surface curvatures are important geometric features for the computer-aided analysis and detection of polyps in CT colonography (CTC). However, the general kernel approach for curvature computation can yield erroneous results for small polyps and for polyps that lie on haustral folds. Those erroneous curvatures will reduce the performance of polyp detection. This work presents an analysis of interpolation’s effect on curvature estimation for thin structures and its application on computer-aided detection of small polyps in CTC. In this work, we demonstrated that a simple technique, image interpolation, can improve the accuracy of curvature estimation for thin structures and thus significantly improve the sensitivity of small polyp detection in CTC. Our experiments showed that the merits of interpolating included more accurate curvature values for simulated data, and isolation of polyps near folds for clinical data. After testing on a large clinical data set, it was observed that sensitivities with linear, quadratic B-spline and cubic B-spline interpolations significantly improved the sensitivity for small polyp detection. In conclusion, the image interpolation can improve the accuracy of curvature estimation for thin structures and thus improve the computer-aided detection of small polyps in CTC.

This work has been published in [16]
6. New Results

6.1. Sex-specific impact of meiotic recombination on nucleotide composition

Meiotic recombination is an important evolutionary force shaping the nucleotide landscape of genomes. For most vertebrates, the frequency of recombination varies slightly to considerably between the sexes (heterochiasmy). We extended the examination of the evolutionary impact of heterochiasmy beyond primates to include four additional eutherian mammals (mouse, dog, pig, and sheep), a metatherian mammal (opossum), and a bird (chicken). We compared sex-specific recombination rates with nucleotide substitution patterns evaluated on transposable elements. Our results, based on a comparative approach, revealed a great diversity of the relationship between heterochiasmy and nucleotide composition. We found that the stronger male impact on this relationship is a conserved feature of human, mouse, dog, and sheep. In contrast, variation in genomic GC content in pig and opossum is more strongly correlated with female, rather than male, recombination rate. We also showed that the sex-differential impact of recombination is mainly driven by the chromosomal localisation of recombination events, not the overall average recombination rate. We proposed a new explanation for the evolutionary impact of heterochiasmy on nucleotide composition. This work has been submitted for publication. This work was done in collaboration with D. Mouchiroud.

6.2. Modelling the influence of karyotype on the distribution of meiotic recombination

Given the important evolutionary role of recombination, recent work has focused on understanding the dynamics of this molecular process. We analysed the variation of recombination rate among species in relation to their karyotype. Specifically, we developed a non-linear model between the total genetic and physical lengths of chromosomes. Our model incorporates important biological knowledge of the recombination process. It further allows the estimation of two main parameters of recombination: the additional recombination rate per Mb and the per-species average strength of interference. Since the model is defined on data from genetic maps, at the global level of the karyotype, it can be applied even on low-resolution data and, hence, can result in the exploration of multiple species. By analysing the variability of our models recombination parameters among species, we showed that the recombination rate and the interference strength are regulated at the Mb scale of the genomes. We found that the genome size is a strong predictor of the recombination rate, while the average physical length of chromosomes is positively correlated with the interference parameter of our model. These relations represent valuable tools for the estimation of recombination parameters even for species lacking genetic maps. This work is in submission. This work was done in collaboration with D. Mouchiroud.

6.3. Finding long and multiple repeats with edit distance

We developed an algorithm, FILMRED, for detecting long similar fragments occurring at least twice in a set of biological sequences (a conference paper [25] has already appeared, a journal version is in preparation). The problem becomes computationally challenging when a non negligible number of insertions, deletions and substitutions are allowed. The algorithm is exact and manages instances whose size and combination of parameters cannot be handled by other currently existing method. This is achieved by using a filter as a preprocessing step, and then the information that this filter has gathered in the following inference phase. FILMRED can deal with very long repeats (up to a few thousands) occurring possibly several times, with a difference rate (substitutions and indels) of 10% or more. This work was done in collaboration with N. Pisanti and P. Peterlongo. The software will be made available in a near future.
6.4. Genomics of symbiosis
Insect symbioses are model systems for studying the evolution of bacterial genomes. Importantly, evolution is directly related to the type of interaction and may also be influenced by the presence of other symbionts [11]. We are currently studying the genomes of different symbionts. The first one is the genome of the *Wolbachia* strain that has recently become obligatory for the reproduction of *Asobara tabida*. First results indicate that this genome may have been recently invaded by mobile elements. The second project concerns the sequencing of the different symbionts that co-exist in the insect *Bemisia tabaci*. The initial sequences are promising and we hope to close the genome of four different symbionts in this system, which will allow the study of the complementation between symbionts and lateral gene transfers among symbionts sharing the same intracellular arena. This work is done in collaboration among others with L. Mouton.

6.5. Bacterial syntenies
The automatic identification of syntenies across multiple species is a key step in comparative genomics that helps biologists shed light both on evolutionary and functional problems. We developed a versatile algorithm to extract all syntenies from multiple bacterial species based on a clear-cut and very flexible definition of the synteny blocks that allows for gene quorum, partial gene correspondence, gaps, and a partial or total conservation of the gene order [8]. We then applied this algorithm to two different kinds of studies. The first one is a search for functional gene associations. In this context, we compared our algorithm to a widely used heuristic - I-ADHORE - and showed that at least up to ten genomes, the problem remains tractable with our exact definition and algorithm. The second application was linked to evolutionary studies: we verified in a multiple alignment setting that pairs of orthologs in synteny are more conserved than pairs outside, thus extending a previous pairwise study. We then showed that this observation is in fact a function of the size of the synteny: the larger the block of synteny is, the more conserved the genes are. This work was done in collaboration with F. Boyer.

6.6. Spatial synteny in Eukaryotes
Folding and intermingling of chromosomes has the potential of bringing close to each other loci that are very distant genomically or even on different chromosomes. On the other hand, genomic rearrangements also play a major role in the reorganisation of loci proximities. Whether the same loci are involved in both mechanisms has been studied in the case of somatic rearrangements, but never from an evolutionary standpoint. From the joint study of the network of spatial proximities of human genomic loci and a dataset of evolutionary breakpoints between human and mouse, we were able to provide evidence that evolutionary breakpoints tended to cluster spatially in human cells, which led us to propose the new notion of spatial synteny, which generalises the concept of genomic synteny. This work was submitted in 2010 and was accepted in 2011 [24]. It was done in collaboration with C. Lemaitre.

6.7. Chimeric transcripts in Eukaryotes
In the framework of the extension of the ENCODE project to chromosomes 21 and 22, we had the opportunity to identify a new category of transcripts, which we call chimeric transcripts, since they contain exons from different genes, which can themselves be located far apart on the same chromosome, or on different chromosomes [9]. We further found that the network formed by these connected genes is enriched in cliques of sizes 3 and 4, which seems to indicate that transcription and/or splicing of these sets of genes co-occur in time and space, as is confirmed by the confrontation of our expression dataset to a dataset indicating the co-localisation of DNA fragments in 3D. This work was done in collaboration with, among others, R. Guigó.

6.8. KisSplice: de-novo calling alternative splicing events from RNA-seq data
We addressed the problem of identifying polymorphisms in RNA-seq data when no reference genome is available, and avoiding an assembly. Based on the fundamental idea that each polymorphism will correspond to a recognisable pattern in a De Bruijn graph constructed from the RNA-seq reads, we propose a general
model for all polymorphisms in such graphs. We then introduce an exact algorithm to extract alternative splicing events and show that it enables to identify more correct events than current transcriptome assemblers. Additionally, when we applied our method on a 50M reads dataset from human, we were able to identify 3884 events, out of which 57\% are not present in the annotations, which confirms recent estimates showing that the complexity of alternative splicing has been largely underestimated so far. This work has been submitted to publication. This work was done in collaboration with P. Peterlongo.

6.9. Transcriptomics of symbiosis in the Asobara tabida-Wolbachia association

*Wolbachia* has evolved a very peculiar phenotype in the host *Asobara tabida* where it is obligatory for oogenesis. Through transcriptomic approaches (Sanger sequencing of mRNA, in vitro subtraction of transcriptomes), we have established a first reference transcriptome of this insect. The analyses done demonstrate that *Wolbachia* interferes with different host pathways, and notably regulation of oxidative stress, apoptosis and autophagy, which are known to be involved in host-pathogen interactions. RNAseq has now been performed on this system and analyses are underway, which will allow a finer investigation of the interaction using the algorithm KISSPLICE (see above) developed in the EPI for the analysis of NGS data without reference genome.

6.10. Navigating through unexplored pre-miRNA candidates

The computational search for novel miRNA precursors often involves some sort of structural analysis with the aim of identifying which type of structures are prone to being recognised and processed by the cellular miRNA-maturation machinery. A natural way to tackle this problem is to perform clustering over the candidate structures along with known miRNA precursor structures. Mixed clusters allows then the identification of candidates that are similar to known precursors. Given the large number of candidate pre-miRNAs that can be identified in single-genome approaches, even after applying several filters for robustness and stability, a conventional structural clustering approach is unfeasible. We presented a method, MinDist, to represent candidate structures in a feature space which summarises key sequence/structure characteristics of each candidate. We demonstrated that proximity in this feature space is related to sequence/structure similarity, and we selected candidates which have a high similarity to known precursors. Additional filtering steps were then applied to further reduce the number of candidates to those with greater transcriptional potential. Our method was compared to another single-genome method (TRIPLET SVM) in two datasets, showing better performance in one and comparable performance in the other. Additionally, we showed that our approach allows for a better interpretation of the results. The MinDist method is available upon request and will be made available online. This work has been submitted to publication. This work was done in collaboration with A. T. Freitas and R. Backofen.

6.11. Exploration of the genetic network of Buchnera aphidicola

Aphids are important agricultural pests which can grow and reproduce thanks to their intimate symbiosis with the \(\gamma\)-proteobacterium *Buchnera aphidicola* that furnishes them with essential amino acids lacking in their phloem sap diet. We investigated how *B. aphidicola*, with its reduced genome containing very few transcriptional regulators, responds to variations in the metabolic requirements of its host by concentrating attention on the leucine metabolic pathway [23]. We showed that leucine is a limiting factor for aphid growth and displays a stimulatory feeding effect. Our metabolic analyses demonstrated that symbiotic aphids are able to respond to leucine starvation or excess by modulating the neosynthesis of this amino acid. Taken together, our data showed that the response of *B. aphidicola* to the leucine demand of its host is multimodal and dynamically regulated, providing new insights concerning the genetic regulation capabilities of this bacterium in relation to its symbiotic functions.
6.12. Annotation database system to ease the development and update of BioCyc databases

In recent years, genomes from an increasing number of organisms have been sequenced, but their annotation remains a time-consuming process. The BioCyc databases offer a framework for the integrated analysis of metabolic networks. The PATHWAY TOOL SOFTWARE SUITE allows the automated construction of a database starting from an annotated genome, but it requires prior integration of all annotations into a specific summary file or into a GenBank file. To allow the easy creation and update of a BioCyc database starting from the multiple genome annotation resources available over time, we developed an ad hoc data management system that we called Cyc Annotation Database System (CYCADS) [22]. The CYCADS pipeline for annotation management was used to build the ACYPICYC database for the pea aphid Acyrthosiphon pisum, TRICACYC for Tribolium castaneum and DROMEYC for Drosophila melanogaster. This work will be extended to create a database for other arthropods. This work was done in collaboration among others with S. Collela.

6.13. Representation and curation of metabolic pathways: UniPathway

UNIPATHWAY (http://www.unipathway.org) is a manually curated database for the representation and annotation of metabolic pathways developed in collaboration with the Swiss Institute of Bioinformatics (SwissProt group). UNIPATHWAY provides explicit chemical representations of enzyme-catalysed and spontaneous chemical reactions, as well as a hierarchical representation of metabolic pathways. This hierarchy uses linear subpathways as the basic building block for the assembly of larger and more complex pathways, including species-specific pathway variants. All of the pathway data in UNIPATHWAY has been extensively cross-linked to existing pathway resources such as KEGG and METAcyC, as well as sequence resources such as the UNIPROT KNOWLEDGE BASE (UNIPROT KB). UNIPATHWAY has been used to provide a controlled vocabulary for pathway annotation within UNIPROT KB records since UNIPROT release 14.7 (January 2009). In release 2011_08 of UNIPROT, UNIPATHWAY provides annotation for 118,390 distinct Swiss-Prot protein records and 783,299 TrEMBL protein records. On the UNIPROT KB web site, each of these records is linked to the appropriate pathway description in the UNIPATHWAY web site. In 2011, the complete description of the UNIPATHWAY database has been published in Nucleic Acids Research (Jan. 2012 Database Issue) and has been chosen by the editors of Nucleic Acids Research to appear on the Featured Articles page (top 5% of NAR papers: http://www.oxfordjournals.org/our_journals/nar/featured_articles.html) [16].


RHEA (http://www.ebi.ac.uk/rhea) is a project developed in collaboration with the Swiss Institute of Bioinformatics (SIB) and the European Institute for Bioinformatics (EBI). It aims at providing a comprehensive resource of expert-curated biochemical reactions. RHEA provides a non-redundant set of chemical transformations for use in a broad spectrum of applications, including metabolic network reconstruction and pathway inference. RHEA includes enzyme-catalysed reactions (covering the IUBMB Enzyme Nomenclature list), transport reactions and spontaneously occurring reactions. RHEA reactions are described using chemical species from the Chemical Entities of Biological Interest ontology (ChEBI) and are stoichiometrically balanced for mass and charge. They are extensively manually curated with links to source literature and other public resources on metabolism including enzyme and pathway databases. This cross-referencing facilitates the mapping and reconciliation of common reactions and compounds between distinct resources, which is a common first step in the reconstruction of genome scale metabolic networks and models. The complete description of the database will appear in the Jan. 2012 NAR Database issue [5].

6.15. Metabolic reconstruction of Klebsiella pneumoniae str. Kp13

Klebsiella pneumoniae str. Kp13 is a multidrug resistant pathogen involved in nosocomial outbreaks in Brazil. The objectives of this study still underway are two-fold: (1) to perform a graph-based metabolic reconstruction of the small-molecules network of strain Kp13 and (2) from the reconstructed network evaluate what makes this pathogen so successful in colonising its human host. Manual annotation of the network was performed...
and a choke-point analysis was carried out, which yielded interesting targets such as L-rhamnose biosynthesis enzymes that may be related to the virulence of this bacterium. The MetAnnot platform within MetExplore was used for the manual curation of the network and graph export/import. A paper is in preparation. This work is being done in collaboration with A. T. Vasconcelos and M. Nicolás.

6.16. Clustering of elementary modes and metabolic modules identification

While it is commonly admitted that metabolism is modular, the identification of metabolic modules remains an open topic. In fact, what remains open comes even upstream of any identification problem, and refers instead to the question of defining a good model for modules in metabolic networks. One would hope that such a model might enable, for instance, to automatically derive the metabolic pathways that have been discovered and painstakingly established by biochemists over the years. Elementary modes, that are informally defined as metabolic subnetworks that can function at steady state, meaning that all internal metabolites are produced and consumed in equal rates (that is, nothing accumulates internally), represent one starting point for a definition of modules that has been considered. There are two difficulties related to this however. One is that enumerating elementary modes has itself been proven (by members of the EPI) to be a hard problem, while the second is that even small networks (around 100 reaction nodes) can have millions of elementary modes. Clustering them based on, for instance, the amount of overlap, that is of shared reactions, is one idea that has been used. We attempted another definition of modules using elementary modes that is related to a form of node-covering. This has been submitted to publication. The corresponding software is available on request. This work was done in collaboration with C. E. Ferreira, E. Moreno, and P. Crescenzi.

6.17. Enumeration of metabolic stories

In many cases, we are interested in understanding how metabolism reacts when an organism is submitted to some environmental stress, that is, to establish which metabolic processes are involved in an organism’s adaptation to such stress. In order to do it, elements of the metabolism such as metabolites are monitored to determine which are over- or under-produced during the stress response as compared to the organism’s state in normal conditions. Such quantitative and qualitative measurements are called metabolomics. The affected part may represent only a small portion of the network, that is, involve a small subset of metabolites. The aim then is to identify subnetworks that enable to link together all elements in this subset. More formally, given such subset and a metabolic network represented as a digraph where nodes are metabolites and edges link two metabolites when one is the input of a reaction that produces the other, we are interested in identifying all maximal directed acyclic graphs that cover all the metabolites in the subset of interest, and have no sources or targets that are not one of these metabolites. Such maximal DAGs are called metabolic stories. We established already that finding one metabolic story is easy (paper submitted) and used the algorithm developed, GOBBOLINO (see the Softwares Section), for practical purposes (second publication in preparation). This work was done in collaboration with P. Crescenzi, A. Marchetti-Spaccamela, L. Stougie, A. Marino, F. Jourdan, and L. Cottret.

6.18. Identification of all the minimal precursor sets for a given set of targets

Once the metabolic network of an organism has been defined, the question of how are produced the essential metabolites for the organism arises. In particular, it is important to know which are the metabolites that the organism needs to obtain from its environment to produce those essential metabolites. In the case of symbiosis, this environment could be the host, and determining such metabolites one way of exploring the dialog that is established between different organisms that entertain a close and often long term relation. We call such metabolites that must be obtained from the environment precursors. We had already in 2008 established the complexity of the problem of, given a network and a set of targets of interest, enumerating all minimal sets of precursors enabling to produce the targets, and given one first algorithm. The algorithm has since been much improved (journal paper submitted). The algorithm has been developed into a software called PITUFO (see the Softwares Section). This work was done in collaboration with A. Marchetti-Spaccamela, L. Stougie, and L. Cottret.
6.19. Metabolic network comparison

All living organisms have very similar metabolic needs as they all uptake nutrients from their environment, degrade them into basic building blocks such as amino acids and nucleotides, which in turn are essential input for protein and DNA synthesis. A natural expectation is therefore that they share an even reduced core of metabolic functions necessary to carry out this basic small molecule metabolism. Comparing the small molecule metabolism of 58 bacteria carefully selected and representing a range of lifestyles, we found not a single enzymatic reaction common to all of them. This absence of a metabolic core is essentially due to intracellular symbionts. These results were in preparation in 2010 and are now submitted. The work was done in collaboration with Ludovic Cottret (INSA Toulouse) and Ana Tereza Vasconcelos.

6.20. Wolbachia detection

Wolbachia is a large monophyletic genus of intracellular bacteria, traditionally detected using PCR assays. Its considerable phylogenetic diversity and impact on arthropods and nematodes make it urgent to assess the efficiency of these screening protocols. The sensitivity and range of commonly used PCR primers and of a new set of 16S primers were evaluated on a wide range of hosts and Wolbachia strains [20]. We showed that certain primer sets are significantly more efficient than others but that no single protocol can ensure the specific detection of all known Wolbachia infections. This work was done in collaboration among others with S. Charlat.

6.21. Genetic architecture of parasite infection

The problem here is to understand the genetic architecture of a parasitic invasion by investigating the different phenotypes such invasion produces in the host. One such phenotype is called "cytoplasmic incompatibility". Briefly, when a parasite invades a male host, it induces developmental arrest, ultimately, death of the host’s offspring unless the fertilised embryo carries the same symbiont inherited from its mother, that is, unless the female is also infected. This has been tentatively explained by a toxin/antitoxin model that involves a toxin deposited by the parasites in the male’s sperm that induces the death of the zygote unless neutralised by an antidote produced by the parasites present in the egg. One toxin/antitoxin pair is linked to one gene. Given a set of observed cytoplasmic incompatibilities, the question is how many genes are required to explain it. Formally, and skipping many intermediate modelling steps, this translates into, given a 0/1 matrix M for pairs of male/female (a 0 indicating that either the male is not infected or, if it is, then so is the female meaning that there is no incompatibility, and a 1 indicating that the male is infected while the female is not), what is the minimum number of "rectangles" that enable to cover all the 1s in M? A "rectangle" in this case is a subset of columns and rows such that, if permuted, they can be arranged in the form of a rectangle with only 1s (publication Nor et al., 2010 by the BAMBOO Team and collaborators). One rectangle corresponds to a gene. The model can then be made more complex by considering that genes may have different alleles (different forms), and are expressed in variable quantities. The quantitative version of the problem in particular translates into having to find a minimum number of "triangles" that cover all 1s. All the above problems translate also into different versions of edge covers of a bipartite graph that are for the most part algorithmically original, and always not fully resolved (meaning, there remains open questions, notably regarding complexity). Work on these problems within the Associate Team SIMBIOSI and the results already obtained should lead to a publication in 2012. This work was done in collaboration among others with S. Charlat.

6.22. Population structure and dynamics

We recently started a collaboration with the Pasteur Institute in Cambodia (Dr. P. Buchy) and the CIRAD at Montpellier (Dr. R. Frutos) on viral population structure and dynamics. In this context, we developed an exploratory statistical approach to characterise mutational patterns in viral populations. The basic idea is to use Multiple Correspondence Analysis (MCoA) on a multiple alignment of nucleic sequences. To this purpose, the alignment is encoded as a boolean table where rows correspond to the sequences and columns to the presence/absence of characters. This can be done simply by considering each of the four possible bases as a
different character, or by considering only two possible states: one for the major base and one for all other (minor) bases. This technique turned out to be very effective in representing co-mutation patterns within a population of sequences. As for simple CoA, the plot is quite easy to interpret by biologists, both in terms of proximities of sequences and characters \( (i.e. \) mutations). Moreover, it has some strong relationships with parsimony phylogeny that need to be clarified. In [ 10 ], we applied this technique to study the structure and time evolution of the Dengue virus (serotype I) population in Cambodia, using heterochronous sequence samples. Beside its methodological aspect, this work also introduced new topics related to population genetics in BAMBOO \( (e.g. \) the use of coalescent theory to reconstruct population dynamics). Another approach, widely used in ecology, to measure biodiversity and to determine the species composition of environmental samples is the "DNA-barcoding" technique. In collaboration with the Laboratoire d’Ecologie Alpine (Univ. Joseph Fourier Grenoble), we developed a new software for identifying new barcode markers and their associated PCR primers [ 19 ].
5. New Results

5.1. Proliferation dynamics and its control

5.1.1. Cell division dynamics in structured cell populations

Participants: José Luis Avila Alonso [DISCO project-team, INRIA Saclay IdF], Annabelle Ballesta, Houda Benjelloun [INSA Rouen], Frédérique Billy, Frédéric Bonnans [Commands project-team, INRIA Saclay IdF], Catherine Bonnet [DISCO project-team, INRIA Saclay IdF], Jean Clairambault, Luna Dimitrio, Marie Doumic-Jaffret, Xavier Dupuis [Commands project-team], Olivier Fercoq [MaxPlus project-team, INRIA Saclay IdF], Stéphane Gaubert [MaxPlus project-team, INRIA Saclay IdF], Germain Gillet [IBCP, Université Cl. Bernard Lyon 1], Philippe Gonzalo [IBCP, Université Cl. Bernard Lyon 1], Pierre Hirsch [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Thomas Lepoutre [now in DRACULA project-team, INRIA Rhône-Alpes, Lyon], Jonathan Lopez [IBCP, Université Cl. Bernard Lyon 1], Pierre Magal [University Bordeaux II], Anna Marciniak-Czochra [Institute of Applied Mathematics, Universität Heidelberg], Jean-Pierre Marie [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Faten Merhi [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Roberto Natalini [IAC-CNR, Università Sapienza, Rome], Silviu Niculescu [DISCO project-team, INRIA Saclay IdF], Hitay Özbay [Bilkent University, Ankara, Turkey], Benoît Perthame, Ruopeng Tang [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Vitaly Volpert [CNRS Lyon, UMR5208, Camille Jordan Institute, Lyon], Jorge Zubelli [IMPA, Rio de Janeiro).

1. Transition kernels in a McKendrick model of the cell division cycle. A focus has been set on transitions between phases of the cell division cycle. The underlying biological question is: “Is desynchronisation between cells in proliferating cell populations a hallmark of cancer?” It has been considered by relating in a natural way transition kernels with the probability density functions of transition times in the cell population. It has been shown -which was expected, but never proved to our knowledge so far- that the more desynchronised cells are with respect to cell cycle phase transitions, the higher is the growth exponent of the cell population. Otherwise, desynchronised cell populations grow faster. This has been proven when transition kernels are time-independent, i.e., when no external control is exerted on transitions. The same question is currently experimentally investigated by our biologist partners in the European network ERASysBio+ C5Sys, coordinated by F. Lévi (Villejuif) and D. Rand (Warwick). Simulations using experimentally identified transition kernels in proliferating cell cultures controlled by theoretical time-dependent (circadian) control functions have verified the relevance of this mathematical result for theoretical cancer treatment optimisation (cf. infra “Periodic (circadian) control of cell proliferation in a theoretical model of the McKendrick type”).

2. Modelling haematopoiesis with applications to AML. The stability of a delay system based on a PDE model designed by M. Adimy and F. Crauste, structured by a discrete differentiation variable and multiple delays, with applications to Acute Myeloblastic Leukaemia (AML, clinical advisers: J.-P. Marie and P. Hirsch; technical adviser: RP Tang) is studied with possible therapeutic implications. This model is currently experimentally investigated, with the aim to identify its parameters in leukaemic cells, in the DIGITEO project ALMA (cf. infra “DIGITEO and Cancéropôle IdF” in “Regional initiatives”), coordinated by C. Bonnet (DISCO team, INRIA Saclay IdF) and in the recently launched DIGITEO project ALMA2 (coordinated by J. Clairambault), that takes over the combined experimental-modelling activity in ALMA. Two INRIA postdocs, F. Merhi (in ALMA, 2010-2011) and A. Ballesta (in ALMA2, 2011-2013) have been devoted to this task. From a theoretical point of view, the Adimy-Crauste model has been modified so as a) to include quick
self-renewal of cells in each stage of maturation and b) to represent each phase of the proliferating compartment (i.e., $G_1$, $S$, $G_2$ and $M$) separately. For the time being, only the $M$ phase is supposed to have a fixed time duration as it is generally admitted that the short time (typically half an hour if the total proliferating phase duration is normalised to 24 hours) necessary to perform mitosis is hardly submitted to any variation.

In a complementary manner, a new model for cell differentiation was introduced and analysed in [17], in collaboration with A. Marciniak and J.P. Zubelli. It assumed that differentiation of progenitor cells is a continuous process. From the mathematical point of view, it is based on partial differential equations of transport type. Specifically, it consists of a structured population equation with a nonlinear feedback loop. This models the signaling process due to cytokines, which regulate the differentiation and proliferation process. We compared the continuous model to its discrete counterpart, a multicompartmental model of a discrete collection of cell subpopulations recently proposed by Marciniak-Czochra et al. [Stem Cells Dev., 18 (2009), pp. 377–386] to investigate the dynamics of the hematopoietic system. We obtained uniform bounds for the solutions, characterized steady state solutions, and analyzed their linearized stability. We showed how persistence or extinction might occur according to values of parameters that characterize the stem cells’ self-renewal. We also performed numerical simulations and discuss the qualitative behavior of the continuous model vis-à-vis the discrete one.

3. Hybrid models

Systems combining PDEs and discrete representations in hybrid models, with applications to cancer growth and therapy, in particular for AML, are the object of study of the ANR program Bimod, coordinated by V. Volpert (Lyon), associating CNRS (V. Volpert, Lyon), Bordeaux II University (P. Magal) and the Bang project-team.


With G. Gilllet (prof. at IBCP/Lyon), we have designed a mathematical ODE model for the mitochondrial pathway of apoptosis, focused on the early phase of apoptosis (before the cytochrome C release). We have validated it with experimental data carried out in G. Gillet’s lab and applied it to propose new therapeutic strategies against cancer. This work has led to a nearly submitted article [47].


Following her first year of PhD in Rome with R. Natalini, working on cytoplasmic transport along microtubules presented in [38], L. Dimitrio has begun her third PhD year, going on studying at INRIA nucleocytoplasmic transport with applications to p53 activity. Her PhD thesis work is supervised in co-tutela between Sapienza University in Rome (R. Natalini) and INRIA (J. Clairambault). The protein p53 plays a capital part as “guardian of the genome”, arresting the cell cycle and launching cell apoptosis or DNA repair in case of DNA damage. Results expected from this newly developed theme will provide a rational link between molecular pharmacokinetics-pharmacodynamics (cf. infra) of anticancer drugs and modelling of the cell division cycle in proliferating cell populations. L. Dimitrio has presented her ongoing work in different meetings in France and in Italy, and a paper in preparation will be submitted in 2012.

5.1.2. Physiological and pharmacological control of cell proliferation

Participants: Annabelle Ballesta, Frédérique Billy, Jean Clairambault, Sandrine Duloung [INSERM Villejuif (U 776)], Olivier Fercq [MaxPlus project-team], Stéphane Gaubert [MaxPlus project-team], Thomas Lepoutre [Dracula project-team], Francis Lévi [INSERM Villejuif (U 776)].

1. Periodic (circadian) control of cell proliferation in a theoretical model of the McKendrick type. The impact of a periodic control exerted on cell cycle phase transitions has continued to be studied [16] with the collaboration of S. Gaubert (MaxPlus INRIA project-team, Saclay IDF) and T. Lepoutre (Dracula INRIA project-team, Lyon) and is currently investigated experimentally in the new C5Sys European network (cf. supra “Transition kernels in a McKendrick model of the cell division cycle”
and “). Thanks to the work of Frédérique Billy (Postdoc in Bang) and Olivier Fercoq (PhD student in MaxPlus), together with permanent members of Bang, Dracula and MaxPlus teams, it has led to three publications [37], [39], [48].

2. **Intracellular pharmacokinetic-pharmacodynamic (PK-PD) models for anticancer drugs.** This theme is actively worked on in collaboration, mainly with the teams of F. Lévi and J.-P. Marie (cf. supra “Transition kernels in a McKendrick model of the cell division cycle” and “Modelling haematopoiesis with applications to AML”). After a PK-PD model for 5-FU with folinic acid [86], it has led for the anticancer drug Irinotecan, the main object of A. Ballesta’s PhD thesis [1], to an article published in PLoS Computational Biology [8], reporting a combined modelling and experimental approach to the effects of a combination of mathematical modelling and experimentation in cell cultures, and to another one [7], focusing on drug delivery optimisation.

3. **Whole body physiologically based model of anticancer drug pharmacokinetics.** This theme has also been studied in A. Ballesta’s PhD thesis. The use of identification, in genetically different laboratory mouse strains, of parameters characterising an ODE model of the action of Irinotecan (cf. supra “Intracellular pharmacokinetic-pharmacodynamic (PK-PD) models for anticancer drugs”) in cell cultures, transposed at the whole-body level, has been designed as a proof of concept for individual adaptation of drug delivery in the context of (future) personalised medicine, a perspective sketched in [15] and among other collaborative contexts linking mathematics and medicine in [14], [3].

5.1.3. **Optimisation of cancer chemotherapy**

**Participants:** Annabelle Ballesta, Frédérique Billy, Frédéric Bonnans [Commands project-team], Jean Clairambault, Sandrine Dulong [INSERM Villejuif (U 776)], Xavier Dupuis [Commands project-team], Olivier Fercoq [MaxPlus project-team], Stéphane Gaubert [MaxPlus project-team], Thomas Lepoutre [Dracula project-team], Alexander Lorz, Francis Lévi [INSERM U 776, Villejuif], Michael Hochberg [ISEM, CNRS, Montpellier], Benoît Perthame.

Optimising cancer chemotherapy, especially chronotherapy, is the final aim of the activities mentioned above. This has been lately discussed in [16] and also in works involving the C5Sys network [37], [48], and in the more general review [39]. Until now had been taken into account as constraints in optimisation strategies only the unwanted toxic side effects of anticancer drugs on healthy cells. More recently, another issue of anticancer treatment has been considered, namely the different mechanisms of resistance to drugs in cancer cells. This has led to include the effect of ABC transporters (active efflux pumps, as is the P-glycoprotein) in the intracellular PK-PD models mentioned above [86], in A. Ballesta’s PhD joint work with F. Lévi’s team [1], [8], [7], and to a perspective paper [15].

In project is also the use of methods of optimal control developed by the Commands project-team (F. Bonnans, X. Dupuis) to optimise therapies in the treatment of Acute Myeloblastic Leukaemia (AML, cf. supra “Modelling haematopoiesis with applications to AML”).

Another way to represent and overcome drug resistance in cancer from a cell Darwinian point of view using concepts of adaptive dynamics in proliferating cell populations is also currently being investigated along the line of other recent works [28] and is currently developed within the multidisciplinary GDR DarEvCan coordinated by M. Hochberg, Montpellier (cf. infra “GDR DarEvCan” in “National initiatives”) and in a proposed ANR project also coordinated by M. Hochberg.

An open question that should have therapeutic implications consists of the interrogation: Is the emergence of drug resistance in cell populations a genetic (resulting from mutations at mitosis) or an epigenetic phenomenon (resulting from amplification of physiological mechanisms, such as ABC transport, which has nothing to do with genetic mutations)? And is it a reversible or irreversible phenomenon? These questions will be studied both theoretically and experimentally within the DarEvCan consortium and could result in new developments in the so-called Darwinian medicine.
5.1.4. Protein polymerisation and application to amyloid diseases (ANR grant TOPPAZ)

**Participants:** Annabelle Ballesta, Vincent Calvez [ENS Lyon], Frédérique Charles, Marie Doumic-Jauffret, Pierre Gabriel, Hadjer Wafaâ Haffaf, Benoît Perthame, Stéphanie Prigent [BPCP, INRA Jouy-en-Josas], Human Rezaei [BPCP, INRA Jouy-en-Josas], Léon Matar Tine [SIMPAF project-team, INRIA Lille Nord-Europe].

With H. Rezaei, a new and very complete PDE model for protein polymerisation has been designed. Following F. Charles’ work, A. Ballesta has applied this model to Huntington’s disease (PolyQ expansion) and compared it with its ODE counterpart, leading to a better understanding of the leading mechanisms responsible for PolyQ fibrillation. This part is nearly submitted.

The eigenvalue problem playing a major role in the representation of Prion proliferation dynamics and, in a more general way, of many fragmentation-coalescence phenomena, the article [11] investigated the dependency of the principal eigenvector and eigenvalue upon its parameters. We exhibited possible non-monotonic dependency on the parameters, conversely to what would have been conjectured on the basis of some simple cases.

5.1.5. Inverse problem in growth-fragmentation equations

**Participants:** Marie Doumic-Jauffret, Marc Hoffmann [ENSAE], Patricia Reynaud [CNRS, Nice Univ.], Vincent Rivoirard [Paris IX Univ.], Léon Matar Tine [SIMPAF project-team, INRIA Lille Nord-Europe].

In collaboration with statisticians (M. Hoffman, Professor at Université de Marne-la-Vallée, V. Rivoirard, MC at Université d’Orsay, and P. Reynaud, CR CNRS at Université de Nice), in [18] we have explored a statistical viewpoint on the cell division problem. In contrast to a deterministic inverse problem approach, we take the perspective of statistical inference. By estimating statistically each term of the eigenvalue problem and by suitably inverting a certain linear operator, we are able to construct an estimator of the division rate that achieves the same optimal error bound as in related deterministic inverse problems. Our procedure relies on kernel methods with automatic bandwidth selection. It is inspired by model selection and recent results of Goldenschluger and Lepski. This work is accepted in SIAM J. Num. Anal..

With L. Matar Tine, in [53] we have generalized the inverse techniques proposed in [88], [67] and [66], in order to adapt them to general fragmentation kernels and growth speeds. The potential applications of this problem are numerous, ranging from polymerisation processes to the cell division cycle. This work is submitted.

5.2. Tissue growth, regeneration and cell movements

5.2.1. Chemotaxis, self-organisation of cell communities

**Participants:** Nikolaos Bournaveas [Univ. Edinburgh], Axel Buguin [UPMC, Institut Curie], Vincent Calvez [ENS Lyon], François James [univ. Orléans], Alexander Lorz, Grégoire Nadin [UPMC], Benoît Perthame, Jonathan Saragosti [Institut Curie], Pascal Silberzan [Institut Curie], Min Tang [SJTU], Nicolas Vauchelet.

We have continued our analysis and simulation of models for large bacterial communities and more generally cells self-organisation as initiated several years ago [12], [31]. This is a rich domain because on the one hand several Partial Differential Equations arise, parabolic models, kinetic equations, hyperbolic systems and on the other hand complex patterns occur that are a sign of the complex underlying dynamics. In their article [65], Y. Dolak and C. Schmeiser have proposed a mathematical model describing the individual behavior of bacteria responding to a chemical substance. Numerical simulations of this kinetic model have been investigated in [93]. In a macroscopic level of description, we perform a hydrodynamical limit which leads to an aggregation model, for which regular solutions blows up in final time [82]. The rigorous study of the behavior of the aggregates relies on a careful analysis of some non-linear scalar conservation laws in the framework of measures [24] and has been investigated in [55].
With the group of P. Silberzan in Curie Institute, in 2010 we have given an explanation of traveling bands first observed by Adler in the 80’s for *E. coli* in microchannels. In a continuation of this work, based on analysis of individual trajectories, we have shown directional persistence which improves the efficiency of collective migration. Kinetic models with tumbling kernels that keep memory of the incoming velocity are able to reproduce accurately the wave parameters \[ 35 \].

Computing effectively traveling wave can be difficult in unstable cases. An algorithm is proposed in \[ 29 \] for the NonLocalFisher equation which catches the traveling wave and not the more stable pulsating wave.

A. Lorz has studied a system consisting of the elliptic-parabolic Keller–Segel equations coupled to Stokes equations by transport and gravitational forcing. We show global-in-time existence of solutions for small initial mass in 2D. In 3D we establish global existence assuming that the initial \( L^{3/2} \)-norm is small. Moreover, we give numerical evidence that for this extension of the Keller–Segel system in 2D, solutions exist with mass above \( 8\pi \), which is the critical mass for the system without fluid. The model is written as

\[
\begin{align*}
  u \cdot \nabla c &= \Delta c + n - a_1 c, \\
  n_t + u \cdot \nabla n &= \Delta n - \nabla \cdot (\chi n \nabla c), \\
  a_2 u_t + \nabla P - \eta \Delta u + n \nabla \phi &= 0, \\
  \nabla \cdot u &= 0.
\end{align*}
\]

(34)

Here \( c \) denotes the concentration of a chemical, \( n \) a cell density and \( u \) a fluid velocity field described by Stokes equations. The fluid couples to \( n \) and \( c \) through transport and gravitational forcing modelled by \( \nabla \phi \). The pressure \( P \) can be seen as the Lagrange multiplier enforcing the incompressibility constraint. The chemical \( c \) diffuses, it is produced by the cells and it degrades. The cell density diffuses and it moves in the direction...
of the chemical gradient. The constant $a_1 \geq 0$ measures self-degradation of the chemical and the constants $a_2 \geq 0$, $\eta > 0$ determine the evolution undergone by $u$.

5.2.2. **Single cell-based models of tumour growth, tissue regeneration, embryonic development**

**Participants:** Annabelle Ballesta, Gregory Batt [CONTRAINTES project-team], François Bertaux, Chadha Chettaoui, Ibrahim Cheddadi, Dirk Drasdo, Adrian Friebel, Rolf Gebhardt [Univ. of Leipzig, Germany], Adriano Henney [Director Virtual Liver Network and VLN consortium], Jan G. Hengstler [Leibniz Research Center, Dortmund, Germany and CANCERSYS consortium], Stefan Höhme, Elmar Heinzle [University of Saarbrücken and NOTOX consortium], Isabelle Hue [INRA], Nick Jagiella, Ursula Klingmüller [German Cancer Center, Heidelberg and LungSys Consortium], Axel Krinner, Emanuele Leoncini, Johannes Neitsch, Benoît Perthame, Ignacio Ramis-Conde, Luc Soler [ IRCAD, Coordinator EU-project PASSPORT and PASSPORT consortium], Irène Vignon-Clémentel [REO project-team], Juhui Wang [INRA], William Weens.

Structure formation in tissues as well as malfunctions on the multi-cellular level are inherently of multi-scale nature. Modifications on the molecular level by intrinsic or extrinsic factors affect the architecture and function on the multi-cellular tissue level. Much of the current research so far focuses on the analysis of intracellular pathways, genetic and metabolic regulation on the intracellular scale and on continuum equations for local densities of cells to capture multi-cellular objects on large spatial scales but only recently have increasing efforts been made at the interface between these two: individual cell based models (IBMs) which permit to include the molecular information on the one hand and to extrapolate to the multi-cellular tissue level on the other hand and hybrid models that combine continuum with individual-based models for different components.

![Figure 2. Left: Modeling ammonia detoxification during regeneration after drug-induced damage in a multi-lobule model. Here, a compartment model of ammonia detoxification has been coupled to the regeneration of liver mass. The colour indicates the gradient of ammonia after regeneration. The upper graphics indicate that our model directly calculates the blood concentrations of ammonia, urea and glutamine during the liver regeneration process. Right: In the meantime it is possible to simulate whole mice lobes up to 10 cells thick.](image)

In order to fill the existing gap we have studied intracellular regulation networks [87], [72], multi-scale IBMs where intracellular regulation and differentiation was explicitly represented within each individual cell [90], [85], [91], lattice-free IBMs [69] and continuum models that can capture their large scale behaviour [63], and cellular automaton (CA) models where each lattice site can be occupied either by at most one cell [61] or by many cells [89], [71] and their corresponding continuum equation [68]. Moreover, for a simple, but for rigorous coarse graining not accessible, growth situation we were able to obtain quantitatively matching results with continuum and individual-cell-based models without any fit parameter.
Besides the methodical aspects we focus on a number of applications:

- Unstructured cell populations growing in a monolayer with free border \[69\], \[74\], or constraint by the presence of a granular or cellular embedding medium \[19\].

- Multicellular spheroids in liquid suspension \[69\], \[70\], and embedding granular or cellular matter \[19\]. For non-small-lung-cancer cell lines growing as multi-cellular spheroids, we could, starting with a complete parameterisation of the model by labelling experiments, simultaneously explain the proliferation, apoptosis, extracellular matrix and growth pattern of multi-cellular spheroids under different nutrient conditions within one consistent mathematical model.

- Vascular tumour growth.

- Regulatory and evolutionary aspects in tumour growth \[81\], \[78\], recently with resolution of intracellular signal transduction pathway variants found in normal vs. malignant cells \[33\].

- Cell differentiation and lineage commitment of mesenchymal stem cells \[85\], \[73\]. In our earlier work we have established a model of cell aging for in-vitro cultured stem cell populations. Stem cell concepts developed earlier \[85\], \[73\] have been extended to include cell aging \[83\]. By this extension it is possible to explain the clonal heterogeneity that was not captured by the previous model. The cell age was coupled with the generation number. It is published in ref. \[84\] and \[56\].

- Complex tissue architectures in regenerative tissues, particularly in the liver.

Examples are:

- Regeneration of liver lobules after toxic damage \[80\], \[79\], \[76\], \[40\], \[41\], \[42\] within the German BMBF-funded network “Systems Biology of the Hepatocyte”). As extension of this project we linked the regeneration of liver architecture after toxic damage to a model of ammonia detoxification by the individual hepatocyte and the liver as a proof of concept to study the link between architecture and function. The comparison of experimental findings by our collaboration partners with our model results suggests that the detoxification during regeneration after drug-induced damage is mainly determined by the total population size of healthy hepatocytes (Fig. 2). Adjustments of enzymatic activities of the individual hepatocytes seem to have only minor effects.

- Liver regeneration after partial hepatectomy \([44]\). Based on the work on regeneration of a liver after toxic damage where we focused on a single liver lobule, we within the EU project
CANCERSYS set up a model on liver regeneration after partial hepatectomy enabling us to model up to the whole liver lobe scale of a mouse, 4 cells thick. This models permits to bridge the gap between the single-cell-model scale and the whole-liver organ scale. Calibrating this model with mouse data we were able to predict the proliferation pattern in pig as a proof of principle that modelling can be used to bridge the gap between different animals. Experiments performed so far confirm the prediction. This is a fundamental issue as it is a longstanding unsolved question in how far experiments in animal models can be used to predict therapeutic responses in the Human. We also expanded the software towards whole liver lobe bright field image analysis. Being able to use a mathematical model calibrated with data from a model animal, for example, mouse to predict tissue organisation processes in another animal, for example, human opens new possibilities to assess drug toxicity [43].

- The fundamental objective is to move towards modelling from the molecular up to the whole organ scale. A conceptual framework for this was presented in ref. [23].

- Cancer development in an environment of granular particles and cells. For the first part we studied how an embedding medium such as granular particles or cells modify the spatial and temporal growth pattern of expanding cell populations [19]. The model could explain the growth kinetics found by Helmlinger et. al. (1997) for growing tumor spheroids embedded in agarose gel. If the friction between the embedding objects and the environment is larger than the friction between the growing clone and its environment we found a fingering instability reminiscent of a Saffman-Taylor instability observed in a Hele-Shaw cell. We systematically studied which model parameters promote the instability. The motivation for this project was to analyse in how far invading tumour fronts can be explained by physical mechanisms alone as in some tumour phenotypes invasive fronts can be observed.

In order to extend this project towards carcinoma development mechanisms of liver cancer development in mice were studied [42]. Within systematic sensitivity analyses we were able to identify parameters explaining the experimentally found tumor phenotypes (Fig. 3). The critical parameters were stiffness of sinusoids, the micro-blood vessels within the liver lobules, tumor cell - sinusoidal adhesion, tumor cell polarity, and the ability of tumour cells to digest neighbouring vessels. As part of the project, in close collaboration with experimental partners, critical differences between liver in normal and transgenic mice have been studied by image analysis [62].

- Synthetic biology. By multi-scale simulations including intracellular pathways in our single-cell-based simulation framework we were able to mimic conditions under which tissue homeostasis and tissue location could be achieved in monolayer culture.

The applications are guided by quantitative comparisons to experimental data either from published knowledge or - in most cases - generated by experimental partners. One main focus is on the understanding of mechanisms that control the growth dynamics and growth phenotypes of multi-cellular systems and use these later to predict and optimise therapy or biotechnological growth processes.

The adjustment of the models developed to applications requires data analysis both, of molecular data such as gene expression profiles and of image data such as spatial-temporal growth pattern. For this purpose we recently considered the geometric and topological measures to quantify tumour shapes [92], and developed an image processing chain to quantitatively analyse liver regeneration processes in liver lobules [79], [76], [44], [40], [41]. As a further step we published executables and descriptions of important elements of our code to spread our model as it turns out that agent-based cell modelling enjoys increasing interest in different communities (engineering, mathematical biology, systems biology, physics) [81]. Current directions moreover include a stronger focus on models of in-vivo systems (within the German medical systems biology consortium “LungSys” (lung cancer treatment); and within the EU-network “CancerSys” (cancerogenesis in liver)). Within LungSys we recently developed a realistic 2D and 3D spatial temporal model of blood flow in xenografts to compare to DCE MRI images visualising the tumour perfusion. Modelling cancer development requires to take into account invasion, mutations and angiogenesis, three hallmarks of cancer and of linking the molecular to the multicellular scale [71]. Moreover, we extend the topic of liver regeneration to regeneration
after partial hepatectomy (within the EU-project “Passport”), and extend our modelling activities to understand early embryonic development (Trophoblast development, collaboration with INRA).
Almost each of our projects is in close collaboration with experimental partners within grant projects performing experiments to permit parameterisation and validation of our models.

5.3. Modeling in computational neurosciences

Participants: Maria Caceres [Univ. Granada], Jose Carrillo [ICREA Barcelona], Benoît Perthame, Jonathan Touboul.

Networks of interacting neurons can be well described by nonlinear PDEs like the Noisy Integrate and Fire model. These are Fokker-Planck-Kolmogorov equations on the probability density of neurons, the main parameters in the model being the connectivity of the network and the noise. In [ 10 ], we analyse several aspects of the NNLIF model: the number of steady states, a priori estimates, blow-up issues and convergence toward equilibrium in the linear case. In particular, for excitatory networks, blow-up always occurs for initial data concentrated close to the firing potential. These results show how critical is the balance between noise and excitatory/inhibitory interactions to the connectivity parameter.

At a larger scale, neurons form large-scale spatially extended populations receiving similar input and interconnected in a specific way. Each neuron receives noisy inputs, and as such their membrane potential is adequately described as the solution of stochastic network equations. In [ 58 ] we study the asymptotic regimes of such spatially extended networks with delays and obtain a complex mean-field equation the dynamics of which is analyze in [ 57 ]. We observe that noise induces transitions from stationary spatially homogeneous solutions to oscillatory solutions, and the transition is characterized by chaotic Turing patterns of activity.

5.4. Free surface geophysical flows

Participants: Emmanuel Audusse [LAGA - Université Paris 13, Institut Galilée], Sakina Ayata, Anne-Céline Boulanger, Marie-Odile Bristeau, Benoît Perthame, Jacques Sainte-Marie [CETMEF and MACS project-team].

We are involved in research concerning the numerical simulation of free surface geophysical flows such as rivers, lakes, coastal areas and also overland flows. Many applications related to environmental problems are concerned: floodings, dam breaks, swell, transport and diffusion of pollutants, water quality, upwellings, sustainability of aquatic ecosystems, ...

The basic model for these problems is the 3D free surface Navier-Stokes system leading to a 3D solver [ 64 ] with a moving mesh. However for efficiency reasons, vertically averaged models such as the Saint-Venant system [ 75 ] are often used.

The Saint-Venant equations are deduced of the Navier-Stokes system with two main assumptions:

- the pressure is hydrostatic,
- the horizontal velocity is represented by its average.

We have developed extensions of the Saint-Venant system where the basic Saint-Venant solver [ 60 ] is still used and, in that way, the robustness, the efficiency and the easiness to treat the free surface are preserved while the domain of validity is larger.

In these extensions, we relax the two above assumptions. Actually, we have derived a non-hydrostatic shallow water model and a multilayer Saint-Venant system.

We have coupled the hydrodynamics of free surface flows with other phenomena such as biology (phytoplankton culture) or erosion.
Figure 4. Map of Japan with the seism epicentre and the DART buoys 21418 and 21413.

Figure 5. Free surface elevation of the sea, comparison between the recorded data by the buoys 21418 and 21413 and the simulation obtained with our 3d Navier-Stokes code.
5.4.1. Hydrodynamics and biology coupling

Cultivating oleaginous microalgae in specific culturing devices is seen as a potential source of biofuel for the future. The complexity of this process coupling non linear biological activity to hydrodynamics makes the optimization problem very delicate. The large amount of parameters to be taken into account paves the way for a useful mathematical modeling. Due to the high heterogeneity of raceways along the depth dimension regarding temperature, light intensity or nutrients availability, we adopt a multilayer approach for hydrodynamics and biology. For hydrodynamics, we use a multilayer Saint-Venant model that allows mass exchanges, forced by a simplified representation of the paddlewheel. Then, starting from an improved Droop model that includes light effect on algae growth, we derive a similar multilayer system for the biological part. A kinetic interpretation of the whole system results in an efficient numerical scheme. We show through numerical simulations in two dimensions that our approach is capable of discriminating between situations of moving water or calm pond and show the influence of light intensity. Moreover, in this paper [49] we exhibit that a posteriori treatment of our velocity fields can provide lagrangian trajectories which are of great interest to assess the actual light pattern perceived by the algal cells and therefore understand its impact on the cell factory.

![Figure 6. Trajectories of three particles during the simulations. In every figure, the large curve represents the water height at the middle of the pool. The other plot is the height of a given particle through time. The algae undergo sudden changes of depth every time it meets the wheel.](image)

5.4.2. Analytical solutions for the free surface hydrostatic Euler equations

In this paper [50] we propose a large set of analytical solutions for the hydrostatic incompressible Euler system in 2d and 3d. These solutions mainly concern free surface flows but flows with partially free surface or in a deformable pipe are also considered. These analytical solutions that can admit entropic shocks can be especially useful for the validation of numerical schemes.

5.4.3. Phytoplankton growth in marine ecosystem

Four different phytoplankton growth models have been implemented. The simplest model assumes constant chlorophyll/carbon and carbon/nitrogen ratios. The more complex ones take into account photoadaptation through a variable chlorophyll/carbon ratio and they also assume a variable carbon/nitrogen cellular quota.
(non-redfieldian stochiometry). These models have been coupled to a 1D ecosystem model at BATS, a station located in an oligotrophic area of the North-Western Atlantic Ocean. The different models have been calibrated from in situ data recorded at BATS using a micro-genetic algorithm to optimize the parameter values. The models with optimized parameters were then compared with each others. The results highlighted the necessity to take into account photoadaptation and variable cellular quotas to simulate the seasonal dynamics of chlorophyll and primary production in oligotrophic areas. They also demonstrated that the chlorophyll did not have to be represented by a prognostic variable and could be represented by a diagnostic variable instead, see [46].

5.4.4. Erosion processes: modelling and simulation

We are interested in the modelling of sediment transport phenomena. We mostly focus on bedload transport and we do not consider suspension sediment processes. We first propose a coupled numerical scheme for the classical Saint-Venant – Exner model. It is based on a relaxation approach and it works with all sediment flux function. We exhibit that this coupled approach is more stable than the splitting approach that is mostly used in industrial softwares. Then we derive an original three layers model in order to overcome the difficulties that are encountered when using the classical Exner approach and we write a related relaxation model, see [45].
5. New Results

5.1. Astrocyte Regulation of Synaptic Depression and Facilitation

Participants: Hugues Berry, Maurizio De Pitta, Vladislav Volman, Eshel Ben-Jacob.

Synaptic plasticity is the capacity of a preexisting connection between two neurons to change in strength as a function of neuronal activity. Because it admittedly underlies learning and memory, the elucidation of its constituting mechanisms is of crucial importance in many aspects of normal and pathological brain function. Short-term presynaptic plasticity refers to changes occurring over short time scales (milliseconds to seconds) that are mediated by frequency-dependent modifications of the amount of neurotransmitter released by presynaptic stimulation. Recent experiments have reported that glial cells, especially hippocampal astrocytes, can modulate short-term plasticity, but the mechanism of such modulation is poorly understood. Here, we explore a plausible form of modulation of short-term plasticity by astrocytes using a biophysically realistic computational model. Our analysis indicates that astrocytes could simultaneously affect synaptic release in two ways. First, they either decrease or increase the overall synaptic release of neurotransmitter. Second, for stimuli that are delivered as pairs within short intervals, they systematically increase or decrease the synaptic response to the second one. Hence, our model suggests that astrocytes could transiently trigger switches between paired-pulse depression and facilitation. This property explains several challenging experimental observations and has a deep impact on our understanding of synaptic information transfer [16].

5.2. A theory of rate coding control by intrinsic plasticity effects

Participants: J Naudé, J Paz, Hugues Berry, Bruno Delord.

Intrinsic plasticity (IP) is a ubiquitous activity-dependent process regulating neuronal excitability and a cellular correlate of behavioral learning and neuronal homeostasis. Because IP is induced rapidly and maintained long-term, it likely represents a major determinant of adaptive collective neuronal dynamics. However, assessing the exact impact of IP has remained elusive. Indeed, it is extremely difficult to disentangle the complex non-linear interaction between IP effects, by which conductance changes alter neuronal activity, and IP rules, whereby activity modifies conductance via signaling pathways. Moreover, the two major IP effects on firing rate, threshold and gain modulation, remain unknown in their very mechanisms. Here, using extensive simulations and sensitivity analysis of Hodgkin-Huxley models, we show that threshold and gain modulation are accounted for by maximal conductance plasticity of conductance in two separate domains of the parameter space corresponding to sub- and supra-threshold conductance (i.e. activating below or above the spike onset threshold potential). Analyzing equivalent integrate-and-fire models, we provide formal expressions of sensitivities relating to conductance parameters, unraveling unprecedented mechanisms governing IP effects. Our results generalize to the IP of other conductance parameters and allow inference of calcium-gated conductance, yielding a general picture that accounts for a large repertoire of experimental observations. The expressions we provide can be combined with IP rules in rate or spiking models, offering a general framework to systematically assess the computational consequences of IP of pharmacologically identified conductance with both fine grain description and mathematical tractability. We provide an example of such IP loop model addressing the important issue of the homeostatic regulation of spontaneous discharge. Because we do not formulate any assumptions on modification rules, the present theory is also relevant to other neural processes involving excitability changes, such as neuromodulation, development, aging and neural disorders [20].

5.3. Impact of receptor clustering on ligand binding

Participants: Hédi A. Soula, Bertrand Caré.
Cellular response to changes in the concentration of different chemical species in the extracellular medium is induced by ligand binding to dedicated transmembrane receptors. Receptor density, distribution, and clustering may be key spatial features that influence effective and proper physical and biochemical cellular responses to many regulatory signals. Classical equations describing this kind of binding kinetics assume the distributions of interacting species to be homogeneous, neglecting by doing so the impact of clustering. As there is experimental evidence that receptors tend to group in clusters inside membrane domains, we investigated the effects of receptor clustering on cellular receptor ligand binding. We implemented a model of receptor binding using a Monte-Carlo algorithm to simulate ligand diffusion and binding. In some simple cases, analytic solutions for binding equilibrium of ligand on clusters of receptors are provided, and supported by simulation results. Our simulations show that the so-called “apparent” affinity of the ligand for the receptor decreases with clustering although the microscopic affinity remains constant. Changing membrane receptors clustering could be a simple mechanism that allows cells to change and adapt their affinity/sensitivity toward a given stimulus [14].

5.4. Illegitimate and Homologous Rearrangements, an intricate relationship?

Participants: David P. Parsons, Guillaume Beslon, Carole Knibbe.

We have introduced homologous horizontal transfer in the aevol model. First results show that this process interacts in a complex way with both non-homologous transfer (that creates homologous sequences) and mutation rate (that destroys homologous sequences). We have shown that homologous transfer is useful only in strict conditions (small mutation rate and small - but non-null - non-homologous rate) [27]. This result confirms the genericity of the indirect selection mechanism we previously shown to be at the origin of scaling laws in genomes and transcriptomes [25].

5.5. An error threshold due to chromosomal rearrangements

Participants: Stephan Fischer, Guillaume Beslon, Carole Knibbe, Samuel Bernard.

Error threshold is a well known theory in evolutionary biology. However, the theory of error threshold only takes into account point mutation rate and states that this rate generates a maximum level of coding sequences in a given genome. Using aevol and mathematical formulations, we have shown other types of mutations are also likely to create error thresholds. In particular we have shown the chromosomal rearrangement (duplications and deletions) generate a very strong threshold.

5.6. Enhanced Stimulus Encoding Capabilities with Spectral Selectivity in Inhibitory Circuits by STDP

Participants: Guillaume Beslon, Hédi A. Soula, Antoine Coulon.

The ability to encode and transmit a signal is an essential property that must demonstrate many neuronal circuits in sensory areas in addition to any processing they may provide. It is known that an appropriate level of lateral inhibition, as observed in these areas, can significantly improve the encoding ability of a population of neurons. We have shown that a homeostatic mechanism by which a spike-timing-dependent plasticity (STDP) rule with a symmetric timing window (swSTDP) spontaneously drives the inhibitory coupling to a level that ensures accurate encoding in response to input signals within a certain frequency range. Interpreting these results mathematically, we find that this coupling level depends on the overlap of spectral information between stimulus and STDP window function. Generalization to arbitrary swSTDP and arbitrary stimuli reveals that the signals for which this improvement of encoding takes place can be finely selected on spectral criteria. We finally show that this spectral overlap principle holds for a variety of neuron types and network characteristics. The highly tunable frequency-power domain of efficiency of this mechanism, together with its ability to operate in very various neuronal contexts, suggest that it may be at work in most sensory areas [15].
6. New Results

6.1. Modern methods of data analysis

Participants: H. Cardot, P. Cénac, O. Collignon, J-M. Monnez, P. Vallois.

In 2011, our contributions to data analysis in a Biological context are twofold:

- At a theoretical level, we have kept on working on the so-called online data analysis alluded to at the Scientific Foundations Section. Specifically, we have carried on the construction of a fast and recursive algorithm for clustering large data sets with the $k$-medians methods.
- At a practical level, our efforts have focused on an interesting study concerning peanuts allergy, for which our expertise in data analysis allows for a good prediction of allergy severity by means of rigorous methods.

Let us now describe more precisely our articles:

(i) A fast and recursive algorithm for clustering large data sets with $k$-medians. Clustering with fast algorithms large samples of high dimensional data is an important challenge in computational statistics. Borrowing ideas from MacQueen [56], who introduced a sequential version of the $k$-means algorithm, a new class of recursive stochastic gradient algorithms designed for the $k$-medians loss criterion is proposed in [16], [17]. By their recursive nature, these algorithms are very fast and well adapted to deal with large samples of data that are allowed to arrive sequentially. It is proved that the stochastic gradient algorithm converges almost surely to the set of stationary points of the underlying criterion. A particular attention is paid to the averaged versions, which are known to have better performances, and a data-driven procedure that allows automatic selection of the value of the descent step is proposed. The performance of the averaged sequential estimator is compared on a simulation study, both in terms of computation speed and accuracy of the estimations, with more classical partitioning techniques such as $k$-means, trimmed $k$-means and PAM (partitioning around medoids). Finally, this new on-line clustering technique is illustrated on determining television audience profiles with a sample of more than 5000 individual television audience measured every minute over a period of 24 hours.

(ii) Discriminant analyses of peanut allergy severity scores. Peanut allergy is one of the most prevalent food allergies. The possibility of a lethal accidental exposure and the persistence of the disease make it a public health problem. Evaluating the intensity of symptoms is accomplished with a double blind placebo-controlled food challenge (DBPCFC), which scores the severity of reactions and measures the dose of peanut that elicits the first reaction. Since DBPCFC can result in life-threatening responses, we propose in [2] an alternate procedure with the long-term goal of replacing invasive allergy tests. Discriminant analysis of DBPCFC score, the eliciting dose and the first accidental exposure score were performed in 76 allergic patients using 6 immunoassays and 28 skin prick tests. A multiple factorial analysis was performed to assign equal weights to both groups of variables, and predictive models were built by cross-validation with linear discriminant analysis, $k$-nearest neighbors, classification and regression trees, penalized support vector machine, stepwise logistic regression and Adaboost methods. We developed an algorithm for simultaneously clustering eliciting doses and selecting discriminant variables. Our main conclusion is that antibody measurements offer information on the allergy severity, especially those directed against $rAra\_h1$ and $rAra\_h3$. Further independent validation of these results and the use of new predictors will help extend this study to clinical practices.

6.2. Local linear estimator of the conditional distribution function

Participants: S. Ferrigno, M. Maumy, A. Muller.
Consider \((X, Y)\), a random vector defined in \(\mathbb{R} \times \mathbb{R}\). Here \(Y\) is the variable of interest and \(X\) the concomitant variable. As usual in the statistics literature, we work under the assumption that a sample \(\{(X_i, Y_i)_{1\leq i\leq n}\}\) of independent and identically distributed \((X, Y)\) is available.

In order to explain the relationship between the variable of interest \(Y\) and the factor \(X\), the standard way is to rely on the regression function \(E(Y|X = x)\). Because of numerous applications, the problem of estimating the regression function has been the subject of considerable interest during the last decades. However, it can be easily argued that the function \(x \mapsto E(Y|X = x)\) alone does not capture the complexity of the relations between \(X\) and \(Y\).

In order to go one step further in this direction, we have chosen to work with another function. Namely, we study the conditional distribution function \(F(y|X = x) = P(Y \leq y|X = x)\) and a nonparametric estimator associated to this quantity. The distribution function has the advantage of completely characterizing the law of the random variable at stake, allowing to obtain the regression function, the density function, the moments and the quantile function. It should also be noticed that conditional distribution functions are used for the estimation of references curves in medical applications.

At a more technical level, our study is based on a local linear nonparametric estimator of the conditional distribution function instead of the widely spread Nadaraya-Watson estimator. Indeed, it is a well-known fact that the asymptotic bias of the Nadaraya-Watson estimator behaves somehow badly. Observe however that local polynomial techniques are good alternatives. Based on these techniques, here are the steps we have focused on in 2010-2011:

- Our main result is the uniform law of the logarithm concerning the local linear estimator of the conditional distribution function (see [21]). We investigate convergence in probability and almost sure convergence results.
- The uniform law of the logarithm has then been used to construct uniform asymptotic certainty bands for the conditional distribution function.
- The certainty bands alluded to above have been applied to simulated data.
- A variant of the test has been introduced in [20].

Let us also mention that applications of these theoretical results to survival analysis are currently the object of active research.

### 6.3. Markovian models for tumor growth

Participants: T. Bastogne, R. Keinj, P. Vallois.

Our research in this direction includes two contributions in 2011:

- A multinomial model for cell growth allowing to calibrate radiotherapies given in [3].
- A study of tumor growth based on the lifespan of each cell (see [13]).

More specifically, our two contributions can be summarized as follows:

(i) Hit and target models of tumor growth typically assume that all surviving cells have a constant and homogeneous sensitivity during the radiotherapy period. In [3], we propose a multinomial model based on a discrete-time Markov chain, able to take into account cell repair, cell damage heterogeneity and cell proliferation. The proposed model relies on the ‘Hit paradigm’ and ‘Target’ theory in radiobiology and assumes that a cancer cell contains \(m\) targets which must be all deactivated to produce cell death. The surviving cell population is then split up into \(m\) categories to introduce the variation of cancer cell radio-sensitivity according to their damage states. Two other parameters have been introduced: the probability \(q\) for a target to be deactivated by radiation and the probability \(r\) for an inactive target in an alive cell to be reactivated. The parameter \(q\) is related to the radiation dose \(u_0\) through the intrinsic sensitivity of a target to radiation. Moreover, the multinomial model is a generalization of typical hit models. Based on the multinomial model, new expressions of the TCP (Tumor Control Probability) and NTCP (Normal Tissue Complication Probability)
have been proposed for nonuniform radiations which permits to deduce the optimal total dose to be delivered. We point out the important influence of the repair parameter \( r \) which could lead to reduce both the total radiation dose to be delivered and the risk of side effects.


\( \text{(ii)} \) We have proposed in [13] an original approach that expresses the probability distribution of the cancer and normal cells lifespans in terms of the number of dose fractions in radiotherapy. Conversely to previous models that examines the number of surviving cells in the treated population at fixed time instants, our modeling approach better reveals the dynamics of the tumor response.

We start by considering the lifespan of a single cancer cell that behaves as described in [3]. We study this random time by calculating its mean, variance and cumulative distribution function. We then assume that a tumor is a group of independent cells. This allows to define the lifespan of the tumor as the maximum of individual lifespans. When the initial number \( n_0 \) of cancer cells is not too large, then we can explicitly calculate the mean, variance and the cumulative distribution function of the tumor lifespan. When \( n_0 \) is large, the previous parameters are no longer calculable. However, we are able to show that, under some assumptions, the mean lifespan of the tumor behaves as a logarithmic function of the initial number \( n_0 \). The second goal is to show that TCP and NTCP can be completely formulated with respect to the tumor and normal tissue lifespans. These expressions of TCP and NTCP are finally used to propose a ROC curve, called ECT (Efficiency-Complication Trade-off), suited to the determination of the appropriate treatment schedule. This synthetic representation summarizes both efficiency and complication of the treatment. Moreover, it allows several possibilities of choice for the radiotherapist: treatment efficiency, priority to safety of normal tissue, or a trade-off between them.

6.4. A stochastic model for bacteriophage therapies


In the last years Bacteriophage therapies are attracting the attention of several scientific studies. They can be a new and powerful tool to treat bacterial infections or to prevent them applying the treatment to animals such as poultry or swine. Very roughly speaking, they consist in inoculating a (benign) virus in order to kill the bacteria known to be responsible of a certain disease. This kind of treatment is known since the beginning of the 20th century, but has been in disuse in the Western world, erased by antibiotic therapies. However, a small activity in this domain has survived in the USSR, and it is now re-emerging (at least at an experimental level). Among the reasons of this re-emersion we can find the progressive slowdown in antibiotic efficiency (antibiotic resistance). Reported recent experiments include animal diseases like hemorrhagic septicemia in cattle or atrophic rhinitis in swine, and a need for suitable mathematical models is now expressed by the community.

Let us be a little more specific about the (lytic) bacteriophage mechanism: after attachment, the virus’ genetic material penetrates into the bacteria and use the host’s replication mechanism to self-replicate. Once this is done, the bacteria is completely spoiled while new viruses are released, ready to attack other bacteria. It should be noticed at this point that among the advantages expected from the therapy is the fact that it focuses on one specific bacteria, while antibiotics also attack autochthonous microbiota. Roughly speaking, it is also believed that viruses are likely to adapt themselves to mutations of their host bacteria.

At a mathematical level, whenever the mobility of the different biological actors is high enough, bacteriophage systems can be modeled by a kind of predator-prey equation. Namely, set \( S_t \) (resp. \( Q_t \)) for the bacteria (resp. bacteriophages) concentration at time \( t \). Then a model for the evolution of the couple \((S, Q)\) is as follows:

\[
\begin{align*}
\frac{dS_t}{dt} &= [\alpha - k Q_t] S_t dt + \varepsilon S_t dW^1_t \\
\frac{dQ_t}{dt} &= [d - mQ_t - k Q_t S_t + k b e^{-\mu_0} Q_{t-\zeta} S_{t-\zeta} - \mu Q_t - \varepsilon] dt + \varepsilon Q_t dW^2_t,
\end{align*}
\]

(35)
where $\alpha$ is the reproducing rate of the bacteria and $k$ is the adsorption rate. In equation (1), $d$ also stands for the quantity of bacteriophages inoculated per unit of time, $m$ is their death rate, we denote by $b$ the number of bacteriophages which is released after replication within the bacteria cell, $\zeta$ is the delay necessary to the reproduction of bacteriophages (called latency time) and the coefficient $e^{-\mu\zeta}$ represents an attenuation in the release of bacteriophages (given by the expected number of bacteria cell’s deaths during the latency time, where $\mu$ is the bacteria’s death rate). A given initial condition $(S_0, Q_0)$ is also specified, and the term $\varepsilon \, dW_i$ takes into account a small external noise standing for both uncertainties on the measures and the experiment conditions (for similar modeling see e.g. [34]). One should be aware of the fact that the latency time $\zeta$ (which can be seen as the reproduction time of the bacteriophages within the bacteria) cannot be neglected, and is generally of the same order (about 20mn) as the experiment length (about 60mn).

With this model in hand, our main results in this direction (see [15]) have been the following:

- Quantification of the exponential convergence to a bacteria-free equilibrium of equation (1) when $d$ is large enough.
- Use of the previous result plus concentration inequalities in order to study the convergence of the noisy system to equilibrium in a reasonable time range.
- Simulation of the stochastic processes at stake in order to observe the convergence to equilibrium.

### 6.5. Convergence of stochastic gene networks

Participants: A. Crudu, A. Debussche, A. Muller, Aurélie, O. Radulescu.

We propose simplified models for the stochastic dynamics of gene network models arising in molecular biology. Those gene networks are classically modeled by Markov jump processes, which are extremely time consuming. To overcome this drawback, we study the asymptotic behavior of multiscale stochastic gene networks using weak limits of Markov jump processes.

We consider a set of chemical reactions $R_i$, $i \in \mathbb{R}$; $\mathbb{R}$ is supposed to be finite. These reactions involve species indexed by a set $S = 1, \cdots, M$, the number of molecules of the species $i$ is denoted by $n_i$ and $X \in \mathbb{N}^M$ is the vector consisting of the $n_i$’s. Each reaction $R_i$ has a rate $\lambda_i(X)$ which depends on the state of the system, described by $X$ and corresponds to a change $X \rightarrow X + \gamma_i$, $\gamma_i \in \mathbb{Z}^M$.

Mathematically, this evolution can be described by the following Markov jump process. It is based on a sequence $(\tau_k)_{k \geq 1}$ of random waiting times with exponential distribution. Setting $T_0 = 0$, $T_i = \tau_1 + \cdots + \tau_i$, $X$ is constant on $[T_{i-1}, T_i)$ and has a jump at $T_i$. The parameter of $\tau_i$ is given by $\sum_{r \in \mathbb{R}} \lambda_r(X(T_{i-1}))$:

$$P(\tau_i > t) = \exp \left( - \sum_{r \in \mathbb{R}} \lambda_r(X(T_{i-1})) t \right).$$

At time $T_i$, a reaction $r \in \mathbb{R}$ is chosen with probability $\lambda_r(X(T_{i-1}))/\sum_{r \in \mathbb{R}} \lambda_r(X(T_{i-1}))$ and the state changes according to $X \rightarrow X + \gamma_r$: $X(T_i) = X(T_{i-1}) + \gamma_r$. This Markov process has the following generator:

$$Af(X) = \sum_{r \in \mathbb{R}} [f(X + \gamma_r) - f(X)] \lambda_r(X).$$

In the applications we have in mind, the numbers of molecules have different scales. Some of the molecules are in small numbers and some are in large numbers. Accordingly, we split the set of species into two sets $C$ and $D$ with cardinals $M_C$ and $M_D$. This induces the decomposition $X = (X_C, X_D)$, $\gamma_r = (\gamma_r^C, \gamma_r^D)$. For $i \in D$, $n_i$ is of order 1 while for $i \in C$, $n_i$ is proportional to $N$ where $N$ is a large number. For $i \in C$, setting $\tilde{n}_i = n_i/N$, $\tilde{n}_i$ is of order 1. We define $x_C = X_C/N$ and $x = (x_C, X_D)$. 


---

**Note:** The page contains mathematical and scientific content related to stochastic gene networks and their convergence, involving reaction rates, probabilities, and exponential distributions. The text is a detailed exposition of the mathematical models and results, including the formulation of the Markov jump processes and the conditions for convergence.
For this kind of system, we are able to give in [18] some relevant information on the asymptotic regime $N \to \infty$ when different type of reactions are involved. Depending on the time and concentration scales of the system we distinguish four types of limits:

- Continuous piecewise deterministic processes (PDP) with switching.
- PDP with jumps in the continuous variables.
- Averaged PDP.
- PDP with singular switching.

We justify rigorously the convergence for the four types of limits.

### 6.6. Inference for Gaussian systems


(i) **LAN property for fractional Brownian motion.** Local asymptotic normality (LAN) property is a fundamental concept in asymptotic statistics, which gives the asymptotic normality of certain estimators such as the maximum likelihood estimator for instance (see [66] for details on this property). In [11], we focus on the LAN property for the model where we observe a sample of $n$ observations $X_n = (X_1, \ldots, X_n)$ of a Gaussian stationary sequence. The sequence $(X_n)_{n \in \mathbb{N}}$, whose spectral density $f_\theta$ is indexed by a parameter $\theta$, can admit antipersistence, long memory or short memory and be noninvertible. To be more specific, our main assumption is:

$$f_\theta(x) \sim_{x \to 0} |x|^{-\alpha(\theta)} L_\theta(x)$$

with $L_\theta$ a slowly varying function and $\alpha(\theta) \in (-\infty, 1)$. We prove the LAN property by studying an asymptotic expansion of the log likelihood and using some results on Toeplitz matrices (see [39], [53]). In particular, our assumptions are fulfilled by fractional Gaussian noises and autoregressive fractionally integrated moving average processes (ARFIMA($p,d,q$)). We also obtain the LAN property for fractional Brownian motion.

(ii) **Inference for dynamical systems driven by Gaussian noises.** As mentioned at the Scientific Foundations Section, the problem of estimating the coefficients of a general differential equation driven by a Gaussian process is still largely unsolved. To be more specific, the most general ($\mathbb{R}$-valued) equation handled up to now as far as parameter estimation is concerned (see [64]) is of the form:

$$X_i^\theta = a + \theta \int_0^t b(X_u) \, du + B_t,$$

where $\theta$ is the unknown parameter, $b$ is a smooth enough coefficient and $B$ is a one-dimensional fractional Brownian motion. In contrast with this simple situation, our applications of interest (see the Application Domains Section) require the analysis of the following $\mathbb{R}^n$-valued equation:

$$X_i^\theta = a + \int_0^t b(\theta; X_u) \, du + \int_0^t \sigma(\theta; X_u) \, dB_t,$$

(36)

where $\theta$ enters non linearly in the coefficient, where $\sigma$ is a non-trivial diffusion term and $B$ is a $d$-dimensional fractional Brownian motion. We have thus decided to tackle this important scientific challenge first.
To this aim, here are the steps we have focused on in 2011:

- A better understanding of the underlying rough path structure for equation (2), carried out in [4], [5]. This step allows a proper definition of our equation of interest in a wide range of contexts.
- Gaussian type bounds for equations driven by a fractional Brownian motion, obtained in [9]. This is an important preliminary step for likelihood estimates for stochastic processes.
- Numerical aspects of a maximum likelihood type procedure for an equation of the form (2), expressed in terms of Malliavin calculus tools (see [10]).
- Convergence of a least square type estimator for an equation of the form (2) where the noise enters additively, handled in [14]. This is the first occurrence of a converging estimator for a general coefficient $b(\theta, \cdot)$.

### 6.7. Local self-similarity properties and stable or Gaussian random fields

Participants: Hermine Biermé, Jacques Istas, Céline Lacaux, Renaud Marty, Hans-Peter Scheffler.

- Recently, an important class of anisotropic random fields called operator scaling random fields has been studied in [30]. To be more specific, the classical self-similarity property is replaced in [30] by the following operator scaling property:

$$\forall c > 0, \left( X(c \cdot x) \right)_{x \in \mathbb{R}^d} \overset{d}{=} c \left( X(x) \right)_{x \in \mathbb{R}^d}, \text{ where } c^E := \exp (E \ln(c)).$$

The Hölder regularity properties of operator scaling Gaussian or stable harmonizable random fields have been studied in [30] and can be expressed in terms of the matrix $E$. In particular, they do not vary along the trajectories, which can be too restrictive for some applications (see our osteoporosis project at the Application Domains Section). In order to obtain some anisotropic random fields whose Hölder regularity properties are allowed to vary, we introduce in [1] a local version of the operator scaling property (similar to the local version of the classical self-similarity property defined in [27]). This local property is illustrated in [1], where we also define and study harmonicizable multi-operator scaling stable random fields. For such a multi-operator random field, we obtain an accurate upper bound of both the modulus of continuity and global and directional Hölder regularities at any point $x$. As expected, the Hölder regularity properties vary along the trajectories.

- In [24], we study the sample paths properties of an anisotropic random field, which is defined as limit of an invariance principle and is of the same type as a multifractional Brownian sheet. Our first aim was to generalize [37], that is to obtain some multifractional random fields indexed by $\mathbb{R}^d$ with $d \geq 2$ and to allow Hurst indices to be lower than $1/2$. To overcome the problem of the values of the Hurst indices which characterize the limit field, we focus on stationary sequences $(X_n(H))_{n \in \mathbb{N}}$, where $H \in (0,1)^d$, defined by an harmonizable representation. Then, our limit field $S_h$ is defined as the limit of

$$S_h^N = \left\{ \sum_{n_1=1}^{[N_1]} \ldots \sum_{n_d=1}^{[N_d]} \frac{X_n(h_{N_n})}{N^{r_N}}; t \in [0, +\infty)^d \right\}$$

for some suitable families $(h_{N_n})_{n,N}$ and $(r_{N_n})_{n,N}$. We then study the sample paths property of this limit field. In particular, we obtain some local self-similarity properties for its increments of order $k$ and its pointwise global and directional Hölder exponents. We also define (and obtain) some pointwise multi-Hölder exponents which characterize the Hölder property satisfied by the increments of order $d$ of $S_h$.

- We are also interested in self-similar processes indexed by manifolds in [23]. This study is motivated by the fact various spatial data are indexed by a manifold and not by the Euclidean space $\mathbb{R}^d$ in practical situations such as image analysis.
6. New Results

6.1. Mathematical methods and methodological approach to biology

6.1.1. Mathematical analysis of biological models

Participants: Jean-Luc Gouzé, Olivier Bernard, Frédéric Grognard, Ludovic Mailleret, Pierre Bernhard, Andrei Akhmetzhanov.

Mathematical study of semi-discrete models

Semi-discrete models have shown their relevance in the modelling of biological phenomena whose nature presents abrupt changes over the course of their evolution [77]. We used such models and analysed their properties in several situations that are developed in 6.2.3., most of them requiring such a modelling in order to take seasonality into account. Such is the case when the year is divided into a cropping season and a ‘winter’ season, where the crop is absent, as in our analysis of the sustainable management of crop resistance to pathogens [58] or in the co-existence analysis of epidemiological strains [17], [21]. Seasonality also plays a big role in the semi-discrete modelling required for the analysis of consumers’ adaptive behaviour in seasonal consumer-resource dynamics, where only dormant offspring survives the ‘winter’ [12], [52].

Mathematical study of models of competing species

When several species are in competition for a single substrate in a chemostat, and when the growth rates of the different species only depend on the substrate, it is known that the generic equilibrium state for a given dilution rate consists in the survival of only one of the species. In [28], we propose a model of competition of \( n \) species in a chemostat, where we add constant inputs of some species. We achieve a thorough study of all the situations that can arise when having an arbitrary number of species in the chemostat inputs; this always results in a Globally Asymptotically Stable equilibrium where all input species are present with at most one of the other species.

6.1.2. Model design, identification and validation

Participants: Olivier Bernard, Jean-Luc Gouzé.

Model design and identification

One of the main families of biological systems that we have studied involves mass transfer between compartments, whether these compartments are microorganisms or substrates in a bioreactor, or species populations in an ecosystem. We have developed methods to estimate the models of such systems [62]. These systems can be represented by models having the general structure popularized by [61], [65], [66], and based on an underlying reaction network:

\[
\frac{d\xi}{dt} = K r(\xi, \psi) + D(\xi_{in} - \xi) - Q(\xi)
\]

We address two problems: the determination of the pseudo-stoichiometric matrix \( K \) and the modelling of the reaction rates \( r(\xi, \psi) \).

In order to identify \( K \), a two-step procedure has been proposed. The first step is the identification of the minimum number of reactions to be taken into account to explain a set of data. If additional information on the process structure is available, we showed how to apply the second step: the estimation of the pseudo-stoichiometric coefficients.

This approach has been applied to various bioproduction processes, most recently on activated sludge processes [60], anaerobic digestion [71], [87] and anaerobic digestion of microalgae [22].
6.1.3. Nonlinear observers

Participants: Jean-Luc Gouzé, Olivier Bernard.

Interval observers

Interval observers give an interval estimation of the state variables, provided that intervals for the unknown quantities (initial conditions, parameters, inputs) are known [8]. We have extended the interval observer design to new classes of systems. First, we designed interval observers, even when it was not possible in the original basis, by introducing a linear, time-varying change of coordinates, [80]. This approach was then extended to n-dimensional linear systems, leading to the design of interval observers in high dimensions [25]. Extension to time-delay systems have also been proposed [26]. The combination of the observers has also been improved in the case where various types of interval observers are run in parallel in a so-called “bundle of observers” [64]. The approach has been applied to estimation of the microalga growth and lipid production [39].

In order to demonstrate the efficiency of the interval observer design, even with chaotic systems, a special application of the interval observer has been developed for Chua’s chaotic systems. The interval estimation of the state variables are performed considering parameters uncertainties of the system and biased output [83], [80].

6.1.4. Metabolic and genomic models

Participants: Jean-Luc Gouzé, Madalena Chaves, Alfonso Carta, Ismail Belgacem, Xiao Dong Li, Olivier Bernard, Christian Breindl, Frédéric Grognard.

Qualitative control of piecewise affine models

In the control of genetic networks, the construction of feedback control laws is subject to many specific constraints, including positivity, appropriate bounds and forms of the input. In addition, control laws should be liable to implementation in the laboratory using gene and protein components. In this context, under the hypothesis that both the observations and control functions are qualitative (or piecewise constant), and using sliding mode solutions, we analysed the controllability and stabilizability with respect to either of the steady states, for a piecewise affine model of the bistable switch with single input. It is also possible to find a qualitative control law that leads the system to a periodic orbit passing through the unstable steady state [14]. Moreover, we designed some preliminary control laws for the negative loop with two genes, which has an oscillatory behaviour [54].

Interconnections of Boolean modules: asymptotic behaviour

A biological network can be schematically described as an input/output Boolean module: that is, both the states, the outputs and inputs are Boolean. The dynamics of a Boolean network can be represented by an asynchronous transition graph, whose attractors describe the system’s asymptotic behaviour. We have shown that the attractors of the feedback interconnection of two Boolean modules can be fully identified in terms of cross-products of the attractors of each module. Based on this result, a model reduction technique is proposed in [33], to predict the asymptotic dynamics of high-dimensional biological networks through the computation of the dynamics of two isolated smaller subnetworks. Applications include a model of cell-fate decision (represented as an interconnection of two 3-input/3-output modules).

Structure estimation for unate Boolean models of gene regulation networks

Estimation or identification of the network of interactions among a group of genes is a recurrent problem in the biological sciences. Together with collaborators from the University of Stuttgart, we have worked on the reconstruction of the interaction structure of a gene regulation network from qualitative data in a Boolean framework. The idea is to restrict the search space to the class of unate functions. Using sign-representations, the problem of exploring this reduced search space is transformed into a convex feasibility problem. The sign-representation furthermore allows to incorporate robustness considerations and gives rise to a new measure which can be used to further reduce the uncertainties. The proposed methodology is demonstrated with a Boolean apoptosis signaling model [70].
Multistability and oscillations in genetic control of metabolism

Genetic feedback is one of the mechanisms that enables metabolic adaptations to environmental changes. The stable equilibria of these feedback circuits determine the observable metabolic phenotypes. Together with D. Oyarzun from Imperial College, we considered an unbranched metabolic network with one metabolite acting as a global regulator of enzyme expression. Under switch-like regulation and exploiting the time scale separation between metabolic and genetic dynamics, we developed geometric criteria to characterize the equilibria of a given network. These results can be used to detect mono- and bistability in terms of the gene regulation parameters for any combination of activation and repression loops [40]. We also find that metabolic oscillations can emerge in the case of operon-controlled networks; further analysis reveals how nutrient-induced bistability and oscillations can emerge as a consequence of the transcriptional feedback [27].

Uniqueness and global stability for metabolic models

We are interested in the uniqueness and stability of the equilibrium of reversible metabolic models. For biologists, it seems clear that realistic metabolic systems have a single stable equilibrium. However, it is known that some type of metabolic systems can have no or multiple equilibria. We have made some contribution to this problem, in the case of a totally reversible enzymatic system. We prove that the equilibrium is globally asymptotically stable if it exists; we give conditions for existence and behaviour in a more general genetic-metabolic loop [84].

Birhythmicity in the p53-Mdm2 network

The p53-Mdm2 network is one of the key protein module involved in the control of proliferation of abnormal cells in mammals. Recently, a differential model of the p53-Mdm2 biochemical network which shows birhythmicity has been proposed to reproduce the two experimentally observed frequencies of oscillations of p53. Our study aimed at investigating the mechanisms at the origin of this birhythmic behaviour. To do so, we approximated this continuous non-linear model into a lower dimensional piecewise affine model and performed a first return map analysis. Based on this analysis, an experimental strategy has been proposed to test the existence of birhythmicity in the p53-Mdm2 network [31], [11].

E. coli modelling and control

In the framework of ANR project Gemco, with the aim of better understanding how to build a controller for Escherichia coli growth, three reduced models of E. coli gene expression machinery have been developed: the wild-type model, the open-loop model and the closed-loop model. Each one of these models is made up of two piece-wise non-linear differential equations.

Notably, the wild-type model describes the qualitative dynamics of the unmodified bacterium in terms of two relevant protein concentrations (RNAP and CRP) and a carbon source as input, which can be either glucose or maltose. Bacteria prefer glucose, which leads to a higher growth rate in wild type.

The open-loop model describes the qualitative dynamics of RNAP and CRP concentrations when the gene encoding for RNAP is controlled externally by an IPTG-inducible promoter. This control, biologically implemented by our collaborators in Grenoble—yielding different E. coli growth rates depending on IPTG concentration—shows that a controller for the bacterial growth can be built acting on the gene expression machinery level. Moreover, lumped parameters related to RNAP dynamics and IPTG regulation function have been estimated by means of E. coli growth curves.

Finally, the closed-loop model implements a possible feedback control-loop able to theoretically generate inverse diauxie, i.e. higher growth on maltose than on glucose.

Observation problems of a class of genetic regulatory networks

A state reconstruction problem with Boolean measurements is considered for a piecewise affine genetic network model. The problem has two distinct aspects with respect to classical ones: 1) the model is a hybrid system, and 2) the measurements (of genes expression) are only qualitative due to the experimental techniques. A Luenberger-like observer is proposed which can present some sliding modes and has finite-time convergence. A transition graph is given for the coupled observer-nominal system. To minimize the
convergence time, different convergence scenarios are discussed for optimizing the choice of initial condition of the observer \[35\], \[57\]. Robustness of the observer is checked for two types of parametric perturbed systems: 1) the observer is used to identify an unknown but fixed variation on the synthesis coefficient, via an adaptive dichotomy algorithm; and 2) the observer is robust in practical sense for the model with an uncertainty on the threshold value \[76\].

6.2. Fields of application

6.2.1. Bioenergy

6.2.1.1. Modelling and optimization of microalgae production

**Participants:** Olivier Bernard, Antoine Sciandra, Frédéric Grognard, Philipp Hartmann, Rafael Munoz Tamayo, Andrei Akhmetzhanov, Nina Moelants, Hubert Bonnefond.

Experiments have been carried out to study the effects of nitrogen limitation on the lipid production in microalgae \[74\] and support model development. We have proposed a new model for lipid production by microalgae which describes the fate of the CO\(_2\) incorporated during photosynthesis \[23\]. This model describes the accumulation of neutral lipids (which can be turned into biofuel), carbohydrates and structural carbon. It has been calibrated and validated with experimental data. Experiments have also attempted to simultaneously represent the effect of an osmotic stress \[55\]. This model highlights and explains the phenomenon of hysteresis in lipid production which has been experimentally verified. It has been extended to account for light/dark cycles \[36\].

On the other hand, a new dynamical model has been developed to describe microalgal growth in a photobioreactor under light and nitrogen limitations \[13\]. The strength of this model is that it takes into account the strong interactions between the biological phenomena (effects of light and nitrogen on growth, photoacclimation \[34\], \[48\] ...) and the radiative transfer in the photobioreactor (light attenuation due to the microalgae). Using these two approaches, we have developed a model which describes lipid production in a photobioreactor under light limitation. This model is used to predict lipid production in the perspective of large scale biofuel production. Simpler models have also been developed and have been used to provide optimization strategies: first, biomass production has been optimized in a constant light environment \[79\], yielding results emphasizing the importance of the optical depth of the reactor. In a second work, we focused on the optimal operating conditions for the biomass productivity under day/night cycles using Pontryagin’s maximum principle (assuming a periodic working mode) \[72\], \[73\].

Another model has been developed to represent growth of microalgae colimited by nitrogen and phosphorus \[69\]. It has been shown, from qualitative analysis of the model that uptake of nitrogen and phosphorus are non-symmetric.

6.2.1.2. Modelling the effect of light and temperature on microalgae

**Participants:** Olivier Bernard, Antoine Sciandra, Frédéric Grognard, Philipp Hartmann, Rafael Munoz Tamayo, Kerstin Ebert, Nina Moelants, Hubert Bonnefond.

The light distribution within a photobioreactor was estimated thanks to a multi photon Monte-Carlo simulation. From measurements of absorption and scattering properties, it was thus possible to extrapolate and validate the light distribution within a photobioreactor or a raceway.

The impact of the hydrodynamics on the light percept by a single cell was studied thanks to fluid dynamics simulations of raceway pond \[48\]. The light signals that a cell experiences at the Lagrangian scale, depending on the fluid velocity, were then estimated. A Droop-Han model was used to assess the impact of light fluctuation on photosynthesis \[48\].

Finally, the effect of temperature on microalgae has been represented by adapting the CTMI model developed for bacteria \[88\]. The proposed model \[59\], associated with a parameter identification procedure, was able to correctly represent the growth response to temperature for 12 different species \[48\].
6.2.2. CO$_2$ fixation by microalgae

Participants: Olivier Bernard, Antoine Sciandra, Philipp Hartmann, Nina Moelants.

We have run experiments to observe the response of a population of microalgal cells to various periodic light/dark or nitrate signals. The measurements show the synchronicity of the cells for some conditions. These experiments support the hypothesis that uptake of nitrogen stops during cell division [82]. On this basis, we have developed a structured model representing the development of microalgal cells through three main phases of their cell cycle: G1, G2 and M. The model is made of three interdependent Droop models [13]. The model was validated through extensive comparison with experimental results in both condition of periodic light forcing and nitrogen limitation. The model turns out to accurately reproduce the experimental observations [81]. The effect of cell synchronization on lipid content were experimentally studied [18] and included into microalgae growth models [36].

The effect of CO$_2$ partial pressure increase on photosynthesis and calcification of the calcareous microalgae *Emiliania huxleyi* have been experimentally observed. It results that an increase of the coccolith size together with a decrease in the calcification rate has been observed.

Three models accounting for the possible coupling between photosynthesis and calcification were included in an ocean model, including settling and predation by grazers, and a bloom of coccolithophorids was simulated [67], [68]. It was shown using Monte Carlo simulations that the uncertainty on the mechanisms driving calcification together with parametric uncertainties lead to uncertainties which are in the same range as the effect of an increase or the CO$_2$ partial pressure.

6.2.3. Design of ecologically friendly plant production systems

6.2.3.1. Controlling plant pests

Participants: Frédéric Grognard, Ludovic Mailleret, Mickaël Teixeira-Alves.

The influence of an alternative prey in biological control programs

We have developed a model based on the classical Leslie-Gower predator prey model, that allows for the choice that a predator might have for its diet. In a biological control framework, this choice might be between a pest that we want to eradicate and another prey that could be fed to the predator in order to help the biological control efficiency or between the pest and an alternative prey that is present in the field and might keep the predator from acting as a natural enemy of the pest. We put the problem in a time-partitioning framework: the predator has to split its time between the two prey. We then compared two time-partitioning strategies: one where the predator always spends a fixed proportion of its time on each prey and one where the predator always chooses the prey that is instantaneously most profitable (adaptive foraging). We then studied the effect of the presence of one prey on the other (indirect effect since it is mediated by the presence of the predator). We showed that, in the Leslie-Gower framework, one of the two prey always benefits from the presence of the other and that this effect is even stronger in the adaptive foraging framework, where the presence of the other prey is never detrimental to the one considered. That way, with very little assumptions, we showed the existence of apparent predation, commensalism and apparent mutualism, while most existing theoretical results tend to evidence apparent competition [51]. Such mechanisms may explain why generalist biological control agents are, in general, not as efficient as specialists in controlling crop pests.

The influence of plant dynamics on pest eradication

Pests-biocontrol agents models have been developed in order to build biological control strategies that can achieve pest eradication through augmentative biological control [85]. In the present work, we aim at introducing a plant compartment since its dynamics clearly have an influence on that of the pests and since the final objective of biological control is to maintain a sufficient plant yield. In a first step, we focused on plant-insect interactions and showed how the level and timing of the pest invasion could influence the final plant yield. Introducing pests control interventions and studying its timing, we showed how it eventually could have important effects on the growth pattern and the final biomass. As a reference, we consider the novel invasive pest Tuta absoluta on tomato plants [56].
This work is done in collaboration with Yves Dumont (Cirad).

6.2.3.2. Controlling plant pathogens

Participants: Frédéric Grognard, Ludovic Mailleret.

Sustainable management of plant resistance

The introduction of plants strains that are resistant to one pathogen often leads to the appearance of virulent pathogenic strains that are capable of infecting the resistant plants. The resistance strain then becomes useless. It is therefore necessary to develop ways of introducing such resistance into crop production without jeopardizing its future efficiency. We did so by choosing the proportion of resistant plants that are mixed with the non-resistant ones. In this work, we studied a vector borne pathogen in a seasonal environment, with healthy crop being planted at the beginning of each season and cropped at its end, the pathogen surviving in the environment during the ‘winter’. Two strategies have been proposed, one that aims at minimizing the cumulated damage over a 15 years horizon and one that aims at preventing the virulent strain outbreak. In the first case, both plant strains need to be mixed, but it results in the loss of the resistance at the end of the 15 year period; in the second case, the damage is higher and the maximal proportion of resistant plant is smaller, but the resistance is preserved [58], [16].

This work is done in collaboration with Frédéric Fabre (INRA Avignon).

Plant pathogen dynamics and cropping management practices

The coexistence of closely related plant parasites is widespread. Yet, understanding the ecological determinants of evolutionary divergence in plant parasites remains an issue. Niche differentiation through resource specialization has been widely researched, but it hardly explains the coexistence of parasites exploiting the same host plant. Agricultural systems are characterized by the cyclical presence and absence of the crop, due to cropping practices such as harvest and planting. We studied the influence that time-partitioning, i.e. the specialization of a parasite for the beginning or the end of crop presence, can have on co-existence. In modelling the epidemiology through a semi-discrete model we showed through an evolutionary invasion analysis that evolutionary divergence, and thus co-existence of different strains, of the parasite phenotype can occur [17], [44]. Also, in a similar context, we underlined why modelling seasonal plant epidemiology did not necessarily lead to competitive exclusion; indeed, generating a compact model by rigorously isolating the slow dynamics from a large detailed model of plant epidemiology, we found out the possibility of coexistence [49], [21]. Such a result contrasts with classical competitive exclusion principles found in compact models which rely on the arguable density independent nature of the pathogen infections occurring during the very beginning of the cropping seasons.

This work is done in collaboration with Frédéric Hamelin (Agrocampus Ouest).

6.2.4. Biological depollution

6.2.4.1. Coupling microalgae to anaerobic digestion

Participants: Olivier Bernard, Antoine Scandra, Jean-Philippe Steyer, Frédéric Grognard, Philipp Hartmann.

The coupling between a microalgal pond and an anaerobic digester is a promising alternative for sustainable energy production and wastewater treatment by transforming carbon dioxide into methane using light energy. The ANR Symbiose project is aiming at evaluating the potential of this process [90], [89].

In a first stage, we developed models for anaerobic digestion of microalgae. Two approaches were used: First, a dynamic model has been developed trying to keep a low level of complexity so that it can be mathematically tractable for optimisation [37], [32], [22]. Considering three main reactions, this model fits adequately the experimental data of an anaerobic digester fed with Chlorella vulgaris (data from INRA LBE). On the other hand, we have tested the ability of ADM1 [91] (a reference model which considers 19 biochemical reactions) to represent the same dataset. This model, after modification of the hydrolysis step [24], [38], [41] has then been used to evaluate process performances (methane yield, productivity...) and stability though numerical simulations.
In a second stage, a model describing the coupling between anaerobic digestion process and microalgae culture (including the feeding of the algae with anaerobic digestion effluents) has been developed. The model is based on the three steps model for anaerobic digestion, and on the photoacclimation model for microalgae [13]. The model also includes the modelling of heterotrophs in the microalgae pond.

6.2.4.2. Life Cycle Assessment of microalgae production

Participants: Olivier Bernard, Jean-Philippe Steyer.

This work is the result of a collaboration with Laurent Lardon and Arnaud Helias of INRA-LBE through the co-supervision of Pierre Collet’s PhD thesis.

An analysis of the potential environmental impacts of biodiesel production from microalgae has been carried out using the life cycle assessment (LCA) methodology [75]. This study has allowed to identify the obstacles and limitations which should receive specific research efforts to make this process environmentally sustainable. This study has been updated and the effects of technological improvements (leading to higher productivities) have been compared to the source of electricity. It turns out that the overall environmental balance can much more easily be improved when renewable electricity is produced on the plant [47], [46]. As a consequence, a new paradigm to transform solar energy (in the large) into transportation biofuel is proposed, including a simultaneous energy production stage.

A LCA has been carried out to assess the environmental impact of methane production by coupling microalgae and anaerobic digestion. The study highlights the limitation derived by the low biodegradability of the considered microalgae [15] which induces a large digester design and thus more energy to mix and heat it.

6.2.5. Models of ecosystems

6.2.5.1. Adaptive behaviour in seasonal consumer-resource dynamics

Participants: Frédéric Grognard, Ludovic Mailleret, Pierre Bernhard, Andrei Akhmetzhanov.

In this work we studied the evolution of a consumer-resource (or predator-prey) system with seasonal character of the dynamics. We specified two main parts of the process. First, we considered the system during one season with a fixed length: the prey lay eggs continuously and the predators lay eggs or hunt the preys (choose their behaviour) according to the solution of an optimal control problem [12]. Secondly, we studied the long-scale discrete dynamics over seasons. We investigated the qualitative behaviour of the dynamics with respect to the parameters of the problem and showed that, depending on the parameters of the model, extinction or co-existence of the predators and preys can be evidenced [12].

We then examined how (resident) predators adopting this behaviour would fare when faced with a small population of selfish mutants that would be identical to the resident but would have the freedom to choose a different behaviour. We studied the resulting optimal control problem where the mutants maximize their own number of offspring using the knowledge of the resident’s behaviour, and showed that, in most situations, mutants can take advantage of their low frequency and fare better than the residents. Over the course of a large number of seasons, the mutants replace the residents, only to find themselves applying the original resident behaviour [52]. We have then proposed a strategy for the predator in which it would prevent the invasion by the mutant instead of maximizing its number of offspring, which corresponds to the computation of evolutionarily stable life history strategies.

We have then considered that the resource itself could adapt its behaviour over time to limit the damage caused by the consumer, and maximize its own offspring. This problem requires the solution of a non zero-sum differential game, the consumer and the resource being the two players. We showed that the patterns of the strategies of the consumers and the resources are identical to the ones that can be obtained if the opposing player adopts a constant behaviour; the timing of the switchings varies however [42].

6.2.5.2. Including phytoplankton photoadaptation into biogeochemical models

Participant: Olivier Bernard.
The complexity of the marine ecosystem models and the representation of biological processes, such as photoadaptation, remain open questions. We compared several marine ecosystem models with increasing complexity in the phytoplankton physiology representation in order to assess the consequences of photoadaptation model complexity in biogeochemical model predictions. Three models of increasing complexity were considered, and the models were calibrated to reproduce ocean data acquired at the Bermuda Atlantic Time-series Study (BATS) from in situ JGOFS data. It turns out that the more complex model are trickier to calibrate and that intermediate complexity models, with an adapted calibration procedure, have a better prediction capability

6.2.5.3. Growth models of zooplankton

Participants: Jean-Luc Gouzé, Jonathan Rault, Eric Benoît.

The model built to describe a zooplankton community is some variant of the McKendrick-Von Foerster Equation. The model includes cannibalism within zooplankton and predation on phytoplankton. Dynamic mass budget theory is used in order to describe individual behaviour and allows mass conservation. Also we have added phytoplankton dynamics, and we use environmental data as an input for the model. The aim is to compare simulations with data provided by the Laboratoire d'Océanographie de Villefranche. Since the model incorporates lots of parameters, which are not always known in the literature, we have to use optimization techniques to find them. Further, equilibria of such models and their local stability is studied in using strongly continuous semigroup approach

We have also built a discrete size-structured model. Discrete models are less numerically demanding and so can be more easily incorporated into bigger models. Moreover the study of discrete models are often easier than that of continuous ones. We focus our study on the impact of cannibalism within zooplankton community and show that under some hypotheses, cannibalism can stabilize the equilibrium of the model

6.3. Software design

Participants: Olivier Bernard, Florian Guenn, Mélaïne Gautier.

Over the years, BIOCORE has been developing a software framework for bioprocess control and supervision called ODIN. This C++ application (working under Windows and Linux) enables researchers and industrials to easily develop and deploy advanced control algorithms through the use of a Scilab interpreter. It also contains a Scilab-based process simulator which can be harnessed for experimentation and training purposes. ODIN is primarily developed in the C++ programming language and uses CORBA to define component interfaces and provide component isolation. ODIN is a distributed platform, enabling remote monitoring of the controlled processes as well as remote data acquisition. Recently, a software development effort has been directed to the graphical user interface, a synoptic view component, new drivers for the experimental hardware and integration of the PlantML data exchange format. ODIN has been tested on four different processes and currently supervises the 66m² high rate pond at the LBE, INRA Narbonne.
6. New Results

6.1. High-throughtput sequence processing

- We published a book chapter on bioinformatics algorithms for GPU/manycore processors [16].
- Within the PhD of T. T. Tran, we proposed a new bit-parallel algorithm, extension of [39], as well as a new indexing structure adapted to GPUs [13].
- We proposed a new index structure specifically designed for reads produced by high throughput sequencers. It can deal both with variable or fixed length reads and can index reads in usually less memory than other classical solutions. This index has been implemented and is available online [10].
- We characterised the number of elements to be reordered when updating a full-text index such as a suffix array. We finally concluded that this number tends to be poly-logarithmic in the input length for DNA sequences [5].

6.2. Noncoding RNAs

- A. Saffarian defended her thesis on November 16. Within her thesis, we obtained two results:
  - We designed a new algorithm to produce all locally optimal secondary structures of an RNA Sequence. Locally optimal secondary structures are thermodynamically stable RNA structures that are maximal for inclusion: they cannot be extended without producing a conflict between base pairs in the secondary structure, or increasing the free energy. A journal article is in revision to Journal of Computational Biology.
  - We also proposed an algorithm to match a multi-structure of RNA against a sequence. A multi-structure gather several RNA structures, as real or putative structures on a same sequence, or as similar structures in a family of RNAs. A journal article was submitted to Algorithms for Molecular Biology.
- We published an algorithm for the comparison RNA secondary structures represented as nested arc-annotated sequences [7].
- The non-coding RNA collaborative annotation platform RNAspace is made available to the community, and published in RNA [3].

6.3. Genome rearrangements

- A. Thomas has started his PhD on September 2011. We already obtained two results:
  - We designed an algorithm for finding the minimal number of block interchanges required to transform a duplicated linear genome into a tandem duplicated linear genome. We provide a formula for the distance as well as a polynomial time algorithm for the sorting problem. This work has been accepted in Bioinformatics 2012.
  - We also introduced and study a new combinatorial problem, a biological phenomenon that apparently associates a significant proportion of segmental duplications in mammalians, drosophilas and bacterias to breakpoints in rearrangement events. called the Genome Dedoubling Problem. It consists in finding a minimum length rearrangement scenario required to transform a genome with duplicated segments into a non-duplicated genome such that duplications are caused by rearrangement breakpoints. We introduced new graph data structures to solve these problems. This work was presented at RECOMB’CG 2011 [11]
• We designed and applied new algorithms for inferring ultra-perfect evolution scenarios for Drosophila and mammals species [6].

• We implemented and applied the algorithm of [32] for the reconstruction of species tree from gene trees of Fungal and eukaryotes species [8].

• We proposed a new reconstruction of the architecture of the ancestral amniote genome based on the detection and assembly of ancestral genomic features conserved in extant species [9].

6.4. Non-ribosomal peptides

• A new database, called Doris (for Database Of non-Ribosomal Synthetases), has been created to extend the information we provide about non-ribosomal peptides to their producing enzymes, the non-ribosomal peptide synthetases. For the moment, a first version of the web interface has been developed by Louise Ott, an engineer from Lille1 University. More than 400 enzymes have been automatically extracted from general databases.

• A collaboration started with members of Orpailleur EPI to design a semi-automatic process to collect non-ribosomal synthetases (NRPSs). We already start adapting MODIM, a generic tool developed by Orpailleur EPI to collect and integrate data extracted from various web sources, to the specific needs of NRPSs.

• As mentioned in the software section, the source code has been reorganised by Laurie Tonon, a SED engineer, to use model view controller software architecture, implemented with Struts2.
5. New Results

5.1. New result 1

In [5], we explain the links between the solutions of the bidomain and monodomain models using some analytical arguments. The result is partially based on the theory of the bidomain operator explained in [2]. We can imagine several consequences to this general results, like improving the preconditionner proposed by C. Pierre [7] or derive some intermediate models.

5.2. New result 2

We computed some bidomain solutions for use by M. Pop and M. Sermesant in the STACOM’11 challenge from the MICCAI 2011 conference. They are the only bidomain simulations presented within this collaborative challenge. A collaborative paper will be published, see [8].
6. New Results

6.1. New methods for data assimilation

Since the beginning, CLIME has been focused on new techniques for data assimilation. Last year’s focus of the methodological development was on the use of non-Gaussian approaches for inverse modeling, and the construction of a multiscale data assimilation methodology. Several methodological papers have now been published on these topics. This year, the applications of these methodologies are put forward in the inverse modeling section, although new theoretical developments have been added to these approaches. In addition, new topics have been addressed, such as the ensemble Kalman filter with a theory that puts the EnKF on safer grounds, the use of 4D-Var for the estimation of fields of parameter in dispersion models, and the real-time data assimilation at urban scale.

6.1.1. Ensemble Kalman filtering without the intrinsic need for inflation

Participant: Marc Bocquet.

The main intrinsic source of error in the ensemble Kalman filter (EnKF) is sampling error. External sources of error, such as model error or deviations from Gaussianity, depend on the dynamical properties of the model. Sampling errors can lead to instability of the filter, which, as a consequence, often requires inflation and localization. The goal of this study is to derive an ensemble Kalman filter, which is less sensitive to sampling errors. A prior probability density function conditional on the forecast ensemble is derived using Bayesian principles. Even though this prior is built upon the assumption that the ensemble is Gaussian-distributed, it is different from the Gaussian probability density function defined by the empirical mean and the empirical error covariance matrix of the ensemble, which is implicitly used in traditional EnKFs. This new prior generates a new class of ensemble Kalman filters, called finite-size ensemble Kalman filter (EnKF-N). One deterministic variant, the finite-size ensemble transform Kalman filter (ETKF-N), is derived. It is tested on the Lorenz ’63 and Lorenz ’95 models. In this context, ETKF-N is shown to be stable without inflation for ensemble size greater than the model unstable subspace dimension, at the same numerical cost as the ensemble transform Kalman filter (ETKF). One variant of ETKF-N seems to systematically outperform the ETKF with optimally tuned inflation. However, it is shown that ETKF-N does not account for all sampling errors and necessitates localization like any EnKF, whenever the ensemble size is too small. In order to explore the need for inflation in this small ensemble size regime, a local version of the new class of filters is defined (LETKF-N) and tested on the Lorenz ’95 toy model. Whatever the size of the ensemble, the filter is stable. Its performance without inflation is slightly inferior to that of LETKF with optimally tuned inflation for small interval between updates, and superior to LETKF with optimally tuned inflation for large time interval between updates.

6.1.2. Parameter field estimation for atmospheric dispersion: Application to the Chernobyl accident using 4D-Var

Participant: Marc Bocquet.

Atmospheric chemistry and air quality numerical models are driven by uncertain forcing fields: emissions, boundary conditions, wind fields, vertical turbulent diffusivity, kinetic chemical rates, etc. Data assimilation can help to assess these parameters or fields of parameter. Because those parameters are often much more uncertain than the fields diagnosed in meteorology and oceanography, data assimilation is much more an inverse modeling challenge in this context. In this study, we experiment with these ideas by revisiting the Chernobyl accident dispersion event over Europe. We develop a fast four-dimensional variational scheme (4D-Var), which seems appropriate for the retrieval of large parameter fields from large observations sets, and for the retrieval of parameters that are non-linearly related to concentrations. The 4D-Var, and especially an approximate adjoint of the transport model, is tested and validated using several advection schemes that are
influential on the forward simulation as well as on the data assimilation results. Firstly, the inverse modeling system is applied to the assessment of the dry deposition parameters and of the wet deposition parameters. It is then applied to the retrieval of the emission field alone, to the joint optimization of removal process parameters and source parameters, and to the optimization of larger parameter fields, such as horizontal and vertical diffusivities, or dry deposition velocity field. The physical parameters used so far in the literature for the Chernobyl dispersion simulation are partly supported by the study. The crucial question of deciding whether such an inversion is merely a tuning of parameters, or a retrieval of physically meaningful quantities is discussed. Even though inversion of parameter fields may fail to determine physical values for the parameters, it achieves statistical adaptation that partially corrects for model errors, and, using the inverted parameter fields, leads to considerable improvement in the simulation scores.

![Graph](image)

**Figure 2.** Validation of the parameter estimation approach using 4D-Var for dry deposition and wet scavenging parameters in the context of the Chernobyl dispersion event. The 4D-Var optimization and the brute force screening value approach lead to the same result.

### 6.1.3. Real-time data assimilation

**Participants:** Vivien Mallet, Anne Tilloy, Fabien Brocheton [Numtech], David Poulet [Numtech], Cécile Honoré [Airparif], Édouard Debry [INERIS].

Based on Verdandi, Polyphemus and the “Urban Air Quality Analysis” software, real-time data assimilation was carried out at urban scale. The Best Linear Unbiased Estimator (BLUE) was computed for every hourly concentration map that the ADMS model computed. A posteriori tests were conducted over Clermont-Ferrand and Paris. We addressed the key issue of the covariance of the state error. The form of the error covariance between two points was determined based on the road network, considering the distance between points along the road and the distance of each point to the road. A few parameters (primarily two decorrelation lengths) were determined thanks to cross validation with several months of simulations and observations. The results showed strong improvements even at locations where no data was assimilated. The assimilation was carried out in the prototype “Votre Air” (http://votreair.airparif.fr/) for real-time air quality estimation over a part of Paris, which originally run in the context of the Futur-en-Seine festival and now runs operationally since June 2011.
At larger scale, the data assimilation library Verdandi was used to apply data assimilation (optimal interpolation) with the air quality model Chimere. This preliminary work will help INERIS to apply optimal interpolation in the operational platform Prev’air.

6.2. Inverse modeling

Many of this year’s studies have focused on inverse modeling, including the reconstruction of the Fukushima radionuclide source term. All were targeted to a particular application. However most of them include new methodological developments.

6.2.1. Estimation of errors in the inverse modeling of accidental release of atmospheric pollutant: Application to the reconstruction of the Fukushima Daiichi source term

Participants: Victor Winiarek, Marc Bocquet, Olivier Saunier [IRSN], Anne Mathieu [IRSN].

The aim of this research activity is the implementation of data assimilation methods, particularly inverse modeling methods, in the context of an accidental radiological release from a nuclear power plant and their application in the specific case of the Fukushima Daiichi accident. The particular methodological focus is the a posteriori estimation of the prior errors statistics. In the case of the Fukushima Daiichi accident, the number of available observations is small compared to the number of source parameters to retrieve and the reconstructed source is highly sensitive to the prior errors. That is the why they need to be well established and justified.

In this aim, three methods have been proposed: one method relies on an L-curve estimation technique, another one on the Desroziers’ iterative scheme and the last method, assumed to be the most robust, relies on the maximum likelihood principle (generalized to a non-Gaussian context).

These three methods have been applied to the reconstruction of cesium-137 and iodine-131 source terms from the Fukushima Daiichi accident. Because of the poor observability of the Fukushima Daiichi emissions, these methods provide lower-bounds for cesium-137 and iodine-131 reconstructed activities. Nevertheless, with the new method based on semi-Gaussian statistics for the background errors, these lower-bound estimates, $1.2 - 4.0 \times 10^{16}$ Bq for cesium-137, with an estimated standard deviation range of $15 - 20\%$, and $1.9 - 3.8 \times 10^{17}$ Bq for iodine-131, with an estimated standard deviation range of $5 - 10\%$, are of the same order of magnitude as those provided by the Japanese Nuclear and Industrial Safety Agency, and about 5 to 10 times less than the Chernobyl atmospheric releases.

6.2.2. Optimal Representation of Source-Sink Fluxes for Mesoscale Carbon Dioxide Inversion

Participants: Lin Wu, Marc Bocquet, Frédéric Chevallier [LSCE, CEA], Thomas Lauvaux [Department of Meteorology, Pennsylvania State University, USA], Peter Rayner [School of Earth Sciences, University of Melbourne, Australia], Ken Davis [Department of Meteorology, Pennsylvania State University, USA].

This study is an application of our previous theoretical developments on a consistent Bayesian multiscale formalism to optimally design control space (in which control variables are to be estimated). We construct the optimal adaptive representations of the surface fluxes for mesoscale carbon dioxide inversions. Such representations are taken from a large dictionary of adaptive multiscale grids. These optimal representations are obtained by maximizing the number of Degrees of Freedom for the Signal (DFS) that measures the information gain from observations to resolve the unknown fluxes. Consequently information from observations can be better propagated within the domain through these optimal representations.

The optimal representations are constructed using synthetic continuous hourly carbon dioxide concentration data in the context of the Ring 2 experiment in support of the North American Carbon Program Mid Continent Intensive (MCI). Compared with the regular grid at finest scale, optimal representations can have similar inversion performance with far fewer grid cells. For the Ring 2 network of eight towers, in most cases, the DFS value is relatively small compared to the number of observations $d$ (DFS/$d < 20\%$). In this multiscale setting, scale-dependent errors due to the aggregation of flux variables are identified and explicitly formulated for more reliable inversions. It is recommended that this aggregation error should be taken into account, especially when the correlations in the errors of a priori fluxes are physically unrealistic. The optimal multiscale grids allow to adaptively mitigate the aggregation error.
Figure 3. An optimal multiscale grid maximizing the DFS for a time period of 15 days. It characterizes the information flow from the eight tower sites to the whole Ring 2 domain. In this case, the error covariance matrix for the background carbon fluxes and the CO2 concentrations is assumed to be diagonal, and the total number of grid cells is set to 1024.

6.3. Monitoring network design

In this section, we report studies that are related to the evaluation of monitoring network and to new monitoring strategies. As opposed to last year’s report, they may not be strictly addressing optimal network design.

6.3.1. Eyjafjallajökull ash concentrations derived from both Lidar and modeling

Participants: Patrick Chazette [LSCE], Marc Bocquet, Philippe Royer [LSCE], Victor Winiarek, Jean-Christophe Raut [ATMOS], Philippe Labazuy [OPG/LMV], Mathieu Gouhier [OPG/LMV], Mélody Lardier [LEOSPHERE], Jean-Pierre Cariou [LEOSPHERE].

Following the eruption of the Icelandic volcano Eyjafjallajökull on April 14 2010, ground-based N2-Raman lidar (GBL) measurements were used to trace the temporal evolution of the ash plume from April 16 to April 20 2010, above the southwestern suburb of Paris. The nighttime overpass of the Cloud-Aerosol LIdar with Orthogonal Polarization onboard Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation satellite (CALIPSO/CALIOP), on April 17 2010, was an opportunity to complement GBL observations. The plume shape retrieved from GBL has been used to assess the size range of the particles size. The lidar-derived aerosol mass concentrations (PM) have been compared with model-derived PM concentrations held in the Eulerian model Polair3D transport model, driven by a source term inferred from the SEVIRI sensor on board the satellite Meteosat. The consistency between model and ground-based wind lidar and CALIOP observations has been checked. The spatial and temporal structures of the ash plume as estimated by each instrument and by the Polair3D simulations are in good agreement.

6.3.2. Potential of the International Monitoring System radionuclide network for inverse modeling

Participants: Mohammad Reza Koohkan, Marc Bocquet, Lin Wu, Monika Krysta [The Preparatory Commission for the Comprehensive Nuclear Test-Ban Treaty Organization].
The International Monitoring System (IMS) radionuclide network enforces the Comprehensive Nuclear-Test-Ban Treaty, which bans nuclear explosions. We have evaluated the potential of the IMS radionuclide network for inverse modeling of the source, whereas it is usually assessed by its detection capability. To do so, we have chosen the Degrees of Freedom for the Signal (DFS), a well-established criterion in remote sensing, in order to assess the performance of an inverse modeling system. Using a recent multiscale data assimilation technique, we have computed optimal adaptive grids of the source parameter space by maximizing the DFS. This optimization takes into account the monitoring network, the meteorology over one year (2009) and the relationship between the source parameters and the observations derived from the FLEXPART Lagrangian transport model. Areas of the domain where the grid-cells of the optimal adaptive grid are large emphasize zones where the retrieval is more uncertain, whereas areas where the grid-cells are smaller and denser stress regions where more source variables can be resolved.

The observability of the globe through inverse modeling is studied in strong, realistic and small model error cases. The strong error and realistic error cases yield heterogeneous adaptive grids, indicating that information does not propagate far from the monitoring stations, whereas in the small error case, the grid is much more homogeneous. In all cases, several specific continental regions remain poorly observed such as Africa as well as the tropics, because of the trade winds. The northern hemisphere is better observed through inverse modeling (more than 60% of the total DFS) mostly because it contains more IMS stations. This unbalance leads to a better performance of inverse modeling in the northern hemisphere winter. The methodology is also applied to the subnetwork composed of the stations of the IMS network that measure noble gases.

6.3.3. Optimal redistribution of the background ozone monitoring stations over France

Participants: Lin Wu, Marc Bocquet.

In this study, the BDQA background stations are partially redistributed over France under a set of design objectives, which are defined on a regular grid that covers France. Spatial interpolations are used to extrapolate simulated concentrations (of chemistry-transport models or assimilation results) to these grid nodes. Three types of criteria are considered: the geostatistical, geometrical, and physical ones. Simulated annealing is employed to select optimally the stations. Significant improvement with all the proposed criteria has been found for the optimally redistributed network against the original background BDQA network. For complex objectives, e.g. that addressing the heterogeneity of ozone field, the physical criteria are more appropriate.

6.4. Ensemble methods

Due to the great uncertainties that arise in air quality modeling, relying on a single model may be not sufficient. Therefore ensembles of simulations are now considered in a wide range of applications, from uncertainty estimation to operational forecast.

6.4.1. Uncertainty estimation based on multimodel ensembles and Monte Carlo simulations

Participants: Damien Garaud, Vivien Mallet, Raphaël Pérollat, Irène Korsakissok [IRSN], Denis Quélo [IRSN].

Air quality forecasts are limited by strong uncertainties especially in the input data and in the physical formulation of the models. There is a need to estimate these uncertainties for the evaluation of the forecasts, the production of probabilistic forecasts, and a more accurate estimation of the error covariance matrices required by data assimilation.

Because a large part of the uncertainty in the forecast originates from uncertainties in the model formulation (primarily the physical parameterizations), a multimodel ensemble seems to be the adequate tool for uncertainty estimation. Several 100-member ensembles were generated and calibrated (based on observations – see the example in Figure 4 ) over year 2007 in order to study the impact of EDF thermal plants on air quality. The ensemble simulations were carried out at European scale and at the scale of French regions so as to study the local impact of two EDF plants. This approach allowed us to estimate the uncertainty on the simulated impact of the plants. Based on the calibrated ensembles, we also computed probabilistic forecasts for threshold exceedences.
Specific work has been carried out to decompose the discrepancy between observations and model simulations in three errors: the modeling error, the representativeness error and the observational error. It was shown that the representativeness error might account for a third of the discrepancy.

After the Fukushima nuclear disaster, we worked on uncertainty estimation of the transport simulations at Japan scale. We carried out Monte Carlo simulations with perturbations on most input parameters. The first results showed that the variability of the ensemble simulations is great enough to reasonably sample the uncertainties after a calibration.

**Figure 4.** Uncertainty estimate (standard deviation) of two ensembles: a raw ensemble on the left, a calibrated ensemble on the right. The uncertainty estimation is significantly reduced after calibration.

### 6.4.2. Aggregation of meteorological forecasts

**Participants:** Anne Tilloy, Vivien Mallet, Fabien Brocheton [Numtech], David Poulet [Numtech].

Nowadays it is standard procedure to generate an ensemble of simulations for a meteorological forecast. Usually, meteorological centers produce a single forecast, out of the ensemble forecasts, computing the ensemble mean (where every model receives an equal weight). The generation of a single forecast with a weighted linear combination is called sequential aggregation. Each time new observations are available, the weights of the linear combination are updated and applied for the next forecasts. We applied the discounted ridge regression algorithm, which we previously introduced for sequential aggregation of air quality forecasts, to forecast wind and temperature at given observation stations. The ensemble was generated with forecasts at different range from two models. The aggregation proved to be efficient for one-day forecasts at least.

### 6.5. Image assimilation

Sequences of images, such as satellite acquisitions, display structures evolving in time. This information is recognized of major interest by forecasters (Meteorologists, oceanographers, etc) in order to improve the information provided by numerical models. However, these satellite images are mostly assimilated in geophysical models on a point-wise basis, discarding the space-time coherence visualized by the evolution of structures such as clouds. Assimilating in an optimal way image data is of major interest and this issue should be considered in two ways:
6.5.1. Divergence-free motion estimation

**Participants:** Dominique Béréziat [UPMC/LIP6], Isabelle Herlin, Nicolas Mercier, Sergiy Zhuk.

This research addresses the issue of divergence-free motion estimation on an image sequence, acquired over a given temporal window. Unlike most state-of-the-art technics, which constrain the divergence to be small thanks to Tikhonov regularisation terms, a method that imposes a null value of divergence of the estimated motion is defined.

Motion is characterized by its vorticity value and assumed to satisfy the Lagrangian constancy hypothesis. An image model is then defined: the state vector includes the vorticity, whose evolution equation is derived from that of motion, and a pseudo-image that is transported by motion. An image assimilation method, based on the 4D-Var technics, is defined and developed that estimates motion as a compromise between the evolution equations of vorticity and pseudo-image and the observed sequence of images: the pseudo-images have to be similar to the acquisitions.

As the evolution equations of vorticity and pseudo-image involve the motion value, the motion field has to be retrieved at each time step of the studied temporal window. An algebraic method, based on the projection of vorticity on a subspace of eigenvectors of the Laplace operator, is defined in order to allow Dirichlet boundary conditions for the vorticity field.

The divergence-free motion estimation method is tested and quantified on synthetic data. This shows that it computes a quasi-exact solution and outperforms the state-of-the-art methods that were applied on the same data.

The method is also applied on Sea Surface Temperature (SST) images acquired over Black Sea by NOAA-AVHRR sensors. The divergence-free assumption is roughly valid on these acquisitions, due to the small values of vertical velocity at the surface. Fig. 5 displays data and results. As no ground truth of motion is available, the method is quantified by the value of correlation between the pseudo-images and the real acquisitions. Again, the method provides the best result compared to other state-of-the-art algorithms.

6.5.2. Improvement of motion estimation by assessing errors on the dynamics

**Participants:** Dominique Béréziat [UPMC/LIP6], Isabelle Herlin, Nicolas Mercier.

Data assimilation technics are used to retrieve motion from image sequences. These methods require a model of the underlying dynamics, displayed by the evolution of image data. In order to quantify the approximation linked to the chosen dynamic model, we consider adding a model error term in the evolution equation of motion and design a weak formulation of 4D-Var data assimilation. The cost function to be minimized simultaneously depends on the initial motion field, at the beginning of the studied temporal window, and on the error value at each time step. The result allows to assess the model error and analyze its impact on motion estimation.

This error assessment method is evaluated and quantified on twin experiments, as no ground truth would be available for real image data. Fig. 6 shows four frames of a series of observations obtained by integrating the evolution model from an initial condition on image and velocity field (the ground truth \( w_{\text{ref}}(0) \) displayed on the left of Fig. 7). An error value is added at each time step on the motion value, when integrating the simulation model. This error is a constant bias.

We performed two data assimilation experiments. The first one considers the evolution model as perfect, with no error in the evolution equation. It is denoted PM (for Perfect Model). The second one, denoted IM (for Imperfect Model) involves an error in the motion evolution equation. In fig. 7 are displayed the motion fields retrieved by PM and IM at the beginning of the temporal window.

As it can be seen, IM computes a correct velocity field while PM completely fails.
Figure 5. Results on SST images (observation and motion result).

Figure 6. Observations Images.

Figure 7. Comparaison estimations with ground truth.
The results on this error assessment method are still preliminary. Perspectives are considered in order to correctly retrieve the error on dynamics by constraining its shape. An important application is, for instance, the detection of dynamics changes on long temporal sequences.

### 6.5.3. Nonlinear Observation Equation For Motion Estimation

**Participants:** Dominique Béréziat [UPMC/LIP6], Isabelle Herlin.

In the image processing literature, the optical flow equation is usually chosen to assess motion from an image sequence. However, it corresponds to an approximation that is no more valid in case of large displacements. We evaluate the improvements obtained when using the non linear transport equation of the image brightness by the velocity field. A 4D-Var data assimilation method is designed that simultaneously solves the evolution equation and the observation equation, in its non linear and linearized form. The comparison of results obtained with both observation equations is quantified on synthetic data and discussed on oceanographic Sea Surface Temperature (SST) images. We show that the non linear model outperforms the linear one, which underestimates the motion norm. Fig. 8 illustrates this on SST images (motion vectors are displayed by arrows).

![Figure 8. Results on SST images. Top: Non-linear observation equation. Bottom: linear.](image)

The aim of this research is to achieve a correct estimation of motion when the object displacement is greater than its size. In this case, coarse-to-fine incremental methods as well as the non linear data assimilation method fail to retrieve a correct value. The perspective is then to include, in the state vector, a variable describing the trajectory of pixels. The observation operator will then measure the effective displacement of pixels, according to their trajectories, and allow a better estimation of motion value.

### 6.5.4. Recovering missing data on images

**Participants:** Dominique Béréziat [UPMC/LIP6], Isabelle Herlin, Nicolas Mercier.
A data assimilation method was designed to recover missing data and reduce noise on satellite acquisition. The state vector includes motion and image fields. Its evolution equation is based on assumptions on the underlying dynamics displayed by the sequence of images and considers the passive transport of images by the velocity field. The observation equation compares the image component of the state vector and the real observations. Missing and noisy data regions are characterized by a gaussian observation error, whose covariance matrix $R$ has an approximately infinitesimal inverse. The noise recovering method computes a solution of the state vector that is a compromise between the evolution equation and the observation equation. The image component of the solution satisfies the assumptions on the dynamics and is close to the real acquisition according to the covariance matrix $R$. This image component provides the reconstruction of the noisy acquisitions.

The recovering method was applied on synthetically noised SST images in order to quantify the quality of the recovering (see Fig. 9).

![Figure 9. Recovering of noisy data.](image)

The method is a promising alternative to those such as space-time interpolation. In the experiments, the Lagrangian constancy of the state vector is used as evolution equation. The perspectives concern the use of more advanced dynamic equations, as for instance the shallow water equations that link the motion field to the thickness of the ocean surface layer, and improved modeling of illumination changes over the sequence, due to various order acquisition times.

### 6.5.5. Validation of velocity estimated with image assimilation

**Participants:** Isabelle Herlin, Etienne Huot, Gennady Korotaev [Marine Hydrophysical Institute, Ukraine], Evgeny Plotnikov [Marine Hydrophysical Institute, Ukraine].

This study is achieved in collaboration with the Marine Hydrophysical Institute (MHI) of Sevastopol. The aim is to estimate and further validate the estimation of Black Sea surface velocity from sequences of satellite images in order to allow an optimal assimilation of these pseudo-observations in 3D ocean circulation models. Several Image Models were designed that express the dynamics of velocity and the temporal evolution of image data. An image assimilation method was developed based on the 4D-Var formalism and estimates motion as a compromise between the Image Model, the image acquisitions and regularity heuristics on the velocity field. Two Image Models were qualitatively and quantitatively compared: the **Stationary Image Model** (SIM) based on the heuristics of stationary motion, which is valid at short temporal scale, and the **Shallow Water Image Model** (SWIM), based on the shallow-water equations.

The comparison between SIM and SWIM results confirms that SIM provides correct results only on short temporal windows, while SWIM allows to process longer image sequences.
The validation of motion estimation by image assimilation requires additional observation data, as no measure of motion is available from satellite sensors. Sea Level Anomaly, measured by satellite altimeters, is then compared to the thickness of the surface layer as estimated by the Shallow Water Image Model. This comparison shows a good adequacy of shape and values \[30\], \[32\]. As the velocity field is strongly related to this thickness value from the physical evolution laws, these results further validate the estimation of the velocity and the image assimilation approach.

6.5.6. Velocity estimation under the geostrophic equilibrium assumption

**Participants:** Isabelle Herlin, Etienne Huot.

The surface motion of the Black Sea approximately verifies the geostrophic equilibrium property. As the surface velocity can be directly derived from the surface layer thickness $h$, this allows to simplify the shallow-water equations and the dynamics is expressed by the evolution of $h$. The Geostrophic Shallow Water Image Model (GSWIM) is then designed based on the evolution of $h$ and the image data. A 4D-Var assimilation method was designed and developed in order to estimate $h$ from a sequence of satellite images. The motion field is then computed from the estimation of $h$.

This method was first tested and quantified on twin experiments with satellite data. Figure 10 simultaneously displays the result of the velocity estimation by GSWIM and the ground truth.

![Figure 10. Left: first image of the sequence, Center: motion estimated by GSWIM, Right: ground truth.](image)

6.5.7. Coupling models for motion estimation on long temporal image sequences

**Participants:** Karim Drifi, Isabelle Herlin.

This study concerns the estimation of motion fields from satellite images on long temporal sequences. The huge computational cost and memory required by data assimilation methods on the pixel grid makes impossible to use these techniques on long temporal intervals. For a given dynamic model (named full model on the pixel grid), the Galerkin projection on a subspace provides a reduced model that allows image assimilation at low cost. The definition of this reduced model however requires defining the optimal subspace of motion. A sliding windows method is thus designed:

- The long image sequence is split into small temporal windows that half overlap in time.
- Data assimilation in the full model is applied on the first window to retrieve the motion field.
- The estimate of motion field at the beginning of the second window makes it possible to define the subspace for motion and a reduced model is obtained by Galerkin projection.
- Data assimilation in the reduced model is applied for this second window.
- The process is then iterated for the next window until the end of the whole image sequence.
Figure 11 summarizes the described methodology.

![Diagram](image)

**Figure 11. Sliding windows method.**

Twin experiments were designed to quantify the results of this sliding windows method. Results on motion estimation are given in Figure 12 and compared with the ground truth. The NRMSE (in percentage) ranges from 1.1 to 4.0% from the first to the sixth window. On the first window, 3 hours are required to estimate the motion fields with the full model. For the next 5 windows, less than 1 minute is required to compute motion.

![Images](image)

**Figure 12. Estimated Motion (up) compared to the ground truth (down).**
6.6. Minimax filtering

Participants: Vivien Mallet, Sergiy Zhuk.

In air quality modeling, the model error is supposed to take into account the uncertainty on the meteorological fields (winds and vertical diffusivities), the segregation and mixing in emission areas that affect the effective kinetic rates of reactions, the boundary condition fields, all physical parameterizations (dry deposition, wet scavenging), etc. All the above sources of error have bounded energy and typically are not normally distributed or independent.

In order to take this into account in the data assimilation process, we applied the Minimax State Estimation (MSE) approach. It is well known that a bottle-neck of minimax estimation algorithms as well as of the family of Kalman-type filters is the dimension issue. To solve it, we applied a powerful version of the minimax filter developed for the so-called differential-algebraic equations. This filter works for any linear ordinary differential equation with time-dependent coefficients on any linear manifold, which can also change in time. Based on this novel approach, we derived a computationally tractable reduced version of the minimax filter. The derivation was made in a new and rigorous framework. In addition to the reduction, the new filter shows all the interesting properties inherited from the minimax setting, especially the description of the model and observational errors, which only need to have bounded energy. The later is important in the context of applications because the errors are always bounded. In contrast, most high-dimensional statistical filters are designed for unbounded random errors with special distribution function.

The algorithm, already implemented in the data assimilation library Verdandi, was further developed to compute a better reduction base.

The algorithm was in addition applied for ensemble sequential aggregation. The minimax filter computes weights for each model in the ensemble and a forecast is generated as the weighted linear combination of the ensemble members. In this case, the dimension is small so that no reduction is needed. The approach shows two noteworthy advantages: the observational errors can be taken into account and a dynamics can be given for the weights.

6.6.1. A posteriori minimax motion estimation

Participants: Sergiy Zhuk, Isabelle Herlin.

Data assimilation algorithms based on the 4D-Var formulation look for the so-called conditional mode estimate. The latter maximizes the conditional probability density function, provided the initial condition, model error and observation noise are realizations of independent Gaussian random variables. However this Gaussian assumption is often not satisfied for geophysical flows. Moreover, the estimation error of the conditional mode estimate is not a first-hand result of these methods. The issues above can be addressed by means of the Minimax State Estimation (MSE) approach. It allows to filter out any random (with bounded correlation operator) or deterministic (with bounded energy) noise and assess the worst-case estimation error.

The iterative MSE algorithm was developed for the problem of optical flow estimation from a sequence of 2D images. The main idea of the algorithm is to use the "bi-linear" structure of the Navier-Stokes equations and optical flow constraint in order to iteratively estimate the optical flow. The algorithm consists of the following parts:

1) we construct the pseudo-observations that is the estimate of the image brightness function $I(x, y, t)$ solving the optical flow constraint such that $I$ fits (in the sense of least-squares) the observed sequence of images; to do so we set the velocity field in the optical flow constraint to be the current minimax estimate of the velocity field $\mathbf{v}$, obtained at the previous iteration of the algorithm, and construct the minimax estimate $\hat{I}$ of the solution of the resulting linear advection equation using image sequence as discrete measurements;

2) we plug the estimate of the image gradient, obtained out of pseudo-observations $\hat{I}$ in 1), into the optical flow constraint and the current minimax estimate $\mathbf{v}$ of the velocity field into the non linear part of Navier-Stokes equations so that we end up with a system of linear PDEs, which represents an extended state equation: it contains a linear parabolic equation for the velocity field and linear advection equation for the image brightness.
function; we construct the minimax estimate of the extended state equation using the image sequence as
discrete measurements of the brightness function;
3) we use the minimax estimate of the velocity field obtained in 2) in order to start 1) again.
Currently numerical experiments are carried out in order to study the convergence rate of the algorithm.

6.7. Fire application

6.7.1. Model evaluation for fire propagation

Participants: Hajer Ayed, Vivien Mallet, Jean-Baptiste Fillipi [CNRS], Bahaa Nader [University of Corsica].

In the field of forest fires risk management, important challenges exist in terms of people and goods
preservation. Answering to strong needs from different actors (firefighters, foresters), researchers focus their
efforts to develop operational decision support system tools that may forecast wildfire behavior. This requires
the evaluation of models performance, but currently, simulation errors are not sufficiently qualified and
quantified. As the main objective is to realize a decision support system, it is required to establish robust
forecast evaluations. In the context of the ANR project IDEA, the evaluation of model simulations has been
started with a bibliographical review, the implementation of a series of forecast scores and the definition of a
series of ideal cases where some classical scores may fail (especially in taking into account the dynamics).

In addition, we consider that the proper evaluation of a model requires to apply it to a large number of fires –
instead of carrying out a fine tuning on just one fire. We implemented a software to simulate a large number
of fires (from the Prométhée database, http://www.promethee.com/ ) with the simulation model ForeFire
(CNRS/University of Corsica) and to evaluate the results with our error measures. One simulation requires
mainly the following data: the ignition point, the ground elevation, the vegetation cover and the wind field.
See the illustration of Fig. 13.

![Figure 13. Fire simulation (using ForeFire) in red elevated contour, and observation (from Prométhée) of the
burned area in filled red contour, for a 2003 fire near San-Giovanni-di-Moriani (Corsica).](image-url)
6. New Results

6.1. Spiking neurons

Participants: Hana Belmabrouk, Yann Boniface, Mohamed-Ghaïth Kaabi, Dominique Martinez, Horacio Rostro, Thierry Viéville, Thomas Voegtlin.

6.1.1. Mathematical modeling

- We demystify some aspects of coding with spike-timing, through a simple review of well-understood technical facts regarding spike coding, allowing to better understand to which extent computing and modeling with spiking neuron networks might be biologically plausible and computationally efficient. Considering a deterministic implementation of spiking neuron networks, we are able to propose results, formula and concrete numerical values, on several topics: (i) general time constraints, (ii) links between continuous signals and spike trains, (iii) spiking neuron networks parameter adjustment. This should prevent one from implementing mechanisms that would be meaningless relative to obvious time constraints, or from artificially introducing spikes when continuous calculations would be sufficient and more simple.

- We propose a generalization of the existing maximum entropy models used for spike train statistics analysis, bringing a simple method to estimate statistics and generalizing existing approaches based on Ising model or one step Markov chains to arbitrary parametric potentials. Our method enables one to take into account memory effects in dynamics. It provides directly the “free-energy” density and the Kullback-Leibler divergence between the empirical statistics and the statistical model. Furthermore, it allows the comparison of different statistical models and offers a control of finite-size sampling effects, inherent to empirical statistics, by using large deviations results. This work is submitted for publication.

- Following some theoretical work about back-engineering from spike recordings, we study the possibility to design an artificial vision system based on spiking neurons, for which neural connections and synaptic weights are directly derived from recordings of spiking activities in the human visual system through a back-engineering approach. A specific simple spiking model has been defined that mathematically enables this back-engineering process from biological data. From a hardware point of view, this model results in an efficient implementation on FPGAs.

6.1.2. Biophysical modeling

Our understanding of the computations that take place in the human brain is limited by the extreme complexity of the cortex, and by the difficulty of experimentally recording neural activities, for practical and ethical reasons. The Human Genome Project was preceded by the sequencing of smaller but complete genomes. Similarly, it is likely that future breakthroughs in neuroscience will result from the study of smaller but complete nervous systems, such as the insect brain or the rat olfactory bulb. These relatively small nervous systems exhibit general properties that are also present in humans, such as neural synchronization and network oscillations. Our goal is therefore to understand the role of these phenomena by combining biophysical modelling and experimental recordings, before we can apply this knowledge to humans. In the last year, we obtained the following results:

- We have explored the role of subthreshold membrane potential oscillations in stabilizing the oscillation frequency in a model of the olfactory bulb [9].

- We have developed several biophysical models of the insect olfactory system to explain the transformation from first-order [10] to second-order neurons [6], [16]. We show in particular how cellular and network mechanisms contribute to coding efficiency.
6.2. Dynamic Neural Fields

Participants: Lucian Aleçu, Frédéric Alexandre, Yann Boniface, Laurent Bougrain, Mauricio Cerda, Georgios Detorakis, Hervé Frezza-Buet, Bernard Girau, Axel Hutt, Mathieu Lefort, Jean-Charles Quinton, Nicolas Rougier, Wahiba Taouali, Thierry Viéville, Thomas Voegtlin.

The work reported this year represents both extensions of previous works and new results linked to the notion of neural population, considered at (i) a formal level (theoretical studies of neural fields), (ii) a numerical level (study of functioning and learning rules) and (iii) a more embodied one (implementations of specific functions).

6.2.1. Formal Level

- study of the differences between synchronous and asynchronous (without a central clock) evaluation: The hallmark of most artificial neural networks is their supposed intrinsic parallelism where each unit is evaluated concurrently to other units in a distributed way. However, if one gives a closer look under the hood, one can soon realize that such a parallelism is an illusion since most implementations use what is referred to as synchronous evaluation, or using a central clock. Here we propose to consider different evaluation methods (namely asynchronous and event based evaluation methods) and study their properties in some restricted but illustrative cases. This work is also in preparation for publication.

- taking into account transmission speed between units in a neural field: Neurons in populations are connected to each other by axonal branches sending electric pulses. The pulse propagation with finite speed delays the neuron interactions. The developed numerical algorithm illustrates how to simulate neural fields in two spatial dimensions involving finite axonal transmission speeds. The algorithm is derived analytically shows how to implement a Fast Fourier Transform in the computation scheme.

- study of the bridge between an ensemble of spiking neurons and the population firing rate to extend neural fields by shunting inhibition effects. Shunting inhibition is an important effect in real neural systems, e.g. in the context of general anaesthesia. We re-derive the population firing rate well-known in neural fields from the single neuron firing statistics. This derivation assumes McCulloch-Pitts neurons with a trivial f-I curve. Then we exchange the McCulloch-Pitts neurons by more realistic type I-neurons with a non-trivial f-I curve and gain a different, more realistic population firing rate. This formulation allows to consider some shunting inhibition effects [44].

6.2.2. Numerical Level

6.2.2.1. Numerical studies of DNF and related mechanisms

At the numerical level, specific developments were carried out to assess our software platform, to master functioning rules and to study the performances of new learning rules:

- The problem of adjusting the parameters of a mesoscopic event and valued neural field with delayed connections is addressed here at the programmatic level. An effective computational framework, with the implementation of a general algorithm is developed allowing us to effectively design non-trivial input/output transformations of events and values, using a class of biologically plausible distributed functional models. This work is in preparation for publication.

- In order to clarify the notion of distributed computing, general concepts and definitions in the framework of artificial neural networks have been reviewed, within the scope of dynamic field theory, proposing an unequivocal definition of asynchronous computation. An innovative way to perform such asynchronous computation has been proposed, following theoretical developments in process formalization. Several consequences on both the trajectories and the stability of the whole system have been drawn, including a few practically usable methods and quantitative bounds that can guarantee most of the mesoscopic properties of the system [15].

- Novel numerically efficient algorithm to compute spatio-temporal activity in two-dimensional neural fields involving finite transmission speed.
• Study of the possibility to obtain properties of self-organization with dynamic neural fields and proposition of a new learning rule for self-organization [1], [5].
• We designed a variation of the self-organising map algorithm [14] where the original time-dependent (learning rate and neighbourhood) learning function has been replaced by a time-invariant one. This allows for on-line and continuous learning on both static and dynamic data distributions. One of the property of the newly proposed algorithm is that it does not fit the magnification law and the achieved vector density is not directly proportional to the density of the distribution as found in most vector quantisation algorithms. From a biological point of view, this algorithm sheds light on cortical plasticity seen as a dynamic and tight coupling between the environment and the model.
• Adaptation of the BCM rule to multi-modality by adapting the dynamics of the threshold by the use of a feedback signal generated by a neural field map [22], [39], [45]
• Following [25], we are now studying a computational model of the primary somato-sensory cortex based on the neural field theory where cortical representations develop through the modification of thalamocortical synapses (from thalamus to layer 4), while cortico-cortical synapses (layer 2/3 and 4) provide a distributed competitive mechanism between cortical pyramidal neurons of layer 2/3. Preliminary results explains both the initial development and the self-organization of cortical representations in the primary sensory cortex as well as the dynamic reorganization following a lesion or a sensory deprivation. In this context, the so-called critical period during childhood would correspond to the development and learning of the intra-cortical competitive mechanism that is critical for cortex plasticity.

6.2.2.2. Gaussian mixture based approximation of neural maps
We have studied the advantages of our new implementation of the Continuous Neural Field Theory (CNFT) using a Gaussian mixture based model of the neural field activity, when using high dimensional inputs [40]. It exploits the rapid convergence of the activity to a reduced set of localized bubbles when competition occurs. These bubbles of activity can be accurately approximated by Gaussian distributions, that are directly computed in any n-dimensional space, instead of projecting high dimensional inputs onto 2D maps (which generally leads to topological distortions). This implementation is thus used to evaluate the possibilities of sensorimotor or multimodal associations without prior self-organization on 2D cortical maps, and could be directly interfaced with high dimensional artificial systems.

6.2.3. Embodied Level
6.2.3.1. Motion detection
We develop bio-inspired neural architectures to extract and segment the direction and speed components of the optical flow from sequences of images. Following this line, we have recently built additional models to code and distinguish different visual sequences. The structure of these models takes inspiration from the course of visual movement processing in the human brain, such as in area MT (middle temporal) that detects patterns of movement, or area FBA where neurons have been found to be sensitive to single spatio-temporal patterns. This work has been recently extended to complex movements: to fight, to wave, to clap, using real-world video databases [2], as well as using speech-driven visual animations of faces [28].

6.2.3.2. Modeling the superior colliculus by mean of a neural field.
In the context of the ANR MAPS project (cf. § 7.2), we have been studying the superior colliculus in tight collaboration with Laurent Goffart from the Institut de Neurosciences Cognitives de la Méditerranée. Considering the cortical magnification induced by the non homogeneous distribution of retina rods and cones on the retina surface, we modeled the superior colliculus using a dynamic neural field that may explain the stereotyped nature of colliculus activity. This year, we have extended this approach to wider contexts:
• Using Neural Fields to model the Superior Colliculus in a task of saccade generation
• Arrangement of several neural fields to model several cortical areas engaged in visual attention
6.2.3.3. Modeling of neural activity during anaesthesia.

Anaesthesia plays an important role in medical surgery though its neural mechanism is still poorly understood. Besides several different molecular and behavioral phenomena, the administration of anaesthetic agents affects the power spectrum of electro-encephalographic activity (EEG) in a characteristic way. The theoretical study aims to model the power spectrum changes in EEG subject to the concentration of the specific anaesthetic agent propofol. The work developed a neural model [38] involving two neuron types and synapse types while taking into account the synaptic effect of propofol. The mathematical derivation of the power spectrum allows for the investigation of suitable physiological parameters which reproduce the experimental effect of propofol. Several mathematical conditions on physiological parameters have been derived and the EEG-power spectrum during the administration of different concentration levels of propofol has been modeled successfully.

6.3. Higher level functions


Our activities concerned information analysis and interpretation and the design of numerical distributed and adaptive algorithms in interaction with biology and medical science. To better understand cortical signals, we choose a top-down approach for which data analysis techniques extract properties of underlying neural activity. To this end several unsupervised methods and supervised methods are investigated and integrated to extract features in measured brain signals. More specifically, we worked on Brain Computer Interfaces (BCI).

6.3.1. Detection of partial amplitude synchronization in multivariate data

To gain information on the interactions between neural structures, several electrodes may be implanted in cortical areas to measure Local Field Potentials. The developed method aims to extract time windows in which a subset of measured time series exhibit an amplitude synchronization in certain frequency bands [12].

6.3.2. Brain-Computer Interface based on motor imagery to control a robotic arm in 3D

The interface we develop aims to control in 3D a Jaco robotic arm by Kinova, using the Graz Motor Imagery detection paradigm for two or three motor actions in an online situation. The interface is part of the OpenViBE software. The user can switch in different modes to control a specific part of the robotic device (arm, wrist, fingers). We plan to use five different motor imageries: right hand, left hand, foot, rest and both hands. The actions are not available all together for a specific control. The interface is already done. More experiments will be done to adjust the classifier.

6.3.3. Reinforcement learning to better control a robotic arm

The approach we proposed in Cobras is innovative. Many studies attempts improve the recognition rate of a BCI order with new methods for treatment of signal. These studies are placed upstream of the BCI to facilitate the retrieval of information in the signal. However, the signal to noise ratio is so low that the improvements are limited. Rather than improving signal processing upstream, we wanted to improve the recognition rate by adding information in the controlled system. Thus, we placed downstream and added, as inputs of our control system, mechanical data concerning robotic arm. Initially, we validated the possibility of finding -using an inverse algorithm of reinforcement learning- the policy of the expert from a set of trajectories followed in a maze. We defined then a scenario to achieve different trajectories with the robotic arm to reach several buttons. In a third step, we used this algorithm on a maze-type problem but for which we have completed the state vector with the classifier outputs. This study is ongoing.
6.3.4. Mutual influence of firing rates of corticomotoneuronal cells for learning a precision grip task

As a part of a Brain-Machine Interface, we define a model for learning and forecasting muscular activity, given sparse cortical activity in the form of action potential signals (spike trains). We have a collection of experiments in which a trained monkey performs a precision grip. More precisely, its neuronal activity is partially recorded from corticomotoneuron cells of the hand area (area 4) as the monkey clasps two levers between its index finger and thumb. The underlying model parameters are interpreted with respect to the physiological aspects, though the model itself is not bio-physical. The method used is based on a system of first degree linear equations involving the firing rate of the recorded neurons, two sets of thresholds associated to them, and the variation of the global neuronal activity. We build a module to translate the data in the form of spikes trains into the event structure of OpenViBE triggers which is more appropriate than signals. The enslavement of the clamp according to the order generated by OpenViBE was also done. These solutions can demonstrate the capabilities of our algorithms for decoding cortical signals in the task of handling.

6.3.5. Hysteresis thresholding for Wavelet denoising applied to P300 single-trial detection

Template-based analysis techniques are good candidates to robustly detect transient temporal graphic elements (e.g. event-related potential, k-complex, sleep spindles, vertex waves, spikes) in noisy and multi-sources electroencephalographic signals. More specifically, we studied the significant impact on a large dataset of wavelet denoisings to detect evoked potentials in a single-trial P300 speller. We applied the classical thresholds selection rules algorithms and compared them with the hysteresis algorithm by R. Ranta which combine the classical thresholds to detect blocks of significant wavelets coefficients based on the graph structure of the wavelet decomposition.

6.4. Embodied and embedded systems

Participants: Yann Boniface, Hervé Frezza-Buet, Bernard Girau, Mathieu Lefort, Dominique Martinez, Jean-Charles Quinton, Nicolas Rougier.

6.4.1. InterCell

Our research in the field of dedicated architectures and connectionist parallelism mostly focuses on embedded systems (cf. § 3.5 ). Nevertheless we are also involved in a project that considers coarse-grain parallel machines as implementation devices. The core idea of this InterCell project (cf. http://intercell.metz.supelec.fr ) is to map fine grain computation (cells) to the actual structure of PC clusters. The latter rather fit coarse grain processing, using relatively few packed communication, which a priori contradicts neural computing. Another fundamental feature of the InterCell project is to promote interaction between the parallel process and the external world. Both features, cellular computing and interaction, allow to consider the use of neural architectures on the cluster on-line, for the control of situated systems, as robots.

6.4.2. Embodied/embedded olfactory systems

6.4.2.1. How can animals successfully locate odour sources?

Our goal is to investigate this question. Two different classes of strategies are possible for olfactory searches: those based on a spatial map, e.g. Infotaxis, and those where the casting-and-zigzagging behaviour observed in insects is purely reactive. We have implemented Infotaxis in a robot and shown that it produces trajectories that feature zigzagging and casting behaviours similar to those of moths. This result however should not be interpreted as evidence that the corresponding moth behaviour is driven by Infotaxis. Whether or not moths use infotactic or reactive strategies is still unclear. To compare both strategies, we have developed a cyborg using the antennae of a tethered moth as sensors (no artificial sensor for pheromone molecules is presently known). Experiments are in progress to compare the trajectories of the cyborg controlled by infotactic and reactive search strategies to those obtained with the same cyborg but driven by the moth’s brain.
6.4.2.2. How can technology emulate biological olfactory processing?

Glomerular microcircuits in the first stage of the olfactory pathway reformat odor representation. First, many ORNs expressing the same receptor protein, yet presenting heterogeneous dose-response properties, converge onto each glomerulus [10]. Second, onset latency of glomerular activation is believed to play a role in encoding odor quality and quantity in the context of fast information processing [6]. Taking inspiration from biology, we designed a simple yet robust glomerular latency coding scheme for processing gas sensor data [7]. The proposed bio-inspired approach was evaluated using an Sn02 sensor array. Glomerular convergence was achieved by noting the possible analogy between receptor protein expressed in ORNs and metal catalyst used across the fabricated gas sensor array. Ion implantation was another technique used to account both for sensor heterogeneity and enhanced sensitivity. The response of the gas sensor array was mapped into glomerular latency patterns, whose rank order is concentration-invariant.

6.4.3. Hardware implementations of neural models

In the field of dedicated embeddable neural implementations, we use our expertise in both neural networks and FPGAs so as to propose efficient implementations of applied neural networks on FPGAs, as well as to define hardware-friendly neural models.

- Following our results on the design of spiking models back-engineered from spike recordings, recent works have focused on the analysis of the influence of precision onto asymptotic dynamics of FPGA-embedded integrate-and-fire neural models [13].

- We design hardware-friendly adaptations of dynamic neural fields that use spiking neurons. In this field, we have derived a highly simplified version of such spiking neural fields, and we have experimentally shown that the main properties of standard neural fields are maintained in the context of visual attention [29].

- We currently intend to minimize the topological constraints of FPGA-embedded spiking neural fields using reduced neighborhoods but randomly propagating spikes. A preliminary result has been obtained so as to implement massively distributed pseudo-random number generators based on cellular automata that use minimal areas [21].

6.4.4. Towards brain-inspired hardware

Our activities on dedicated architectures have strongly evolved in the last years. We now focus on the definition of brain-inspired hardware-adapted frameworks of neural computation. Our current works aim at defining hardware-compatible protocols to assemble various perception-action modalities that are implemented and associated by different bio-inspired neural maps.

6.4.4.1. Anticipatory mechanisms in neural fields

We have defined first models of neural fields that include anticipatory mechanisms through the integration of spatiotemporal representations into the lateral interactions of a dynamic neural field [23]. This work targets increased robustness and goal-oriented action selection within sensori-motor systems.

6.4.4.2. Multimodal learning through joint dynamic neural fields

This work relates to the development of a coherent multimodal learning for a system with multiple sensory inputs.

- We have modified the BCM synaptic rule, a local learning rule, to obtain the self organization of our neuronal inputs maps and we use a CNFT based competition to drive the BCM rule. In practice, we introduce a feedback modulation of the learning rule, representing multimodal constraints of the environment [39].

- We have introduced an unlearning term in the BCM equation to solve the problem of the different temporalities between the raise of the activity within modal maps and the multimodal learning of the organization of the maps [22].
6. New Results

6.1. Modelling and Identification

6.1.1. Multi-Dimensional Wrist Musculoskeletal Modeling for Tremor Simulation

Participants: Peng Yao, Mitsuhiro Hayashibe, Dingguo Zhang (Shanghai Jiao Tong Univ.).

In this work, we established multi-dimensional wrist musculoskeletal model to be used for suppressing the wrist joint’s tremor by functional electrical stimulation. Often the wrist model for FES control is based on 1DOF biomechanical model for the simplicity and convenience to develop the controller. However, wrist motion is generated by complex interactions of multiple muscles spanning the wrist joint. Here, we have tried to have 3DOF wrist model considering main muscles involved in flex-extension, radial-ulnar deviation, and pron-supination as in the left of Fig. 1 (opensim model). Inertia of 3DOF joint model was obtained based on the work of de Leva which modified Zatsiorsky parameters. Joint dynamics was formed by inertia, gravitational torque, and passive visco-elastisity. As for the mapping between torque and muscles, the moment arm matrix \((3 \times 4)\) was obtained from opensim software, thus anatomical information could be considered along with muscle parameters such as muscle-length, isometric maximal forces. For this first trial, 4 muscles were considered based on Hill-type model. In order to confirm the generated motion by muscle activation, we activated the ECRB (extensor carpi radialis brevis) and FCU (flexor carpi ulnaris) in antiphase, the corresponding 3DOF wrist angles were obtained as in Fig. 1. Qualitatively, we know that ECRB can generate negative flexion angle, negative deviation angle and positive pronation when it is activated; while FCU can produce positive flexion angle, positive deviation angle and positive pronation angle. When both ECRB and FCU are activated in order, the reasonable wrist angles could be generated. We will work on the development of tremor controller in the future work.

![Figure 1. 3DOF Wrist Musculoskeletal Modeling. Anatomical layout of principle wrist muscles, moment arm matrix was obtained from opensim (left). 3DOF wrist angles were generated by tremor like activation in ECRB and FCU (right).](image)

6.1.2. Multi-functional EMG classification for dynamic EMG-motion modeling

Participants: Lizhi Pan, Mitsuhiro Hayashibe, Dingguo Zhang, Xiangyang Zhu (Shanghai Jiao Tong Univ.).
EMG signal is widely used to control a limb protheses such as exoskeleton. It would be also useful to control prosthetic hand for upper-extremity amputees. Especially for upper limb control, hybrid control both for position and torque is required for dynamic motions. In this work, we aimed at establishing the EMG-joint angle model for dynamic motions, to effectively decode EMG signals to reproduce the corresponding motion in different velocity. In the experiment, the subject performed the wrist flexion-extension with different speeds along with EMG measurements on flexor and extensor muscles. ARX model, of which the parameters are adaptively identified by extended Kalman filter, was applied to represent 1 DoF (degree of freedom) EMG-joint angle model. The result shows that the EMG-angle model could produce good angle tracking as shown in Fig. 2. We have observed that three principal parameters a1, a2, a3 for autoregressive term are almost constant at the same angular speed and different angular speeds correspond to different model parameters. With this regular pattern, SR (Switch Regime) model would be possible to be used to switch the model to reproduce the motion in different conditions. In the further work, we work on creating a SR EMG-angle model using learning and classification technique to dynamically decode the EMG signals into the corresponding motions only based on EMG without angle information.

Figure 2. Dynamic EMG-motion modeling and identification. Subject performed the wrist flexion-extension with different speeds along with EMG measurements on flexor and extensor muscles. Adaptive identification was implemented with Kalman filtering. Model parameters and speed will be used for classification.

6.1.3. Real-time Volumetric Skeletal Muscle Deformation

Participants: Yacine Berranen, Mitsuhiro Hayashibe, Benjamin Gilles.

In this work, we explore skeletal muscle volumetric deformation. The current available simulation for musculoskeletal model is basically using wire-type muscle model which considers only principal longitudinal path of the muscle-tendon units. If we aim at the simulation of the interaction between muscles and objects like orthosis, exoskeleton keeping classical biomechanical property, wire-type modeling is not sufficient. In addition, muscle modeling in volumetric way gives another advantage to reflect microscopic muscle fiber direction and function. We have tried to implement real-time volumetric skeletal muscle deformation as in Fig. 3 using the INRIA SOFA environment. The idea is making more realistic musculoskeletal simulation from the current approximation of muscle model as wire element to physically and functionally detailed simulation as volumetric element.

6.1.4. Muscle Strength and Mass Distribution Identification in Musculoskeletal Modeling

Participants: Mitsuhiro Hayashibe, Gentiane Venture (Tokyo Univ. of Agriculture and Technology), Ko Ayusawa (Univ. of Tokyo), Yoshihiko Nakamura (Univ. of Tokyo).
In current biomechanics approach, the assumptions are commonly used in body-segment parameters and muscle strength parameters due to the difficulty in accessing those subject-specific values. Especially in the rehabilitation and sports science where each subject can easily have quite different anthropometry and muscle condition due to disease, age or training history, it would be important to identify those parameters to take benefits correctly from the recent advances in computational musculoskeletal modeling. In this paper, Mass Distribution Identification to improve the joint torque estimation and Muscle Strength Identification to improve the muscle force estimation were performed combined with previously proposed methods in muscle tension optimization. This first result highlights that the reliable muscle force estimation could be extracted after these identifications. Fig. 4 shows the estimated muscle forces of Rectus Femoris, Vastus Lateralis and Vastus Medialis with different speeds (first two series are normal speed, second two are slow and last two are fast). The corresponding visualizations of estimated muscle tensions at the indicated time instant are depicted in the bottom. The proposed framework toward subject-specific musculoskeletal modeling would contribute to a patient-oriented computational rehabilitation [22].

6.1.5. Joint Angle Estimation with Inertial Sensors Calibrated by Kinect

Participants: Mitsuhiro Hayashibe, Antonio Padilha Lanari Bo (Univ. of Brasilia), Philippe Poignet.

In this work, we explore the combined use of inertial sensors and the Kinect for applications on rehabilitation robotics and assistive devices. In view of the deficiencies of each individual system, a new method based on Kalman Filtering was developed in order to perform online calibration of sensor errors automatically whenever measurements from Kinect are available. The method was evaluated on experiments involving healthy subjects performing multiple DOF tasks. Accelerometers and gyrometers are used to estimate joint angle, while the Kinect is used for initializing the inertial system and for enabling 3D visualization of the performed task as in Fig. 5 [32].

6.1.6. Investigation of fibre size stimulation selectivity using earthworm model

Participants: Pawel Maciejasz, Christine Azevedo Coste, David Andreu, David Guiraud.

Fibre type and diameter selective stimulation may allow to restore various motor and sensory functions of human body that have been lost due to disease or injury. For example in people unable to voluntary empty the bladder, selective stimulation of small fibres within the ventral branch of the sacral nerve roots (S2-S4/5) would induce detrusor contraction and those cause bladder emptying closest to normal physiology. Currently, it is not possible to perform it in such a way, because stimulation of sacral nerve roots activates also bigger fibres innervating the urethral sphincter, which closest the outlet of the bladder.

Already many stimulation techniques have been proposed for fibre type and diameter selective stimulation. They were verified performing computer simulations and in some cases also by in vivo experiments on mammalian models. However, results of computer simulations still need to be confirmed by in vivo experiments, whereas experiments on mammalian models, due to high number of fibres within stimulated nerve, can be very
Figure 4. Estimated muscle forces of Rectus Femoris, Vastus Lateralis and Vastus Medialis using the identified model in squat motion with 4kg load with different speeds (first two series are normal speed, second two are slow and last two are fast) (up). The corresponding visualization of estimated muscle tensions at the indicated time instant (bottom).

Figure 5. Angles estimated using inertial sensors in a sit-to-stand task.
complex to perform and obtained results difficult to interpret. As a result, it is still unclear which stimulation parameters may allow for selective stimulation of only particular group of fibers. Therefore we propose the earthworm (Lumbricus terrestris) as a model for selective stimulation. The earthworm has three giant nerve fibres, with two distinctly different conduction velocities and diameters. Therefore it is very easy to distinguish between fibres that are firing at the moment. As a consequence the selectivity of stimulation may be immediately verified without application of sophisticated signal processing and averaging techniques.

We have investigated influence of various pulse amplitudes and durations on the selectivity of stimulation. Using a simple experimental set-up [29] shown in the fig. 6 A, we were able to achieve selective activation of small (fig. 6 C) and big (fig. 6 D) fibers, as well as concurrent activation of both fibers (fig. 6 B) [28]. For that purpose we have used so called "anodal block" technique.

Based on the results of the above experiments, the recommendations of the optimal parameters for selective stimulation of nerve fibres will be prepared. Afterwards we are going to verified these recommendations in mammalian models (rats).

Figure 6. A. The schematic representation of the stimulation ("-" - cathode, "+" - anode) and recording (R1, R2 and R3) sites during experiments on earthworms; B, C and D - examples of nerve responses recorded for various pulse amplitudes. It may be observed that although all fibres are activated under cathode (R1 recording site, first plot in each line), on the anode side (R2 and R3 recording sites, second and third plot in each line respectively) activation of all (B), only small (C) and only big (D) fibres could be achieved.

6.1.7. Neural network based identification for time-variant dynamics

Participants: Zhan Li, Mitsuhiro Hayashibe, David Guiraud.

Due to high nonlinearity and time variance in muscle dynamics under FES, the identification of muscle model is complex task. The time-variant of muscle response may come from muscle fatigue, but also the electrode attachment condition. Along with such long-term time-variance, short term time-variance may be created by reflex effect. In addition, the characteristics of such variance even may change in time in an unpredicted way.
Reinforcement learning framework may be applied to bring the robustness in adaptive identification. Current work is focused on the usage of discrete-time recurrent neural network for model identification.

6.2. Function control and synthesis

6.2.1. Correction of drop-foot in post-stroke hemiplegic patients

Participants: Christine Azevedo Coste, Roger Pissard-Gibollet (SED INRIA), Fabien Jammes (INRIA RA), Jérôme Froger (Rehab. Centre, Grau du Roi, CHU Nîmes).

Hemiplegia is a condition where one side of the body is paretic or paralyzed; it is usually the consequence of a cerebro-vascular accident. One of the main consequences of hemiplegia is the drop-foot syndrome. Due to lack of controllability of muscles involved in flexing the ankle and toes, the foot drops downward and impede the normal walking motion. Today, there are commercially available assistive systems that use surface electrodes to stimulate Tibialis Anterior (TA) muscle and prevent drop-foot. The efficiency of drop-foot stimulators depends on the timing of stimulation and functionality of dorsiflexion motion. Classically, available stimulators use footswitches to detect foot on/off events. These discrete events allow only for triggering the stimulation and/or playing with the duration of the stimulation pattern, but does not allow for precise online modification of the pattern itself. We have developed algorithms to monitor the ongoing walking cycle by observing the valid limb movements. In order to ensure legs coordination during walking, we propose a robust phase estimation method based on the observer of a nonlinear oscillator. We have modified a commercial stimulator, ODSTOCK, in order to be able to trigger it using our own wireless sensors and algorithms. Agreement from Nîmes CPP (ethical committee) was obtained in June 2010 to run tests on patients. The protocol comprises 1) the control algorithm triggering the stimulator based on signals issued from one wireless inertial sensor placed on healthy shank, 2) a sensor setup including inertial sensors placed on deficient shank and foot, one goniometer measuring deficient ankle angle, one EMG sensor placed on stimulated TA and one instrumented carpet (GAITRITE) (fig. 7). Several patients have been included in the study and data is being processed [15][14].

6.2.2. eEMG Feedback Torque Control in FES

Participants: Mitsuhiro Hayashibe, Qin Zhang, Christine Azevedo Coste.

![Figure 7. Stroke patient study protocol description.](image)
Electrical stimulation (ES) is one of the solutions for drop foot correction. Conventional ES systems deliver predefined stimulation pattern to the affected muscles. However, time-variant muscle response may influence the gait performance as they are difficult to be taken into account in advance. Therefore, closed-loop ES control is important to obtain desired gait in presence of muscle response variation. In this work, a dual predictive control, which consists of two nonlinear generalized predictive controllers, is proposed to track desired torque. The stimulated muscle dynamics are modeled by Hammerstein cascades, with one representing stimulation to activation, the other representing activation to torque. Ankle dorsiflexion torque and ES-evoked EMG of tibialis anterior were recorded experimentally for model identification. The control scheme is validated by following desired torque trajectories with the identified model. The results show that the stimulation pattern obtained from the dual predictive control can produce good torque tracking according to the current muscle condition as shown in Fig. 8. The updates of model parameters were switched off after certain instant for both the excitation and contraction model. Consequently, the model prediction in the control was only driven by the model input and the last model parameter estimates. The dual predictive controller can still generate suitable control signal to obtain desired torque trajectory [23].

Figure 8. Torque reference (solid blue) and reproduced torque output by the proposed controller (dashed red). The updates of model parameters were switched off after 19.2s. Consequently, the model prediction in the control was only driven by the model input and the model parameter estimates at the time of 19.2s.

6.2.3. FES assisted Sit-to-Stand

Participants: Jovana Jovic, Christine Azevedo Coste, Philippe Fraisse, Charles Fattal.

Standing up is a common daily activity and a prerequisite to standing or walking. This frequently executed task is one of the most biomechanically demanding activities. The ability to rise from a sitting to a standing position is very important for individuals with paraplegia in order to achieve minimal mobility and has functional and therapeutic benefits related to bone loading, joint extension, cardio-circulatory stimulation, and pressure sore prevention. One method which has been widely investigated is functional electrical stimulation (FES) of the lower extremities. The sit-to-stand method, which is widely used in clinical practice, involves open-loop stimulation of knee extensors activated by hand switches. This technique works adequately in many cases, however, in applying this strategy, stimulation starts without referencespect to the upper body movement. Hence, the whole-body motion is not optimal and requires a high velocity of the joints and large upper limb forces during the rising motion, which may cause both damage of joint tissues and shoulder complications.
We propose a "patient-driven" FES method that would coordinate motion of the trunk, which is under voluntary control of the patient, and motion of the lower limbs, which are under FES control. The proposed approach is based on the observation of trunk movement during rising motion and a detection algorithm, which triggers a pre-programmed stimulation pattern. Trunk acceleration was acquired by a single one-axis wireless accelerometer positioned on the subject’s back. The detection algorithm consists of an online comparison of the movement acceleration of the ongoing motion with the reference pattern (a typical pattern characterizing the sit-to-stand transfer for each subject) using Pearson’s correlation coefficients [25]. Experiments on paraplegic subjects are ongoing in rehabilitation center PROPARA. The experimental setup and a paraplegic subject of the experiment are presented in Fig. 9.

Figure 9. Experimental setup.

We have shown that in the cases where the acceleration and reference signal are similar, our algorithm is able to recognize sit-to-stand motion and to properly trigger leg stimulation at the desired instant. Also, we have shown that there is an influence of stimulation timing on applied hand forces during the motion. The best results were achieved for trials in which motion was similar to the one of the able bodied subjects in terms of trunk motion and the beginning of the leg motion with respect to the trunk acceleration signal.

We also investigated dynamic optimization as a tool to improve FES assisted sit to stand transfers of paraplegic subjects. The objective would be to find optimal strategy for voluntary trunk movement, which would minimize hip, knee and ankle torques and demand minimal upper limb participation during the motion. Our results suggest paraplegic patients should bend their body forward in order to use linear momentum of the trunk in sit off phase. Figure 10 shows optimal coordination of ankle, knee, and hip angles during sit-to-stand motion [26], [24].

6.2.4. Signal-based segmentation of human locomotion using embedded sensor network

Participants: Maud Pasquier, Christine Azevedo Coste, Bernard Espiau, Christian Geny (CHU Montpellier), Fabien Jammes.

Last year, we introduced a simple approach to segment in homogeneous phases a long-duration record of locomotion data consisting of body segment acceleration and foot pressure information. We used a system based on a network of wireless nodes embedding various types of sensors [3]. Two cases were considered: walk and run around an indoor running track [35] and outdoor marathon [34].
We now use this system as part of a study of mobility impairment caused by Parkinson’s disease (fig. 11). Freezing of gait (FOG) has been identified as one of the main contributors to gait disturbances in this disease. We introduce an ambulatory gait analysis method using body attached gyroscopes and accelerometers to detect the freezing of gait. One hand, we aim at proposing a FOG detection algorithm more robust because the existing algorithms were not able to detect the FOG without tremor. On the other hand, we would like to anticipate the freezing before it is installed in order to reduce the risk of falling.

Before and during a FOG, a patient tends to walk slowly with short strides and fast rhythm. The detection of an increase of frequency is not enough, because there exists similar variations during the initiation of gait or a voluntary acceleration. The association of an increase of the gait rhythm together with a decrease of stride length allows us to detect a FOG. The computation of correlation coefficient in a moving window allows us to estimate the rhythm of strides. We are also working on different methods to estimate the stride length using one or two IMU (3-axis gyroscope and 3-axis accelerometer).

A time-frequency representation permits to show an increase of fundamental frequency and a duplication before a FOG, Fig. 12. The variations of the fundamental frequency are already detected with the correlation. In the future works, we aim at characterizing these duplications and to propose an algorithm of automatic detection.

6.2.5. Awake surgery: How to optimize functional brain mapping by improving per-operative testing?

Participants: Cheikh Niang, Pom Charras, Stephane Argon, Christine Azevedo, Hugues Duffau, David Guiraud, François Bonnetblanc.

It is now possible to perform resections of slowgrowing tumors in awake patients. Using direct electrical stimulation (DES), real-time functional mapping of the brain can be used to prevent the resection of essential areas near the tumor. For now, simple clinical tests are performed on conscious patients and combined with DES in order to discriminate functional and non-functional areas invaded by the tumors. In this work we try to develop a simple device based on a simple technology to better quantify the performances of the patients during the surgery itself and give a real-time feedback to the neurosurgeon that will help to further guide the surgery.
Figure 11. nodes disposition on the patient.

Figure 12. FOG: Smoothed pseudo Wigner-Ville time-frequency distribution with causal kernel.
by improving the sensibility of the functional mapping. This procedure should also allow building a strong database that should serve retrospectively to improve the surgical procedure and reinforce the neurosurgeons’ experience as well as to monitor the patients’ performances all along their life.

6.2.6. Closed-loop CoM based posture control in FES

**Participants:** Alejandro González, Mitsuhiro Hayashibe, Philippe Fraisse.

Center of Mass control has been used in humanoid robotics to create stable standing postures and movements. By controlling the CoM position and acceleration, joint trajectories which respect to the ZMP stability criterion can be generated. FES may be used to drive joint torques in order to maintain a standing posture within a closed loop controller. Current work is focused on locating a human’s CoM by creating a statically equivalent serial chain (SESC) model using widely accessible equipment, such as the Kinect camera and the Wii balance board.

6.3. Neuroprostheses

6.3.1. Stimulator calibration

**Participants:** Jérémie Salles, Fabien Soulier, Serge Bernard, Guy Cathébras.

![Figure 13. Stimulator architecture overview.](image)

In the context of the TIME project, one CAFE12-based stimulator will be used for chronic experiment in human. During the validation of the stimulator, it appeared that we needed to improve both the linearity and the current matching of the 12 channels. We thus define a calibration process consisting in:

- **PCB modification:** To take advantage of the 10 bits of the DAC (only 8 were used before for compatibility reasons). The modification give the FPGA access to the two latter bits. Moreover, test points were added between the ASIC and the output capacitors that now can be removed for the calibration phase.
- **Digital interface modification:** To allow a 16th current amplifying ratio (only 0-15 were enable). Improvements in the activation sequences of output current mirrors have also been carried on.
- **Reference voltage and current tuning:** The stimulator use several level of power supply and voltage references. Nominal values are:

\[
\begin{align*}
V_{hv} &= 16.0 \, \text{V}, \\
V_{DD} &= 3.3 \, \text{V}, \\
V_{hv2} &= 13.3 \, \text{V}, \\
V_{ref\, P} &= 13.5 \, \text{V}, \\
V_{ref\, N} &= 1.5 \, \text{V}.
\end{align*}
\]
632

Computational Sciences for Biology, Medicine and the Environment - New Results - Project-Team
DEMAR

The DAC current reference is set to get a 5.46 mA maximum current at the stimulator output for ratio of 15. We have enabled modification of the biasing resistor in order to fine-tune this current reference.

- Raw data acquisition: Measurement of the 12 output currents are carried on independently with the following configuration:
  - all outputs configured as cathodes,
  - all ratios set to 16.

The measurement setup makes use of a characteristic analyzer (HP4156A) to maintain the voltage load to 7.5 V for all the DAC current values.

- Correction: The linearity and matching correction is specific to the association of a particular ASIC with a particular DAC. For human experiment, the configuration will be limited to common anode/controlled cathodes. A first-order linear regression is applied to the 12 raw current measurements. This gives gain and offset adjustments for each channel that are applied by a linear digital correction block (fig. 14). The DAC initial value (8-bit) is multiplied by the correction gain and summed to the correction offset (both channel-dependant), resulting in a 10-bit corrected command. At last, respective correction values are chosen and quantified to lower impact of these modifications to the precision and dynamic range of the stimulator output (no “lost” bit). Concerning the implementation, since the 3 most significant bits of the correction gain appear to be constant, it is possible to use channel-independent bit-shifts and a subtractor to perform the 8 to 10-bit multiplication.

![](image)

*Figure 14. Linear digital correction principle.*

The improvement in the matching of the 12 channels can be seen in figure 15. The results in terms of linearity are an integral non-linearity of ±2.5 LSB and a differential non-linearity of ±0.3 LSB.

6.3.2. Nerve Modelling for ENG recording

**Participants:** Olivier Rossel, Jonathan Coulombe, Fabien Soulier, Serge Bernard, Guy Cathébras.

In the context of FES, neural recording is one of the main issues, as the control requires information carried on afferent peripheral nerves. Because specific information are carried in different fascicles, we propose to realize a non-invasive and spatial-selective electrode. Last year, based on investigation on the topic of extracellular Action Potentials (AP), we proposed a new tripole design, where the tripolar output signal is the image of the activity in the close vicinity of this tripole, providing high spatial selectivity.

We showed however, that this high spatial selectivity is achieved at the expense of signal amplitude. This first result jeopardizes the feasibility of this kind of electrode since the signal amplitude appears to be on the same range of the expected noise. First, we propose to estimate the performance of the proposed electrode with a quantitative study of the electrode selectivity. Then, to conclude on the feasibility of this electrode, the SNR has to be determined. So with a more accurate model, we studied the sensitivity of the proposed tripole, allowing to determine precisely the amplitude level of the expected signal. Thus, the SNR can be estimated knowing the expected noise.
Figure 15. Mean gain (top) and worst case (bottom) before and after correction.
In short, the work of this year aims at characterizing the performances and evaluating the feasibility this new multi-contact cuff electrode.

6.3.2.1. Selectivity

We proposed an electrode configuration inspired from the FINE electrode (figure 16) designed for the same purpose. The electrode is composed of many tripoles, placed around the nerve. This disposition is used for two electrode, state-of-the-art electrode A and the proposed electrode B. The unique difference between both electrode resides on the longitudinal inter-pole distance ($d_e$), which is respectively 5 mm for the electrode A and 0.375 mm for the electrode B.

![Figure 16. Electrodes A and B have the same shape but differ on the longitudinal inter-pole distance ($d_e$). Two fascicles are represented.](image)

The electrodes performances are evaluated based on simulations using a model of a nerve comprising multiple fascicles [38]. The Selectivity Index ($SI$) quantifies the ability to record and distinguish between different active fascicles in such a manner that $SI = 0$ corresponds to a case where an active fascicle yields identical signals at every recording site, while $SI = 1$ occurs when one recorded signal is different from every other. This $SI$ has to be presented according to the inter-fiber spacing.

The result of electrode selectivity are presented in the fig. 17. This figure shows that activity of two fascicles separated by as little as 1 mm can be distinguished for the proposed electrode (for this distance $SI$ for electrode B ($\approx 0.9$) is more than double that of electrode A ($< 0.4$)). The proposed electrode thus appears to be much more selective than the reference electrode.

6.3.2.2. Sensitivity

Using a more realistic model (inhomogeneous and anisotropic), we investigate the spatial properties of extracellular AP and that of the filtering done by the proposed tripole [37]. This allows us to represent the tripopolar sensitivity. It was realized for the proposed tripole B and compared to a state-of-the-art tripole A [18]. This sensitivity represents the amplitude of the tripolar output signal for a single unit action potential.

This figure shows that the classical tripole radial sensitivity is huge compared to that of the proposed electrode. This confirms the high spatial sensibility of the proposed tripole. We can also determine the expected amplitude, where the signal can reach 6 $\mu$V. Considering this amplitude and knowing that in natural case there will be superposition of action potentials, we can conclude that the signal amplitude could be higher than the expected noise (around 1 $\mu$V). So we can conclude positively to the feasibility of this kind of electrode.

6.3.3. Low-noise, low-power ENG amplifier design

Participants: Jonathan Coulombe, Olivier Rossel, Fabien Soulier, Serge Bernard, Guy Cathébras.

This year we proposed a method for enhancing the noise-power tradeoff of front-end amplifiers in parallel recording applications of analog signals with respect to a common reference. One example of application is shown in the Fig. 19 for spatial-selectivity ENG recordings.

The circuit architecture is based on a Shared-Input Amplifier (SIA), composed by shared-input transconductance amplifiers and a differential stage. Averaging null signals and subtracting the result from every signal reduce the noise because the correlated noise between parallel outputs is attenuated. It results significant supply current savings without noise penalty. One example is shown in the Fig. 20 for the specific case of two average of two null signals.
Figure 17. Selectivity index computed for random combinations of simulated fascicles, plotted as a function of the distance between each couple of fascicles.

Figure 18. Comparison between classical and proposed tripole sensitivity. The peak-to-peak amplitude of a single unit action potential is represented, measured by a classical tripole on the left and by the proposed tripole on the right. The tripoles are placed on the surface of a nerve of 300 µm radius.
Figure 19. Conceptual representation of a selective ENG recording system.

Figure 20. SIA modified for the use of transconductance amplifiers in open-loop configuration for ENG recordings. And concept of parallel SIA with noise reduction using null output averaging and differential readout.
Also, a method for reducing the remaining noise with little power penalty is possible. And it is possible to combine both methods, either noise level, total supply current, or both can be significantly reduced. The benefits of combining both methods and the related trade-offs was validated by simulations using models of a 0.35 µm BiCMOS process. So we have shown that the total supply current can be reduced by more than 50% that of a comparable system using conventional differential amplifiers with equivalent output noise. Alternatively, the noise can be reduced by approximately 35% with comparable power consumption.

This should enable low-noise recording of signals with significantly better efficiency than even the theoretical limit of any conventional differential amplifier. Future work will include circuit optimization and investigation of the impact of the architecture over other performances of the system, such as crosstalk, linearity, distortion, and channel mismatch. Implementation of the circuit for full characterization is expected to be completed in the near future.
5. New Results

5.1. Model Design

5.1.1. Sensitivity Analysis of Complex Biophysical Models

5.1.1.1. Background

Sensitivity analysis (SA) is a fundamental tool in the building, use and understanding of mathematical models [44]. Sampling-based approaches to uncertainty and sensitivity analysis are both effective and widely used [38]. For this purpose, Sobol’s method is a key one [47]. Since it is based on variance decomposition, the different types of sensitivity indices that it estimates can fulfill different objectives of sensitivity analysis: factor prioritization, factor fixing, variance cutting or factor mapping [37]. It is a very informative method but potentially computationally expensive [38]. Besides the first-order effects, Sobol’s method also aims at determining the levels of interaction between parameters [48]. In [46], the authors also devised a strategy for sensitivity analysis that could work for correlated input factors, based on the first-order and total-order index from variance decomposition.

5.1.1.2. Algorithm numerical implementation

Computational methods to evaluate Sobol indices sensitivity rely on Monte-Carlo sampling and re-sampling [47], [40]. For $k$ dimensional factor of model uncertainty, the $k$ first-order effects and the ‘$k$’ total-order effects are rather expensive to estimate, needing a number of model evaluations strictly depending upon $k$ [43]. Therefore, it is crucial to not only devise efficient computing techniques, in order to make best use of model evaluations [45], but also to have a good control of the estimation accuracy with respect to the number of samples.

With the objective of an efficient computational method for sensitivity analysis of functional-structural tree growth models, we proposed a new estimator based on Homma-Saltelli method to compute Sobol indices, which improves slightly their use of model evaluations thanks to a more balanced resampling strategy. This new estimator can be considered as an effort to improve the efficiency of SA methods for models.

We also derived a theoretical analysis of the error estimation for the sensitivity analysis for the studied class of Sobol’s estimators (it can be applied to all the three Sobol’s estimators mentioned in this paper) with respect to the sampling size and the number of model evaluations. An analytical test function is used to test the error estimation, and we obtained that the error estimation in this paper gives out a better ‘upper bound’ than the previous works related to this problem. This error estimation directly relates to the variance of the result, so it can also be used for checking the confidence interval, which is usually difficult to attain.

We derived a theoretical analysis of the error estimation for the sensitivity analysis for the studied class of Sobol’s estimators (it can be applied to all the three Sobol’s estimators mentioned in this paper) with respect to the sampling size and the number of model evaluations. An analytical test function is used to test the error estimation, and we obtained that the error estimation in this paper gives out a better ‘upper bound’ than the previous works related to this problem. This error estimation directly relates to the variance of the result, so it can also be used for checking the confidence interval, which is usually difficult to attain.

The result has been accepted to be published in J of Rel. Eng. and Sys. Safety. Based on the published result for first order index of Sobol’s, we also extended this method to the second order index computing which is very important for us to know the precise pairs parameters with interactions between them, with the objective of making best use of the input-output model evaluation matrices that involve in the main part of the sensitivity analysis computing cost. Besides the computing efficiency improvement, one side result we got from the algorithm for second order index is that it can also make the final index has smaller variance so that the result can be more reliable.
Besides Sobol’s method, we also tried Morris method to complete the aim of ‘screening’ parameter of sensitivity analysis.

5.1.1.3. Strategies for FSPM

Global sensitivity analysis (SA) has an important role to play in functional-structural plant growth modeling by assessing the different source of uncertainty help us to gain some insights inner the models so as to explain the behavior of them. Different FSPMs have different scales of model design, which leads to all types of diverse multi-biophysical processes.

To study specifically how global SA can help for FSPMs, SA was applied on a wide variety of functional-structural plant models, typically the 3 FSPMs: firstly a simple source-sink model of maize growth, is used to specifically study the process of carbon (C) allocation among expanding organs during plant growth, with simple plant structure, multi-stage and detailed observations, secondly the GreenLab model of tree growth (applied to poplar tree) characterized by the retroaction of plant functioning on its organogenesis [41], which describes tree structural plasticity in response to trophic competition, lastly a functional-structural model, NEMA [16], describing C and nitrogen (N) acquisition by a wheat plant as well as C and N distributions between plant organs after flowering. This model has the specificity to integrate physiological processes governing N economy within plants: root N uptake is modeled following the transport systems high affinity transport systems (HATS) and low affinity transport systems (LATS), and N is distributed between plant organs according to the turnover of the proteins associated to the photosynthetic apparatus. C assimilation is predicted from the N content of each photosynthetic organ. Consequently, this model is more mechanistic but also more complex than the two previous ones. Another objective is to explore an effective simulation design to help the sensitivity analysis for complicate models with several logically distinct but biological functioning interacted moduls.

All these SA result shadowed a light to the models for us to diagnosis the model behavior and will bring a big step for parameter estimation and experimental simplification in our modelling next.

5.1.2. A model for Cecropia sciadophylla under fluctuating environmental conditions

In collaboration with Patrick Heuret (INRA, JRU Ecofog, Kourou, French Guiana), we developed a tree growth model dedicated to *Cecropia sciadophylla*, a neotropical species from the genus Cecropia. These trees have interesting properties from a modeller’s point of view: they have a simple architecture, their number of phytomers remain limited even for old individuals, and most importantly, [39] and [49] have developed a methodology based on morphological observations to estimate tree age on *C. obtusa* and *C. sciadophylla* respectively. It is therefore possible to fully describe the tree structure and topology from morphological observations, which is very uncommon for trees: for most tree species, their high stature, complex structure, and long life span drastically increase the fieldwork required to collect data at the organ scale and hamper the development, calibration and validation of functional-structural tree growth models and their potential applications in the field of forest management.

We used datasets collected on 18 trees in 2007 and 2008 in French Guiana to develop and evaluate our model. Our objective was to analyse the influence of fluctuating environmental conditions on the dynamics of trophic competition within *C. sciadophylla* trees. We defined an integrated environmental factor that includes meteorological medium-frequency variations and a relative index representing the local site conditions for each plant. The meteorological variations were input from pluviometry data, that could be considered as the main fluctuating environmental stress under that tropical climate. The relative index was estimated based on inversion of our model using data from respectively 11 trees for model calibration (those measured in 2007) and 7 trees for model evaluation (those measured in 2008). This study provided a model that can be seen as a tool to disentangle the ontogenic variations (low-frequency trend) and the environmental variations (medium-frequency variation). One paper was accepted for publication [22].
5.1.3. Using model inversion to analyze the effects of inter-tree competition on four Pine trees grown under two contrasted density conditions

In collaboration with Guo Hong and Lei Xiangdong (Chinese Academy of Forestry, Beijing, China), we analyzed the characteristics of individual tree response to competition on source-sink balance through the calibration of the GreenLab model.

Four Chinese pine trees (*Pinus tabulaeformis* Carr.) were destructively measured in November 2009 from the nursery garden located in the Yuanyiqi forest farm, Beijing, China. Two 13-year-old trees (T1 and T2) were from a high density plantation (3500N/ha) and two 10-year-old trees (T3 and T4) were from a low density plantation (2000N/ha). We first examined the statistical differences in the tree morphologies and topologies. Significant differences were found for internode diameter, internode biomass and needle biomass between the two densities, but not for internode length. In a second step, we studied the ability of the GreenLab model to simulate the plasticity of pine trees grown under different densities. To fulfil these objectives, it was necessary to find a way to characterize the competition conditions of each tree. Given the inherent difficulty of identifying the most relevant experimental measurements for this characterization, we proposed to represent the effects of competition on the tree growth through a single tree-specific parameter of GreenLab, called characteristic surface area, and to estimate it for each tree by model inversion, together with the more classical endogenous species parameters. This will eventually allow us to examine whether the obtained value of this characteristic surface area could be correlated to other possible indicators of competition pressure. This could pave the way to the development of an individual-based stand growth model including the effects of a competition index. One paper was submitted to *Trees - Structure and Functions*.

5.1.4. Coffee trees and genetics

In collaboration with Sylvie Sabatier (INRA, AMAP), Philippe de Reffye (CIRAD, AMAP) and Perla Hamon (IRD Montpellier), we studied the architectural and genetic diversities in 5 Coffea species, native from Madagascar. We explore two complementary methods: the genetic diversity using molecular markers (genomic- and/or EST-microsatellites) and the variability of adaptive traits between populations with different ecological niches. We focused on 5 Coffea species endemic to Madagascar, some of which are classified as critically endangered in the World Conservation Union (IUCN) Red List. For each species, architecture and genetic comparative analyses between individuals growing in situ (natural forest) and ex situ (common garden test) are being performed. In parallel, the same populations are analysed using the GreenLab model. These results will be used to study the potential links between the parameters of GreenLab and the allelic distribution in these populations. This is the subject of the PhD of Domohina Andrianasolo (CIRAD, Montpelier and FOFIFA, Antananarivo, Madagascar). This work was presented at the XVIII International Botanical Congress (IBC) [34].

5.1.5. Methods for tree crown analysis and application to young Eucalyptus

Based on the pioneer work on coffee trees of Philippe de Reffye, a stochastic model was developed to describe the topological development of trees. In the model, growth and branching processes are driven by the respective probabilities of activity, rest or death of apical and lateral buds. Because of its mathematical formulation, the model inversion can be done analytically which is rare and parameter values can be estimated from experimental data. The MATLAB software GLOUPS developped by Philippe de Reffye was used. We explored the feasibility of calibrating this stochastic model for eucalyptus, which presents the additional difficulty of a continuous growth with no marked endogenous cessation. Incomplete systems were also defined for the case, common with trees, of incomplete datasets. An adequate strategy was defined to sample measurements and applied to five eucalyptus trees (data collected by Pr Lei Xiangdong, Guo Hong and Diao Jun, Chinese Academy of Forestry, Beijing, China).

One paper was accepted for publication [20].

5.2. Model Evaluation and Parameter Estimation
5.2.1. Maximum Likelihood Estimation

In [18], a first approach for parameter estimation was introduced based on the assumption of an underlying deterministic model of biomass production and uncorrelated errors in the mass measurements of different types of organs in the plant structure. A novel idea is developed on the modeling plant growth in the framework of non-homogeneous hidden Markov models, for a certain class of plants with known organogenesis (structural development). Unknown parameters of the models are estimated via a stochastic variant of a generalised EM (Expectation-Maximization) algorithm where both steps (E-step, M-step) are non-explicit. For this reason, the E-step is approximated via a sequential Monte-Carlo procedure (sequential importance sampling with resampling) and the M-step is separated into two steps (Conditional-Maximization), where before applying a numerical maximization procedure (quasi-Newton type), a large subset of unknown parameters is updated explicitly conditioned on the other subset. The model is tested with real data and the results are satisfying. Further work is in progress, including MCMC techniques for parameter estimation (with the collaboration of Dr. Sonia Malefaki from the University of Patras, Greece) and Bayesian type estimation, see [33].

5.2.2. Convolution Particle Filter for parameter estimation

Although Kalman filter is applied to various fields and dominated for decades, it is limited by its assumptions of linearity. Particle filter, which combines Bayesian inference with Monte Carlo sequential sampling approach, is a method using different combinations of random variables sampled directly from the parameter space (or the state space) to estimate parameters and states of a complex system. These combinations, generally called particles, propagate by introducing new observations and provide updated posterior distributions by taking into account their weights. Meanwhile, a resampling procedure is used to prevent the degeneracy problem. Since classical filtering methods are generally not able to estimate the dynamical state vector along with the unknown parameters, the convolution particle filter is implemented based on convolution kernel approximation to meet the need while modelling with Markovian dynamical system. Several tests are carried out to examine the performance of the Convolution Particle Filtering method [42], [36], and efforts are made to find the optimal perturbation parameters. The applications of the method rely on the Lotka-Volterra model and the sugar beet model. Since the quality of the estimations is limited by the number of the observations, the Conditional Iterative Bayesian Filtering method is applied. The principal is simply to use the posteriori distributions as the a priori information to re-perform over and over again the estimation algorithm and each time we introduce the same sequence of observations. This approach helps us to improve significantly the final estimation of the hidden states and the unknown parameters while testing with the virtual data. The bootstrapping method is implemented in order to compare with the results from different methods. In the case of applications based on dynamical stochastic systems, two types of noise are introduced, one is involved in the modelization technique and the other is attached to the observation procedure. An alternation of deterministic parameter estimation and stochastic parameter estimation is proposed (in progress) which allows us to estimate these two kinds of parameters at the same time.

5.2.3. Modelling the inter-individual variability of organogenesis in sugar beet populations

Modelling the inter-individual variability in plant populations is a very important issue to enhance the predictive capacity of plant growth models at the field scale. In the case of sugar beet, this variability is well illustrated by the phyllochron (the thermal time elapsing between the appearance of two successive leaves): if the mean phyllochron remains very stable across seasons, there is a high heterogeneity among individuals. Likewise, seedling emergence may strongly vary within a population, potentially inducing important variations in individual plant productions. A hierarchical segmented model was used to describe and study the variability of the dynamics of leaf appearance in sugar beet crops. The use of this nonlinear mixed model allows for a better handling of the heterogeneity in the plant population, and gives estimates of this variability: each model’s parameter is considered as a random variable, varying from one plant to another around a mean population value, with a given variance.
These mean population values and inter-individual variability can be used as input of functional-structural models, the main issue being then to compute the propagation of these sources of probabilistic uncertainty in the dynamic system of Greenlab.

5.3. Methods for the Applications

5.3.1. Optimization of Phenotyping based on a Parameter Selection Methodology

The model Cornflo is a functional plant growth model simulating Corn’s growth and yield. Based on it, the classification of environmental scenarios is researched in term of their influences to corn’s yield, and their parameter estimation capabilities for Cornflo parameters. The initial qualitative analysis of parameter estimation results shows that environmental scenarios’ classification benefit estimation accuracy and identifiability. Currently this project is researched from three aspects. Firstly, different clustering techniques are tested to find the most proper scenarios categories for parameter estimation. Secondly, the scenarios clusters are used for botanical experiments optimization, such as the selection of experimental locations. Lastly, parameter estimation is optimized and researched in a practical use for plant growth models.

5.3.2. Plant-Soil interaction and Optimal Control of Irrigation

This work is performed in collaboration with JC Mailhol (Cemagref). Irrigation scheduling is an important issue for crop management, in a general context of limited water resources and increasing concern about agricultural productivity. Methods to optimize crop irrigation should take into account the impact of water stress on plant growth and the water balance in the plant-soil-atmosphere system. For this purpose, different plant-soil interaction models are proposed to simulate the functional plant growth. In particular, a compartment plant model is designed to integrate water stress impact on different main physiological processes of crop: biomass production, biomass allocation, and foliar senescence. This model is applied and calibrated for maize, in order to predict the harvest index according to the stress undergone by crop during its whole cycle. As for the optimal control problem of irrigation, it can be formulated by considering a price for the crop yield and for the water resource. Dynamic programming is then applied to the plant-soil system to determine an optimal irrigation strategy.
6. New Results

6.1. Modelling of Erythroblastic Islands (red blood cell production)

The production of red blood cells, erythropoiesis, occurs in the bone marrow, where immature erythroid cells differentiate and produce red blood cells. Differentiation and maturation of immature red blood cells occurs in very specific spatial structures called erythroblastic islands. They consists of a macrophage (big white blood cell) surrounded by immature cells and providing them with survival factors. Using a hybrid model, made of a discrete model describing cell-cell interactions and accounting for spatial interactions, and a continuous model describing intracellular protein regulation (deciding for cell fate), we showed the importance of the central macrophage in the erythroblastic island, in order to prevent unstable islands leading either to cell populations extinction or excessive proliferation. This result is actually under review (Journal of Theoretical Biology), partial results have been already published (Math. Model. Nat. Phenom.)

6.2. Modelling of the CD8 T cell Immune Response

The CD8 immune response is a specific immune response triggered by the organism when the innate response is unable to fight a pathogen. We proposed a new model of the CD8 T cell immune response based on the description of feedback controls exerted by the cytotoxic CD8 T cell population on the pathogen and the population itself. This model, a system of ordinary and age-structured partial differential equations, allows to describe a classical response, characterized by a cellular expansion following the pathogen-mediated activation, then a contraction phase and the generation of memory CD8 T cells. Moreover, we showed the global asymptotic stability of this system corresponding to the elimination of the virus. This situation is expected and describes for instance what is observed with the flu virus. We are now confronting the model to experimental data being generated by Jacqueline Marvel’s team in Lyon (immunology team). The analysis of the model and the first results have been published in Journal of Biological Systems.

6.3. Particle Dynamics Methods of Blood Flow Simulations

Various particle methods are widely used to model dynamics of complex media. In our work molecular dynamics and dissipative particles dynamics are applied to model blood flows composed of plasma and erythrocytes. The properties of the homogeneous particle fluid are studied. Capillary flows with erythrocytes are investigated.

6.4. Periodic linear cell cycle models

Several results on periodic linear cell cycle models were obtained in collaboration with Frederique Billy (Inria Bang), Jean Clairambault (Inria Bang), Olivier Fercoq (Inria Maxplus), Stéphane Gaubert (Inria Maxplus) and Thomas Ouillon (Ensta). Those results are currently in revision (minor modifications requested). It deals with the property of the growth rate in such models. Several aspects are discussed, among wich: effect of the variability of the division process on the growth rate, fitting coefficients to data obtained by FUCCI methodology and optimization procedure for the growth rate.

6.5. Relaxed cross diffusion models

A general well posedness result for relaxed cross diffusion models was obtained by members of our team in collaboration with Michel Pierre and Guillaume Rolland (IRMAR, Rennes). A paper in this subject has been submitted.
ESTIME Project-Team (section vide)
6. New Results

6.1. Fluid motion estimation

6.1.1. Multiscale PIV method based on turbulent kinetic energy decay

**Participants:** Patrick Héas, Dominique Heitz, Etienne Mémin.

We have proposed a new multiscale PIV method based on turbulent kinetic energy decay. The technique is based on scaling power laws describing the statistical structure of turbulence. A spatial regularization constraints the solution to behave through scales as a self similar process via second-order structure function and a given power law. The real parameters of the power-law, corresponding to the distribution of the turbulent kinetic energy decay, have been estimated from a simple hot-wire measurement. The method has been assessed in a turbulent wake flow and grid turbulence through comparisons with HWA measurements and other PIV approaches. Results have indicated that the present method is superior because it accounts for the whole dynamic range involved in the flows.

6.1.2. Stochastic uncertainty models for motion estimation

**Participants:** Thomas Corpetti, Etienne Mémin.

In this work we have proposed a stochastic formulation of the brightness consistency used principally in motion estimation problems. In this formalization the image luminance is modeled as a continuous function transported by a flow known only up to some uncertainties. Stochastic calculus enables to built then conservation principles which take into account the motion uncertainties. These uncertainties defined either from isotropic or anisotropic models can be estimated jointly to the motion estimates. Such a formulation besides providing estimates of the velocity field and of its associated uncertainties allows us to define a natural linear scale space multiresolution framework. The corresponding estimator implemented within a local least squares approach has shown to improve significantly the results of the corresponding deterministic estimator (Lucas and Kanade estimator). This fast local motion estimator has been shown to provide results that are in the same order of accuracy than state-of-the-art dense fluid flow motion estimator for particle images. This work has been published in a conference proceeding and has been accepted for publication in the journal *Ieee trans. on image processing* [23], [16]. We intend to pursue this formalization to define dense motion estimators that allows handling in the same way luminance conservation under motion uncertainty principles.

6.1.3. 3D flows reconstruction from image data

**Participants:** Ioana Barbu, Dominique Heitz, Cédric Herzet, Etienne Mémin.

Our work focusses on the design of new tools for the problem of 3D reconstruction of a turbulent flow motion. This task includes both the study of physically-sound models on the observations and the fluid motion, and the design of low-complexity and accurate estimation algorithms. On the one hand, state-of-the-art methodologies such as “sparse representations” will be investigated for the characterization of the observation and fluid motion models. Sparse representations are well-suited to the representation of signals with very few coefficients and offer therefore advantages in terms of computational and storage complexity. On the other hand, the estimation problem will be placed into a probabilistic Bayesian framework. This will allow the use of state-of-the-art inference tools to effectively exploit the strong time-dependence of the fluid motion. In particular, we will investigate the use of “ensemble Kalman” filter to devise low-complexity sequential estimation algorithms.
This year, we have more particularly focussed on the problem of reconstructing the particle positions from several two-dimensional images. Our approach is based on the exploitation of a particular family of sparse representation algorithms, namely the so-called “pursuit algorithms”. Indeed, the pursuit procedures generally allow a good trade-off between performance and complexity. Hence, we have performed a thorough study comparing the reconstruction performance and the complexity of different state-of-the-art algorithms to that achieved with pursuit algorithms. This work has led to the publication of two conference papers in experimental fluid mechanics [21], [34].

6.1.4. Motion estimation techniques for turbulent fluid flows

Participants: Patrick Héas, Dominique Heitz, Cédric Herzet, Etienne Mémin.

Based on physical laws describing the multi-scale structure of turbulent flows, this article proposes a regularizer for fluid motion estimation from an image sequence. Regularization is achieved by imposing some scale invariance property between histograms of motion increments computed at different scales. By reformulating this problem from a Bayesian perspective, an algorithm is proposed to jointly estimate motion, regularization hyper-parameters, and to select the most likely physical prior among a set of models. Hyper-parameter and model inference is conducted by likelihood maximization, obtained by marginalizing out non-Gaussian motion variables. The Bayesian estimator is assessed on several image sequences depicting synthetic and real turbulent fluid flows. Results obtained with the proposed approach in the context of fully developed turbulence improve significantly the results of state of the art fluid flow dedicated motion estimators. This work has been published in several conferences and in the journal Tellus, Serie A [18].

6.1.5. Wavelet basis for multi-scale motion estimation

Participants: Pierre Dérian, Patrick Héas, Cédric Herzet, Souleymane Kadri Harouna, Etienne Mémin.

This work aims at exploring wavelet representations for fluid motion estimation from consecutive images. This scale-space representation, associated to a simple gradient-based optimization algorithm, sets up a natural multi-resolution framework for the optical flow estimation well suited to medium range velocity magnitude. Moreover, a very simple closure mechanism, approaching locally the solution by high-order polynomials, is provided by truncating the wavelet basis at fine scales. Well-known turbulence regularities and multifractal behaviors on the reconstructed motion field can also be imposed on the wavelet coefficients. Accuracy and efficiency of the proposed method has been evaluated on scalar and particles image sequences of turbulent fluid flows. Particularly good results have been observed for particle image velocimetry. This offers a very interesting alternative to traditional PIV techniques. This work has been published in computer vision or turbulence conferences [26], [25].

6.1.6. Divergence-free wavelet basis and high-order regularization

Participants: Pierre Dérian, Patrick Héas, Souleymane Kadri Harouna, Etienne Mémin.

Expanding on a wavelet basis the solution of an inverse problem provides several advantages. Wavelet bases yield a natural multiresolution analysis that may alleviate the use of Gauss Newton strategy for medium range motion amplitude. The continuous representation of the solution with wavelets enables analytical calculation of regularization integrals over the spatial domain. By choosing differentiable wavelets, high-order derivative regularizers can be designed, either taking advantage of the wavelet differentiation properties or via the basis’s mass and stiffness matrices. Moreover, differential constraints on vector solutions, such as the divergence-free volume preserving constraint, can be handled with biorthogonal wavelet bases. Numerical results on synthetic and real images of incompressible turbulence show that divergence-free wavelets and high-order regularizers are particularly relevant in the context of incompressible fluid flows. This work has been partly published in a conference proceeding [25].

6.1.7. Divergence-free wavelet basis and high-order regularization

Participants: Pierre Dérian, Patrick Héas, Souleymane Kadri Harouna.
This work presents a method for regularization of inverse problems. The vectorial bi-dimensional unknown is assumed to be the realization of an isotropic divergence-free fractional Brownian Motion (fBm). The method is based on fractional Laplacian and divergence-free wavelet bases. The main advantage of these bases is to enable an easy formalization in a Bayesian framework of fBm priors, by simply sampling wavelet coefficients according to Gaussian white noise. Fractional Laplacians and the divergence-free projector can naturally be implemented in the Fourier domain. An interesting alternative is to remain in the spatial domain. This is achieved by the analytical computation of the connection coefficients of divergence-free fractional Laplacian wavelets, which enables to easily rotate this simple prior in any sufficiently “regular” wavelet basis. Taking advantage of the tensorial structure of a separable fractional wavelet basis approximation, isotropic regularization is then computed in the spatial domain by low-dimensional matrix products. The method is successfully applied to fractal image restoration and turbulent optic-flow estimation.

6.1.8. Bayesian inference of hyper-parameters and models in motion estimation

Participants: Patrick Héas, Cédric Herzet, Etienne Mémin.

Bayes rule provides a nice framework for motion estimation from image sequences. We rely on a hierarchical modeling linking the image intensity function variable, the motion field variable, hyper-parameters composed of the likelihood and prior model inverse variances and of robust parameters, and finally the observation and prior model. The variable dependence can thus be expressed as a 4-level hierarchy. Applying the Bayes rule on this hierarchy, we obtain three levels of inference, which enable us to obtain, by marginalizing out intermediate variables, a direct dependence of the variable of interest to the image intensity function. Thus, the estimates of regularization parameters, of robust parameters associated to semi-quadratic norms of a family of M-estimators, and of observation and prior models are inferred in a maximum likelihood sense while maximizing jointly the motion field a posteriori probability. The quality of the method is demonstrated on synthetic and real two-dimensional turbulent flows and on several computer vision scenes of the “Middlebury” data-base. This work has been accepted for publication in IEEE transaction on Image Processing (IP) [17].

6.1.9. Method to quantify the uncertainty of motion measurement

Participants: Patrick Héas, Dominique Heitz, Cédric Herzet.

Measurement uncertainty is a general concept associated with any measurement that can be used to quantify the confidence of the estimation. The ‘Guide to the expression of uncertainty in measurement’ (GUM) provides a framework to account for all uncertainties and then to propagate them. However, in particle image velocimetry (PIV), measurement uncertainty estimation is a tricky task since it has to be done through the propagation of distributions via Monte Carlo simulations for each velocity component and all pixel location in the image. Considering a standard Bayesian formulation of the optical flow problem together with a Gaussian assumption the uncertainty associated to the estimated velocity field has been provided and described. First PIV measurement uncertainty estimations and discussions have been recently published in a conference issue [36].

6.1.10. Sparse-representation algorithms

Participant: Cédric Herzet.

We have pursued the study of efficient sparse decomposition algorithms. In particular, we have addressed the problem of finding good sparse representations into a probabilistic framework [24]. First, we have showed that one of the standard formulations - the Lagrangian formulation - of this problem can be interpreted as a limit case of a maximum a posteriori (MAP) problem involving Bernoulli-Gaussian variables. Then, we have proposed different tractable implementations of this MAP problem and explained some well-known pursuit algorithms (MP, OMP, StOMP, CoSaMP and SP) as particular cases of the proposed algorithms. Experimentations led on synthetic data show a good general behavior of the proposed methods. Exploiting further this probabilistic framework, we have then considered the design of soft pursuit algorithms. In particular, instead of making hard decisions on the support of the sparse representation and the amplitude of the non-zero coefficients, our soft procedures iteratively update probability on the latter values. The proposed algorithms are designed within the framework of the mean-field approximations and resort to the so-called variational Bayes EM algorithm to implement an efficient minimization of a Kullback-Leibler criterion.
6.2. Tracking and data assimilation

6.2.1. Stochastic filtering for fluid motion tracking

Participants: Sébastien Béyou, Anne Cuzol, Sai Gorthi, Etienne Mémin.

We investigated the study of a recursive Bayesian filter for tracking velocity fields of fluid flows. The filter combines an Ito diffusion process associated to 2D vorticity-velocity formulation of Navier-Stokes equation and discrete image error reconstruction measurements. In contrast to usual filters designed for visual tracking problems, our filter combines a continuous law for the description of the vorticity evolution with discrete image measurements. We resort to a Monte-Carlo approximation based on particle filtering. The designed tracker provides a robust and consistent estimation of instantaneous motion fields along the whole image sequence.

When the likelihood of the measurement can be modeled as Gaussian law, we have also investigated the use of the so-called ensemble Kalman filtering for fluid tracking problems. This kind of filters introduced for the analysis of geophysical fluids is based on the Kalman filter update equation. Nevertheless, unlike traditional Kalman filtering setting, the covariances of the estimation errors, required to compute the so-called Kalman gain, relies on an ensemble of forecasts. Such a process gives rise to a Monte Carlo approximation for a family of non-linear stochastic filters enabling to handle state spaces of large dimension. We have recently proposed an extension of this technique that combines sequential importance sampling and the propagation law of an ensemble Kalman filter. This technique leads to an ensemble Kalman filter with an improved efficiency. This year we have investigated the introduction of a nonlinear direct image measurements operator within this ensemble Kalman scheme. This modification of the filter provides very good results on 2D numerical and experimental flows even in the presence of strong noises. We are currently assessing its application to oceanic satellite images for the recovering of ocean streams. We are studying also the impact on the stochastic dynamics of turbulent noise defined as auto-similar Gaussian random fields and the introduction within an incremental ensemble analysis scheme of multiscale motion measurements. This work has been published in conference issues [28], [27], [35].

6.2.2. Reduced-order models for flows representation from image data

Participants: Patrick Héas, Cédric Herzet, Etienne Mémin, Véronique Souchaud.

One of the possibilities to neglect the influence of some degrees of freedom over the main characteristics of a flow consists in representing it as a sum of $K$-orthonormal spatial basis functions weighted with temporal coefficients. To determine the basis function of this expansion one of the usual approaches relies on the Karhunen-Loeve decomposition (referred as proper orthogonal decomposition – POD – in the fluid mechanics domain). In practice, the spatial basis functions, also called modes, are the eigen vectors of an empirical auto-correlation matrix which is built from “snapshots” of the considered physical process.

In this axis of work we focus on the case where one does not have a direct access to snapshots of the considered physical process. Instead, the POD has to be built from the partial and noisy observation of the physical process. Instances of such scenarios include situations where real instantaneous vector-field snapshots are estimated from a sequence of images. We have been working on several approaches dealing with such a new paradigm. A first approach consists in extending standard penalized motion-estimation algorithms to the case where the sought velocity field is constrained to span a low-dimensional subspace [38]. Giving a probabilistic interpretation to this problem, we have designed novel optimization procedures in the framework of maximum a posteriori estimation problem. This work has led to the publication of a paper in the Gretsi conference. We are currently working on an EM-algorithm implementation of this approach.

In a second approach, we are considering the design of the POD as the solution of a minimum least squares estimation problem based on the distribution of the (unknown) velocity field given a sequence of images. This alternative formulation allowed us to take explicitly the uncertainty on the velocity field into account into our optimization process. We are currently working on several practical implementations of this problem, relying on Monte-Carlo integration and Krylov subspaces.
In a third axis we have studied two variational data assimilation techniques for the estimation of low order dynamical models for fluid flows. Both methods are built from optimal control recipes and rely on POD representation associated to Galerkin projection of the Navier Stokes equations. The proposed techniques differ in the control variables they involve. The first one introduces a weak dynamical model defined only up to an additional uncertainty time dependent function whereas the second one, handles a strong dynamical constraint in which the coefficients of the dynamical system constitute the control variables. Both choices correspond to different approximations of the relation between the reduced basis on which is expressed the motion field and the basis components that have been neglected in the reduced order model construction. The techniques have been assessed on numerical data and for real experimental conditions with noisy Image Velocimetry data. This work has been presented in several conferences. A journal paper has been recently accepted with minor changes to the journal of computational Physics.

6.2.3. Optimal control techniques for the coupling of large eddy dynamical systems and image data

**Participants:** Dominique Heitz, Etienne Mémin, Cordelia Robinson, Yin Yang.

This work aims at investigating the use of optimal control techniques for the coupling of Large Eddies Simulation (LES) techniques and 2D image data. The objective is to reconstruct a 3D flow from a set of simultaneous time resolved 2D image sequences visualizing the flow on a set of 2D plans enlightened with laser sheets. This approach will be experimented on shear layer flows and on wake flows generated on the wind tunnel of Irstea Rennes. Within this study we wish also to explore techniques to enrich large-scale dynamical models by the introduction of uncertainty terms or through the definition of subgrid models from the image data. This research theme is related to the issue of turbulence characterization from image sequences. Instead of predefined turbulence models, we aim here at tuning from the data the value of coefficients involved in traditional LES subgrid models or in longer-term goal to learn empirical subgrid models directly from image data. An accurate modeling of this term is essential for Large Eddies Simulation as it models all the non resolved motion scales and their interactions with the large scales.

First tests have been conducted with two-dimensional Direct Numerical Simulations (DNS) of mixing layer coupled with noisy observations. By modifying the initial condition of the system, the proposed method recovers the state of an unknown function with good accuracy. This work has been published in the International Symposium on Turbulence and Shear Flow Phenomena (TSFP) 2011 [29].

6.2.4. Free surface flows reconstruction and tracking

**Participants:** Benoît Combes, Dominique Heitz, Etienne Mémin.

Characterising a free-surface flow (space and time-dependent velocity and geometry) given observations/measures at successive times is an ubiquitous problem in fluid mechanic and in hydrology. Observations can consist of e.g. measurements of velocity, or like in this work of measurements of the geometry of the free-surface. Indeed, recently developed depth/range sensors allow to capture directly a rough 3D geometry of surfaces with high space and time resolution. The main purpose of this study is to evaluate the ability of the Kinect sensor to estimate time-dependent 3D free-surface geometries. Then, based on these observations and on a stochastic data assimilation method, we want to estimate both time dependent geometry and displacement field associated to a free-surface flow from a simple temporal sequence of Kinect data. This year we have demonstrated on real data the possibility to measure free surface flow geometry with Kinect sensor and on synthetic data to estimate both time dependent geometry and displacement field associated to a free-surface flow from a simple temporal sequence of Kinect-like data. This work has been published in the new conference Flow Volume Reconstruction [22]. We intend to extend such a study to hydrological applications.

6.2.5. Stochastic filtering technique for the tracking of closed curves

**Participants:** Christophe Avenel, Etienne Mémin.
We have proposed a filtering methodology for the visual tracking of closed curves. Opposite to works of the literature related to this issue, we consider here a curve dynamical model based on a continuous time evolution law with different noise models. This led us to define three different stochastic differential equations that capture the uncertainty relative to curve motions. This new approach provides a natural understanding of classical level-set dynamics in terms of such uncertainties. These evolution laws have been combined with various color and motion measurements to define probabilistic state space models whose associated Bayesian filters can be handled with particle filters. This ongoing work will be continued within extensive curve tracking experiments and extended to the tracking of other very high dimensional entities such as vector fields and surfaces. This work has been published in conference proceedings and a journal article is conditionally accepted to minor changes for publication in a meteorological journal.

### 6.2.6. Sequential smoothing for fluid motion

**Participants:** Anne Cuzol, Etienne Mémin.

In parallel to the construction of stochastic filtering techniques for fluid motions, we have proposed a new sequential smoothing method within a Monte-Carlo framework. This smoothing aims at reducing the temporal discontinuities induced by the sequential assimilation of discrete time data into continuous time dynamical models. The time step between observations can indeed be long in environmental applications for instance, and much longer than the time step used to discretize the model equations. While the filtering aims at estimating the state of the system at observations times in an optimal way, the objective of the smoothing is to improve the estimation of the hidden state between observation times. The method is based on a Monte-Carlo approximation of the filtering and smoothing distributions, and relies on a simulation technique of conditioned diffusions. The proposed smoother can be applied to general non linear and multidimensional models. It has been applied to a turbulent flow in a high-dimensional context, in order to smooth the filtering results obtained from a particle filter with a proposal density built from an Ensemble Kalman procedure.

### 6.2.7. Stochastic fluid flows dynamics under Gaussian uncertainty

**Participant:** Etienne Mémin.

In this research axis we aim at devising stochastic Eulerian expression for the description of fluid flow evolution laws incorporating uncertainty on the particles location. Such an uncertainty modeled through the introduction of a random term allows taking into account approximations or truncation effects performed within the dynamics analytical constitution steps. This includes for instance the modeling of unresolved scales interaction in large eddies simulation (LES) or in Reynolds average numerical simulation (RANS), but also uncertainties attached to non uniform grid discretization. This model is mainly based on a stochastic version of the Reynolds transport theorem. Within this framework various simple expressions of the mean drift component can be exhibited for different models of the random field carrying the uncertainties we have on the flow. We aim at using such a formalization within image based data assimilation framework and to derive appropriate stochastic versions of geophysical flow dynamical modeling.

### 6.2.8. Variational assimilation of images for large scale fluid flow dynamics with uncertainty

**Participants:** Souleymane Kadri Harouna, Etienne Mémin.

In this work we explore the assimilation of a large scale representation of the flow dynamics with image data provided at a finer resolution. The velocity fields at large scales is described as a regular smooth components whereas the complement component is a highly oscillating random velocity field defined on the image grid but living at all the scales. Following this route we have started to assess the performances of a variational assimilation technique with direct image data observation. Preliminary results obtained for a wavelet based 2D Navier Stokes implementation and images of a passive scalar transported by the flow are very encouraging.

### 6.3. Analysis and modeling of turbulent flows

#### 6.3.1. Mixing layers between a uniform flow and a shear flow

**Participant:** Dominique Heitz.
We have addressed the analysis and modelling of non canonical turbulent mixing layers between a uniform flow and a shear flow. From a parametric study by bidimensional direct numerical simulations two mixing layer configurations between a uniform flow and a shear flow have been selected. These two configurations share the same shear flow but have a different uniform flow.

The shear flow was obtained with curved gauze. However the theoretical shear parameter predicted by the literature is different from the value obtained by experiments. In order to study these discrepancies, the flow through a gauze was studied by particle image velocimetry. This allowed the general modeling of the uniform flow through curved wire gauze, leading to linear mean velocity profiles. From a hot-wire anemometry study of the two flow configurations it was observed that one flow behaves like a mixing layer whereas the other flow yields a wake behaviour. The mixing layer indicates an increasing turbulent kinetic energy along its longitudinal development, while the wake exhibits an asymmetry.

6.3.2. Hot-wire anemometry at low velocities

**Participant:** Dominique Heitz.

A new dynamical calibration technique has been developed for hot-wire probes. The technique permits, in a short time range, the combined calibration of velocity, temperature and direction calibration of single and multiple hot-wire probes. The calibration and measurements uncertainties were modeled, simulated and controlled, in order to reduce their estimated values. This year a patent application has been submitted.

6.3.3. Experimental studies for the assessment of turbulence statistical models

**Participants:** Patrick Héas, Dominique Heitz, Etienne Mémin.

[In collaboration with G. Artana and P. Minini (Univ. Bueno Aires)]

Selecting directly from images the most likely scaling motion priors enables the recovery of physical quantities related to the energy flux and the flow regularity. Such measurements are of major interest for turbulence studies. In particular, determining the energy flux across scales and characterizing intermittency is very important to assess the relevance of the statistical models proposed for atmospheric turbulence. Although, the measurement of flux and atmospheric flow regularity has already been obtained previously using in situ data, it required an important measurement campaign lasting several years based on sensors placed on airplanes. Therefore, the proposed motion estimation technique described above represents an attractive tool since it enables the direct estimation of these quantities from a couple of images. A paper concerning an atmospheric turbulence study using Meteosat Second Generation (MSG) images has been accepted in the journal Tellus A. Experimental studies of three-dimensional turbulence behind a grid or in the wake of a cylinder has been performed and will be submitted in the journal Experiments in fluids. New experiments for the assessment of turbulence statistical models are currently going on in collaboration with the laboratory of fluid mechanics and turbulence scientists in Argentina. They focus on two-dimensional turbulence of soap films visualized with a Schlieren imagery system. The goal of this work is to validate experimentally the theoretical model predicting non-intermittent inverse energy cascades in pure two-dimensional flows.

6.4. Visual servoing approach for fluid flow control

6.4.1. Fully exploitation of the controlled degrees of freedom of the 2D plane Poiseuille flow

**Participants:** Christophe Collewet, Xuan Quy Dao.

This works concerns the Phd of Xuan-Quy Dao and can be seen as an extension of the works carried out by Roméo Tatsambon. Since visual measurements are used, we propose here to use advanced visual servoing techniques to fully exploit the controlled degrees of freedom of the 2D plane Poiseuille flow. To achieve this goal we propose to design a control law based on partitioned visual servo control (this approach has been first proposed in the robotics community by [46] but in a very different context). Therefore, we have shown that, following this way and contrary to the literature concerning drag reduction, it becomes easy to simultaneously reduce the drag and the kinetic energy density of the flow. That is of great importance since the controlled flow is in an unstable state and may become turbulent when the kinetic energy density is growing. This key
problem is not well taken into account in the literature. Indeed, either the drag or the kinetic energy density is reduced, but never both of them. Moreover, we have shown that in practice, the way the drag is reduced does not influence the way the kinetic energy density is reduced. In addition, since dense visual measurements are used, our approach is very robust against measurement noise.

6.4.2. Visual servoing for the 3D plane Poiseuille flow  
Participant: Christophe Collewet.

We focus here on the 3D plane Poiseuille flow which is much more realistic than the 2D case. In that case, it can be shown that the reduced linearized flow is in a stable configuration. However, it is possible to find some bad initial conditions which causes the flow to present high transient energy growths. Indeed, a small perturbation velocity value in the reduced linearized system leads to a transient effect which is characterized by a growth in a short-time behaviour of the kinetic density energy, before a decay occurs. Practically, this transient effect, if not controlled, can cause transition to turbulence. Usually, the streamwise shear stress component at a point belonging to the wall is used as the output of the system in order to control it in a closed-loop fashion. We have proposed a vision-based approach to control this flow (see section 3.4 ). Our approach has revealed to be the most efficient approach in comparison to existing ones. Indeed, the transient energy is highly reduced. In addition, as in the 2D plane Poiseuille flow, the initialization problem is not of concerned in the vision-base approach. In addition, our approach is robust to measurement noise when a large number of flow measurements is available, which is possible in real practical situations. This work has been published in several conferences [33], [20], [32] and has been recently accepted for publication in the International Journal of Flow Control [19].

6.5. National Initiatives

6.5.1. ANR-COSINUS PREVASSEMBLE: Ensemble methods for assimilation of observations and for prevision in Meteorology and Oceanography  
Participants: Sébastien Béyou, Anne Cuzol, Etienne Mémin.

duration 36 months.

The purpose of this project is to further study ensemble methods -, and to develop their use for both assimilation of observations and prediction. Among the specific questions to be studied are the theory of Particle Filters and Ensemble Kalman Filters, the possibility of taking temporal correlation into account in ensemble assimilation, the precise assessment of what can and cannot be achieved in ensemble prediction, and the objective validation of ensemble methods.

The partners of this project are Laboratoire de Météorologie Dynamique/ENS (leader), Météo-France and three INRIA groups (ALEA, ASPI, FLUMINANCE).

6.5.2. ANR SYSCOMM MSDAG: MultiScale Data Assimilation in Geophysics  
Participants: Pierre Dérian, Patrick Héas, Dominique Heitz, Cédric Herzet, Etienne Mémin.

duration 36 months.

Changing scale is a well-known topic in physics (geophysics, fluid mechanics and turbulence, theoretical and statistical physics, mechanics, porous media, etc.) It has lead to the creation of powerful sophisticated mathematical tools: renormalization, homogenization, etc. These ideas are also used in numerical analysis (the so-called multigrid approach) for solving efficiently partial differential equations. Data assimilation in geophysics is a set of methods that allows to combine optimally numerical models in large spaces with large dataset of observations. At the confluence of these two topics, the goal of this project is to study how to embed the change of scales (a multiscale point of view) issue into the framework of geophysical data assimilation, which is a largely unexplored subject.

The partners of this 3 years project are the CEREA/CLIME INRIA group (leader), the LSCE/CEA, the INRIA groups MOISE and FLUMINANCE.
6.5.3. ANR SYSCOMM GeoFluids:
Participants: Patrick Héas, Dominique Heitz, Etienne Mémin, Véronique Souchaud.
duration 48 months.
The project Geo-FLUIDS focuses on the specification of tools to analyse geophysical fluid flows from image sequences. Geo-FLUIDS aims at providing image-based methods using physically consistent models to extract meaningful features describing the observed flow and to unveil the dynamical properties of this flow. The main targeted application domains concern Oceanography and Meteorology. The project consortium gathers the INRIA research groups: FLUMINANCE (leader), CLIME, IPSO, and MOISE. The group of the “Laboratoire de Météorologie Dynamique” located at the ENS Paris, the IFREMER-CERSAT group located at Brest and the METEOFRANCE GMAP group in Toulouse.

6.5.4. Brittany concil ARED IMAGEO:
Participants: Cédric Herzet, Etienne Mémin, Véronique Souchaud.
duration 36 months. This project of the Brittany concil, which finances the PhD thesis of Véronique Souchaud, aims at studying methods for the estimation of reduced order modeling of fluid flows evolution laws from image sequences. The goal consists here at defining the estimation of a reduced basis describing the flow evolution as a motion estimation problem.

6.6. International Initiatives
6.6.1. INRIA Associate Teams
6.6.1.1. HURACAN
Title: Analysis and control of fluid flows from image sequences
INRIA principal investigator: Etienne Memin
International Partner:
Institution: Universidad de Buenos Aires (Argentina)
Laboratory: Fluid Dynamics Laboratory
International Partner:
Institution: IRSTEA (France)
Duration: 2010 - 2012
See also: http://huracan.inria.fr
The HURACAN associated team is centered on the analysis and the control of fluid flows from image sequences. The research objectives of this team are organized into two distinct work axes. The first one aims at defining and studying visual servoing techniques for fluid flows control. In addition to the definition of efficient visual servoing schemes this axis of work gathers research issues related to fluid flows velocity measurement from images and to flows excitation through plasma actuators. The second research axis focuses on the coupling between large scales representations of geo-physical flows and image data. More precisely, it aims at studying means to define directly from the image sequences the small scales terms of the dynamics. This research axis includes the study of coupling models and data defined at different scales, problems of multiscale velocities estimation respecting turbulence phenomenological laws and issues of experimental validation.
6. New Results

6.1. Reconstruction

Participants: Panagiotis Koutsourakis, Helene Langet, Loic Simon, Olivier Teboul, Gilles Fleury, Elisabeth Lahalle, Yves Trousset, Cyril Riddell, Nikos Paragios.

- **Image-based Procedural Modeling of Urban environments**: In [20] we develop a multiple hypotheses testing algorithm for image-based/grammar-driven building modeling. Shape grammars are used to express the variation of the observed architecture. Such a model is coupled with the observations through a maximum likelihood principle where the aim is to maximize the posterior segmentation probability in the image plane given the partition being determined from the grammar derivation. The unknown parameters of the process involve the grammar derivation tree and the associated parameters. Such a mixed continuous/discrete problem is solved through a hill climbing approach that involves joint perturbations in the derivation and parameter space. Promising results demonstrated the potentials of such a formulation for complex Parisian architectures. This idea was further extended in [40] where reinforcement learning was used as optimization principle. 2D Image-based grammar parsing was expressed as a Markov decision process where an agent ought to take actions in an environment so as to maximize some notion of cumulative reward. Performance in particular computational gain over [20] demonstrated the extreme potentials of such a formulation. In order to cope with multi-view geometry, the grammar was further derived to include 3D components and the optimization process was amended to deal with multiple views. An evolutionary computation process (based on consistent mutation and recombination of partial grammar trees) was proposed to fuse image and depth-based information. The use of the Pareto frontier between the two concurrent components of the objective function provides a principle way to determine the optimal solution of the designed objective function.

- **Compressed Sensing Digital Subtraction Rotational Angiography**: in [39] we develop an extension of iterative filtered backprojection method for reconstruction of three-dimensional vascular structures from two spins. Our contribution refers to an approach that improves the reconstruction quality of non-sparse volumes when there exists a sparse combination of these volumes. This is achieved through a joint reconstruction of the mask and contrast volumes via \(l'_1\)-minimization of sparse priors. These ideas were further explored to address three-dimensional reconstruction in interventional radiology in [30] through a regularized extension of the iterative filtered backprojection algorithm. To this end the conventional TV-norm was replaced from a new sparsity constraint that relies on the \(l'_1\)-minimization-norm and the positivity constraint. The use of such a constraint allows for removing most of the subsampling artifacts while preserving background structures.

6.2. Matching/Segmentation

Participants: Haithem Boussaid, Iasonas Kokkinos, Chaohui Wang, Bo Xiang, Ahmet Besbes, Ben Glocker, Nikos Komodakis, Nikos Paragios.

- **Rapid Deformable Part Model Detection**: in [27] we introduce a Branch-and-Bound technique which efficiently finds the most promising configuration of a pictorial structure model given an image. The fastest previously known techniques are linear in the image size; our technique has a best-case complexity that is logarithmic in the image size. When evaluated on standard datasets (Pascal benchmark) our technique gives a 5- to 15-fold speedup. Moreover, when evaluated in the multi-object detection problem our technique’s complexity scales sublinearly also in the number of objects, resulting in 20- to 100- fold speedups when evaluated with 20 object categories.
• **Segmentation with Deformable Graph-based Priors**: in [22] we have introduced a novel formulation to address deformable segmentation using graph-based priors while being able to handle partial-correspondences. Segmentation was formulated as a matching task, where candidate correspondences were determined using boosting, and the assignment problem was solved using MAP inference constrained by a graph-based deformable prior. The notion of missing/erroneous correspondences was introduced in the process leading to state-of-the-art results once compared with prior art in the field. The same prior was used in the context of the segmentation of tagging MR heart images [37]. The main contribution of this paper was the exact estimation of the region-based probability likelihood within a pair-wise MRF through the use of Stokes theorem and integral images.

• **Deformable Model-based 3D reconstruction**: in [23] we introduce a model-based optimization approach to the 3D reconstruction of Femur images using a small set of low-dose X-Ray images. We use a parametric deformable model of the Femur surface and fit it to the acquired data by optimizing its parameters. We incorporate in our optimization criterion multiple aspects of the problem, namely the 3D surface- to 2D plane projection, region-based statistics, and edge-based terms. Our evaluation includes both in vitro and in vivo experiments, where our method is shown to yield promising results, while alleviating the need for time demanding, manual annotations.

• **Pose-invariant Higher Order Graph-based Priors**: in [36] we have introduced a novel method for 3D model inference from 2D images in the absence of camera pose parameters. The method exploits higher (fourth) order priors, which alleviate the need of the estimation of the camera parameters. Furthermore, the proposed formulation couples 3D model inference with 2D correspondences and results on a single shot solution for both problems in the absence of knowledge of the observer internal and external parameters.

### 6.3. Fusion/Registration

**Participants:** Stavros Alchatzidis, Nicolas Honnorat, Fabrice Michel, Aristeidis Sotiras, Chaohui Wang, Alex Bronstein, Michael Bronstein, Christos Davatzikos, Ben Glocker, Nikos Komodakis, Yangming Ou, Dimitris Samaras, Regis Vaillant, Yun Zeng, Nikos Paragios.

• **Intrinsic Dense 3D Surface Matching**: in [38] a probabilistic tracking framework for registering two 3D shape that relies on accurate correspondences between all points across the two frames was proposed. The definition of the matching cost is done using the “uniformization” theory that is combined with regularization terms that enforce spatial and temporal motion consistencies, into a maximum a posteriori (MAP) problem which we approximate using a Markov Random Field (MRF).

• **Optimal Linear Registration**: in [26] we proposed a novel formulation to address linear registration of volumetric images (translation, rotation and scale) that guarantees the optimality of the obtained solution. This was achieved through the approximation of the volumetric data using a sparse representation and the expression of the registration criterion in the form of a difference of convex functions. Cutting plain algorithms in the high-dimensional space were used to provide the optimal solution of the registration problem.

• **Quasi-real Time Registration**: in [21] we proposed a novel message-passing based optimization method to for pair-wise Markov Random Fields models and their applications in medical imaging and computer vision. Such a method was integrated to the deformable registration paradigm introduced in [12]. Such an optimization framework was combined with efficient use of modern architectures (Graphics Processing Units) leading to a speed up of at least one order of magnitude with respect to [12] making quasi real-time deformable registration feasible.

• **Metric Learning**: in [31] we extend prior work on similarity sensitive hashing to address multimodal 3D registration. The method consists of combining invariant to translation/rotation/scale features defined at the Gabor space with a machine learning/boosting method that aims to projection corresponding visual patterns to binary vectors with minimal Hamming distance while maximizing the distance between no corresponding samples.
• **Symmetric Deformable Fusion**: in [9] a novel graph-based formulation combining image and geometric terms was proposed for deformable registration. The method aimed at constraining iconic registration using a set of landmark correspondences that are sparse, do not inherit redundancy and are symmetric. The central idea was to simultaneously deform the target and the source image using two symmetric flows such that the similarity criterion is reaching its lowest potential. This was achieved through the use of composite symmetric deformation fields. This formulation was expressed as a graph-based optimization problem leading to promising experimental results.

• **Deformable registration of gene expression data**: in [28] the combined iconic/geometric registration framework introduced in [9] was extended to deal with gene expression data. Similarity Sensitive Hashing was used to establish costs for landmark correspondences, and a graph-based formulations with unknowns the deformation vectors was adopted for the objective function. Such an idea was extended to deal with combined segmentation/registration approach through an atlas in [29] where subdivision surfaces were considered to represent the deformation grid.

### 6.4. Physiological Modeling & Spatio-Temporal Analysis

**Participants**: Nicolas Honnorat, Sarah Parisot, Stephane Chemouny, Hugues Dufaut, Regis Vaillant, Nikos Paragios.

• **Low Gliomas Brain Map**: in [33] we introduce a graph-based modeling approach towards spatial position interpretation of low gliomas brain tumors. This was achieved through unsupervised clustering from exemplars, where spatial and geometric proximity of tumors were used to determine the strength connectivity of a graph. Towards automatic estimation of the lowest rank graph that is able to express the observed variation of tumors, an LP problem was solved that determines automatically the number of clusters and their centers while associating the training exemplars with them. Promising results that are well aligned with observations from neuro-sciences demonstrate the potentials of the proposed formulation.

• **Coupled Iconic/Geometric Spatio-temporal Segmentation**: in [25] we have introduced a combined elongated structures segmentation/tracking approach that was based on a two-layer graphical model. The image layer was exploiting the visual space and was seeking to minimize a data-driven cost while the geometric layers was seeking to establish temporal correspondences of the deforming structure. These two layers were coupled through a common set of variables acting on the deformation of the control points representing the elongated structure. Guide-wire segmentation [24] and tracking in low signal-to-noise ratio interventional images demonstrated the extreme potentials of our approach.
5. New Results

5.1. Qualitative modeling, simulation, analysis, and verification of gene regulatory networks

Genetic Network Analyzer (GNA) is a tool for the qualitative modeling and simulation of the dynamics of gene regulatory networks by means of PL models, as described in Section 4.1. GNA has been integrated with the other bioinformatics tools distributed by Genostar (http://www.genostar.com/). Version 8.3 of GNA was released by IBIS and Genostar this year. This version is an update of version 8.0, deposited at the Agence pour la Protection des Programmes (APP). Some bugs have been corrected in the new version and the program has been adapted to the latest versions of Java and the software platform of Genostar. A book chapter describing the current version of GNA has been published in a volume on the modeling of bacterial molecular networks [15]. The chapter is a tutorial illustrating the practical use of recent functionalities of GNA like the network editor and the formal verification module by means of an example network in *E. coli* (see also [14]).

Notwithstanding the above improvements of the software, most of our efforts in the past year have gone into applications in collaboration with users of GNA. For example, Delphine Ropers has worked with several groups at IST Lisbon on the modeling of the FLR1 network in yeast, resulting in a paper in *IET Systems Biology* [8]. The paper reports on the qualitative modelling and simulation of the transcriptional regulatory network controlling the response of the model eukaryote *Saccharomyces cerevisiae* to the agricultural fungicide mancozeb. The model has allowed the analysis of the regulation level and activity of the components of the mancozeb-induced network controlling the transcriptional activation of FLR1. This gene is proposed to confer multidrug resistance to the cell through its putative role as a drug efflux pump. Formal verification analysis of the network allowed us to confront model predictions with experimental data and to assess the model robustness to parameter ordering and gene deletion. This analysis led to a better understanding of the mechanisms regulating the response of FLR1 to mancozeb and confirmed the need for a new transcription factor to account for the full transcriptional activation of the gene YAP1. The result is a model of the response of FLR1 to mancozeb, permitting a quick and cost-effective test of hypotheses prior to experimental validation.

As another example of the use of GNA, Hidde de Jong has contributed to the modeling of the TOL system in *Pseudomonas putida*, carried out at the Spanish National Biotechnology Center (CNB). The gene regulatory network of the TOL plasmid pWW0 of the soil bacterium *Pseudomonas putida* mt-2 for catabolism of m-xylene is an archetypal model for environmental biodegradation of aromatic pollutants. Although nearly every metabolic and transcriptional component of this regulatory system is known in detail, the complexity of its architecture is still perplexing. To gain an insight into the inner layout of this network a PL model of the TOL system was implemented, simulated and experimentally validated by measuring the expression of the genes encoding the regulators XylR and XylS when specific portions of the network were activated with selected inducers (m-xylene, o-xylene, 3-methylbenzylalcohol and 3-methylbenzoate). This analysis made sense of the specific regulatory topology on the basis of an unprecedented network motif in the genetic circuit for m-xylene catabolism. The motif appears to ensure a simultaneous expression of the upper and lower segments of the m-xylene catabolic route that would be difficult to bring about with a standard substrate-responsive single promoter. Furthermore, it is plausible that the motif helps to avoid biochemical conflicts between competing plasmid-encoded and chromosomally-encoded pathways in this bacterium. The analysis of the TOL system has been published in *BMC Systems Biology* [11].
5.2. Experimental mapping of gene regulatory networks in bacteria

The use of fluorescent and luminescent reporter genes allows real-time monitoring of gene expression, both at the level of individual cells and cell populations (Section 3.3). In order to fully exploit this technology, we need methods to rapidly construct reporter genes, both on plasmids and on the chromosome, mathematical models to infer biologically relevant quantities from the primary data, and computer tools to achieve this in an efficient and user-friendly manner. For instance, in a typical microplate experiment, 96 cultures are followed in parallel, over several hours, resulting in 10,000-100,000 measurements of absorbance and fluorescence and luminescence intensities. Over the past few years, we put into place an experimental platform and data analysis software, notably the WELLEADER program (Section 4.2), to allow biologists to make the most of the information contained in reporter gene expression data. Several improvements of the platform for measuring gene expression are the subject of ongoing work, including a novel method for efficiently cloning reporter gene constructions on the chromosome of *E. coli*.

These tools are actually used in a series of studies directed at the experimental mapping of gene regulatory networks in *E. coli*. One example, carried out in the framework of the PhD thesis of Guillaume Baptist, is the development of a new screening methodology for identifying all genes that control the expression of a target gene through genetic or metabolic interactions. The screen combines mutant libraries with luciferase reporter constructs. Instead of a static picture of gene expression, this method allows dynamical monitoring in different environmental conditions. Mutants with interesting phenotypes can thus be selected based on multiple criteria, and the expression dynamics of the target gene can be extensively characterized. The method has been applied to the identification of the direct and indirect regulators of the gene *acs* in *Escherichia coli*. We confirmed known genetic regulators of the gene and identified new regulatory influences, many of which involve metabolic intermediates or metabolic sensing. An analysis of mutants involved in glycolysis and glucose transport demonstrates that the classical model of catabolite repression in *E. coli* needs to be amended.

A paper describing the above work is currently under revision.

Other examples of on-going work are the analysis of the network involved in motility and sessility and the modulation of the RpoS regulon in *E. coli* by Omaya Dudin and Stephan Lacour, the validation of a model of the network of global regulators of transcription by Sara Berthoumieux and Hidde de Jong, and the analysis of the regulation of cAMP levels in the bacterial cell by Claire Villiers.

5.3. Analysis of metabolic coupling in gene regulatory networks

The regulation of gene expression is tightly interwoven with metabolism and signal transduction. A realistic view of genetic regulatory networks should therefore not only include direct interactions resulting from transcription regulation, but also indirect regulatory interactions mediated by metabolic effectors and signaling molecules. We coined the term metabolic coupling to denote these indirect interactions mediated by metabolism. Ignoring metabolic coupling during the analysis of the network dynamics may lead crucial feedback loops to be missed.

In previous work, published in *PLoS Computational Biology* in 2010, we showed how indirect interactions arising from metabolic coupling can be derived from a model of the underlying biochemical reaction network. We applied this approach to the carbon assimilation network in *Escherichia coli* investigating how the structural properties of the network are modified by the inclusion of metabolic interactions. Our results showed that the derived gene regulatory network is densely connected, contrary to what is usually assumed. Moreover, we found that the signs of the indirect interactions are largely fixed by the direction of metabolic fluxes, independently of specific parameter values and rate laws, and that a change in flux direction may invert the sign of indirect interactions. This leads to a feedback structure that is at the same time robust to changes in the kinetic properties of enzymes and that has the flexibility to accommodate radical changes in the environment.

It remains an open question, however, to which extent the indirect interactions induced by metabolic coupling affect the dynamics of the system. This is a key issue for understanding the relative contributions of the regulation of gene expression and metabolism during the adaptation of the cell to changes in its environment.

In collaboration with Valentina Baldazzi, formerly post-doctoral fellow in IBIS and now research scientist...
at INRA (Avignon), we have carried out a dynamic analysis by developing a qualitative PL model of the gene regulatory network, including both the direct and indirect interactions. We previously showed, in a paper published in the *IEEE/ACM Transactions on Computational Biology and Bioinformatics* this year, that PL models provide a good approximation of the direct and indirect interactions occurring in gene regulation [10].

In order to obtain a clearer view of the dynamic role of metabolic coupling in the adaptation of gene expression, we developed several qualitative models corresponding to a network topology including all, some, or none of the indirect interactions. The dynamical properties of the models were analyzed and compared with available experimental data using the computer tool GNA (Section 4.1). In particular, we compared the steady-state concentrations of enzymes and transcription regulators during growth on glucose and acetate, as well as the dynamic response of gene expression to the exhaustion of glucose and the subsequent assimilation of acetate. We find significant differences between the dynamics of the system in the absence and presence of metabolic coupling. This confirms that indirect interactions are essential for correctly reproducing the observed adaptation of gene expression to a change in carbon source. Our work thus underlines the importance of metabolic coupling in gene regulatory networks, and shows that such indirect interactions cannot be neglected when studying the adaptation of an organism to changes in its environment. A short, preliminary paper on this work was presented at an invited session of the 18th IFAC World Congress held in Milano [12] and a long paper has been accepted for publication in the *Journal of Theoretical Biology* [2].

### 5.4. Parameter estimation for kinetic models of carbon metabolism in bacteria

Kinetic models capture the dynamics of the large and complex networks of biochemical reactions that endow bacteria with the capacity to adapt their functioning to changes in the environment. In comparison with the qualitative PL models described in Sections 5.1 and 5.3, these more general classes of ODE models are intended to provide a quantitative description of the network dynamics, both on the genetic and metabolic level. New experimental techniques have led to the accumulation of large amounts of data, such as time-course measurements of metabolite, mRNA and protein concentrations and measurements of metabolic fluxes under different growth conditions. However, the estimation of parameter values in the kinetic models from these data remains particularly challenging in biology, mostly because of incomplete knowledge of the molecular mechanisms, noisy, indirect, heterogeneous, and partial observations, and the large size of the systems, with dynamics on different time-scales. We have addressed parameter estimation in the context of the analysis of the interactions between metabolism and gene expression in carbon metabolism in *E. coli*.

In collaboration with Matteo Brilli and Daniel Kahn (INRA and Université Claude Bernard in Lyon), we have developed an approximate model of central metabolism of *E. coli*, using so-called linlog functions to approximately describe the rates of the enzymatic reactions. More precisely, linlog models describe metabolic kinetics by means of a linear model of the logarithms of metabolite concentrations. We have used metabolome and transcriptome data sets from the literature to estimate the parameters of the linlog models, a task in principle greatly simplified by the mathematical form of the latter. However, a major problem encountered during parameter estimation was the occurrence of missing data, due to experimental problems or instrument failures. In the framework of her PhD thesis, Sara Berthoumieux has addressed the missing-data problem by developing an iterative parameter estimation approach based on an Expectation-Maximization (EM) procedure. This approach adapted from the statistical literature has the advantage of being well-defined analytically and applicable to other kinds of linear regression problems with missing data. It has been tested on simulations experiments with missing data and performs well compared to basic and advanced regression methods.

On the biological side, we have applied the method to a linlog model of central metabolism in *Escherichia coli*, consisting of some 23 variables. We estimated the 100 parameters of this model from a high-throughput dataset published in the literature. The data consists of measurements of metabolic fluxes and metabolite and enzyme levels in glucose-limited chemostat under 29 different conditions such as wild-type strain and single-gene mutant strains or different dilution rates. Standard linear regression is difficult to apply in this case due to missing data, which disqualifies for 7 reactions too many datapoints, leaving a dataset of size inferior to the
number of parameters to estimate. Application of our approach allows one to compute reasonable estimates for most of the identifiable model parameters even when regression is inapplicable. The method and its application to the linlog model of central metabolism in *E. coli* are the subject of a paper accepted for the ISMB/ECCB conference this year and published in a special issue of *Bioinformatics* [3]. Sara Berthoumieux received the Ian Lawson Van Toch Memorial Award for outstanding student paper at ISMB/ECCB. In the continuation of this work, we are currently preparing for submission a journal paper on the identifiability of linlog models.

A second line of work is based on the use of classical kinetic models that are, in comparison with the above-mentioned linlog models, much reduced in scope (the focus is on the metabolic and genetic regulation of the glycolysis pathway) and granularity (individual reactions are lumped together). The models, developed by Delphine Ropers, have been calibrated using experimental data from the experimental part of the IBIS group for the gene expression measurements and the group of Jean-Charles Portais at INSA in Toulouse for the measurements of metabolism. The model with the estimated parameter values is currently being tested and used to understand some key mechanisms in the adaptation of *E. coli* to the exhaustion of glucose. The PhD theses of Stéphane Pinhal and Valentin Zulkower, which started at the end of this year, will further develop these research directions.

### 5.5. Structural identification of gene regulatory networks

In general, structural identification of genetic regulatory networks involves fitting appropriate network structures and parameters to the data. While modern measurement techniques such as reporter gene systems provide data of ever-increasing quality, the problem remains challenging because exploring all possible network structures in the search of the best fitting model is prohibitive.

In order to address the structural identification problem, Eugenio Cinquemani proposed in collaboration with the Automatic Control Lab at ETH Zürich (Switzerland) and the Computer Engineering & Systems Science Department of the University of Pavia (Italy), an ODE modelling framework which we refer to as models with unate-like structure. In Boolean network modeling, unate functions are argued to capture virtually all observable interactions in gene regulatory networks. In our quantitative framework, unate logics are encoded in the structure of the nonlinear synthesis rates of the network proteins. This framework allows us to integrate *a-priori* information on the most likely network structures, and the models enjoy monotonicity properties that can be exploited to simplify the identification task.

As described in previous work, published in *Bioinformatics* in 2010, the key idea is to divide the identification process into two steps. In the first step, different monotonicity properties of different model structures are exploited to discard those structures whose property is falsified by the observed data points (time-lapse protein concentrations and synthesis rates). In the second step, the parameters of the model structures not discarded in the first step are fitted to the data in the search of the simplest structure explaining the data with sufficient accuracy. The procedure was validated on challenging data from the literature.

On the methodological side, in the context of the same international collaboration, the identification approach has been further developed. For important subclasses of unate models, larger sets of network structures can now be discarded in the hypothesis falsification step, based on additional properties other than monotonicity (namely quasi-convexity). These improvements have been presented at the 18th IFAC World Congress held in Milan, and are reported in a journal paper that has been accepted for publication in the *International Journal of Robust and Nonlinear Control*, in a special issue on system identification for biological systems. In the framework of the PhD thesis of Diana Stefan, in collaboration with Eugenio Cinquemani, Stephan Lacour and Omaya Dudin, the method is now being applied to experimental data produced within IBIS for the study of the gene network regulating motility of *E. coli* bacteria.

Woei-Fuh Wang, who defended her PhD thesis carried out under the supervision of Johannes Geiselmann and Chung-Ming Chen in December 2011, addressed a different problem in the structural identification of gene regulatory networks. The inference of the network topology and regulatory mechanisms is complicated by the fact that we usually do not know all relevant genes that need to be taken into account for explaining the observed expression patterns. The aim of the thesis was to detect the presence of such “missing genes”, as
well as their regulatory roles and expression patterns. Using a well-known class of simplified kinetic models, based on power-law approximations of synthesis rate functions, an inference algorithm was developed. The algorithms are based on factor analysis, a well-developed multivariate statistical analysis approach that is used to investigate unknown, underlying features of a dataset, as well as independent component analysis. The proposed method of inferring the expression profile of a missing gene and connecting it to a known network structure has been applied to artificial networks, as well as a real network studied within IBIS: the \textit{acs} regulatory network in \textit{Escherichia coli}.

### 5.6. Stochastic modeling and identification of gene regulatory networks in bacteria

At the single cell level, the processes that govern gene expression are often better described by stochastic models. Modern techniques for the real-time monitoring of gene expression in single cells enable one to apply stochastic modelling to study the origins and consequences of random noise in response to various environmental stresses, and the emergence of phenotypic variability. The potential impact of single-cell stochastic analysis and modelling is tremendous, ranging from a better comprehension of the biochemical regulatory mechanisms underlying life, to the development of new strategies for the control of bacterial populations and even of single cells, with applications in for example biotechnology and medicine.

In the literature, much effort has been devoted to the analysis of stochastic gene expression models derived from biochemical kinetics and specific knowledge of the systems at hand. Less effort has been dedicated to developing general methods for inferring unknown parameter values of these stochastic models from single-cell experimental data. While some strategies have been proposed in the recent literature, no method of general applicability exists. IBIS recently started a new line of research dedicated to the study of stochastic modelling and identification of gene regulatory networks in single cells. This work, coordinated by Eugenio Cinquemani, focuses on simple network modules in bacterial cells. Our reference system is the regulation of the onset of fructose uptake in \textit{E. coli} upon depletion of glucose.

In the past year we developed a working method for the estimation of unknown network parameters of a simple stochastic model of the fructose uptake process. The method was tested on simulated data and applied with success to time-lapse fluorescence microscopy data acquired by Guillaume Baptist. This application involved the development by Michel Page of a microscopy data processing program based on a customization of the freely accessible Matlab tool CellTracer. Preliminary results were presented in the poster session of the Conference on Stochastic Systems Biology held in Monte Verità (Switzerland). The work is currently being extended in preparation for a journal publication. A generalization of the method and the investigation of alternative stochastic modelling and identification methodologies are being pursued in parallel. Other ongoing work concerns the study of noise propagation in gene regulatory networks in collaboration with Irina Mihalcescu (Université Joseph Fourier).

### 5.7. Control of regulatory networks in bacteria

While systems biology is primarily concerned with natural systems shaped by evolution, synthetic biology opens up a new generation of fundamental research by trying to redesign natural systems or create novel systems from scratch. Mathematical modeling and analysis are essential components of synthetic biology, as they help understanding the consequences of (changes in) the network of interactions on the dynamical behavior of the system. More specifically, a model can be a powerful tool for the control and regulation of the system towards a desired goal.

Within the projects ColAge and GeMCo (Section [5.1](#)), we attempt to control one of the fundamental physiological properties of bacterial cells, their growth rate. In particular, in order to control the growth rate, we propose to focus on the gene expression machinery of \textit{E. coli}, whose activity is controlled by a complex regulatory network with many components and intertwined feedback loops. Delphine Ropers is developing models of the gene expression machinery and Jérome Izard, in the context of his PhD thesis, is rewiring part of the network to enable control of the network dynamics. The results on these projects are currently being prepared for publication.
6. New Results

6.1. Modeling and estimation in biomechanics

6.1.1. Patient-Specific Electromechanical Models of the Heart for the Prediction of Pacing Acute Effects in CRT: a Preliminary Clinical Validation

In collaboration with Project-Team Asclepios from INRIA Sophia-Antipolis-Méditerranée and the Division of Imaging Sciences of St Thomas’ Hospital, King’s College London we demonstrated the benefits of using patient-specific electromechanical models of the heart for the prediction of pacing acute effects in CRT, see [5].

Cardiac resynchronisation therapy (CRT) is an effective treatment for patients with congestive heart failure and a wide QRS complex. However, up to 30% of patients are non-responders to therapy in terms of exercise capacity or left ventricular reverse remodelling. A number of controversies still remain surrounding patient selection, targeted lead implantation and optimisation of this important treatment. The development of biophysical models to predict the response to CRT represents a potential strategy to address these issues. We present how the personalisation of an electromechanical model of the myocardium can predict the acute haemodynamic changes associated with CRT. In order to introduce such an approach as a clinical application, we needed to design models that can be individualised from images and electrophysiological mapping of the left ventricle. We performed the personalisation of the anatomy, the electrophysiology, the kinematics and the mechanics. The acute effects of pacing on pressure development were predicted with the in silico model for several pacing conditions on two patients, achieving good agreement with invasive haemodynamic measurements: the mean error on dP/dtmax is 47.5 $\pm$ 35 mmHg.s$^{-1}$, less than 5% error.

6.1.2. Estimation of tissue contractility from cardiac MRI using a biomechanical heart model

Participants: Radomir Chabiniok, Dominique Chapelle, Alexandre Imperiale, Philippe Moireau.

In collaboration with P.-F. Lesault, A. Rahmouni and J.-F. Deux from Hospital H. Mondor, Créteil we proposed and assessed an estimation procedure – based on data assimilation principles – well-suited to obtain some regional values of key biophysical parameters in a beating heart model, using actual Cine-MR images, see [8], [1]. The motivation is twofold: (1) to provide an automatic tool for personalizing the characteristics of a cardiac model in order to achieve predictivity in patient-specific modeling, and (2) to obtain some useful information for diagnosis purposes in the estimated quantities themselves. In order to assess the global methodology we specifically devised an animal experiment in which a controlled infarct was produced and data acquired before and after infarction, with an estimation of regional tissue contractility – a key parameter directly affected by the pathology – performed for every measured stage. After performing a preliminary assessment of our proposed methodology using synthetic data, we then demonstrate a full-scale application by first estimating contractility values associated with 6 regions based on the AHA subdivision, before running a more detailed estimation using the actual AHA segments. The estimation results are assessed by comparison with the medical knowledge of the specific infarct, and with late enhancement MR images. We discuss their accuracy at the various subdivision levels, in the light of the inherent modeling limitations and of the intrinsic information contents featured in the data.

We are now working on improving these results by the use of Tagged-MRI, see [9]. The first approach consists in assuming that the image data is processed in the form of deforming tag planes, which we employ to obtain a discrepancy between the model and the data by computing distances to these surfaces. We assess our procedure using synthetic measurements produced with a model representing an infarcted heart as observed in an animal experiment, and the estimation results are found to be of superior accuracy compared to assimilation based on segmented endo- and epicardium surfaces. Then we extend this strategy to tagged lines instead of tagged planes or even directly with apparent displacements extracted from tagged images by optical flow methods.
### 6.1.3. Convergence of observers based on partial field measurements for the wave equation

**Participants:** Dominique Chapelle, Nicolae Cîndea, Maya de Buhan, Philippe Moireau.

We analyzed an observer strategy based on partial – i.e. in a subdomain – measurements of the solution of a wave equation, in order to compensate for unknown initial conditions, see [17], [18]. We proved the exponential convergence of this observer under a non-standard observability condition, whereas using measurements of the time-derivative of the solution would lead to a standard observability condition arising in stabilization and exact controllability. Nevertheless, we directly related our specific condition to the classical geometric control condition. This results justify in a linear framework the use of our observer-based filter in cardiac modeling.

### 6.1.4. Reduced nonlinear optimal filtering

**Participants:** Dominique Chapelle, Akos Matszangosz, Philippe Moireau.

We investigated some issues pertaining to reduced-order considerations in nonlinear optimal filtering. Classically, optimal filtering formulations lead to Hamilton-Jacobi-Bellman (HJB) equations posed in the complete “space of uncertainty”, namely, including the state space. This makes such methods generally untractable for PDE-based models. However, under certain assumptions pertaining to reduced uncertainties we can transform the HJB equations into a form posed in the reduced uncertainty space, and with only time derivatives involved. This form can be solved for – including with PDEs – provided this reduced space is of limited size, and then gives a reference “optimal” method to which other filtering procedures can be compared. The subject of Akos Matszangosz’ internship (from “Ecole des Mines de Paris”, duration 4 months) was to design and perform an adequate implementation of this reduced-order optimal filter, based on a sparse-grid discretization of the uncertainty space. In addition, we are currently working on discrete-time optimal filtering formulations, which are distinct – and preferable in principle – to discretizing the continuous forms.

### 6.2. Asymptotic and multiscale modeling

#### 6.2.1. Modeling and simulation of multi-layers mechanical structures

**Participants:** Marina Vidrascu, Sofiene Hendili.

The collaboration with Françoise Krasucki (Montpellier University) and Giuseppe Geymonat (Ecole polytechnique) on the modeling of 3D materials connected by stiff interfaces continues within the Epsilon ANR project (Domain decomposition and multi-scale computations of singularities in mechanical structures 7.1.1). In the framework of matched asymptotic expansions we introduced a new effective and robust method to approximate the behavior of a structure containing a thin layer with periodically distributed heterogeneities (holes, rigid bodies...), see [3], [10], [11].

#### 6.2.2. Multi-scale modeling and simulation of rubber

**Participants:** Maya de Buhan, Marina Vidrascu, Antoine Gloria [SIMPAF], François Lequeux [ESPCI], Patrick Le Tallec [Ecole Polytechnique].

In collaboration with A. Gloria (project-team SIMPAF) and P. Le Tallec (Ecole Polytechnique), we are currently working on a multiscale model for rubber based on the statistical physics description of a network of polymer chains. The numerical simulation of the model has been addressed within the ARC Disco using the Sheldon software. Comparisons with mechanical experiments are promising, and related inverse problems have been addressed in the post-doc of M. de Buhan. Two publications are in preparation.

This work is supported by the ARC DISCO (7.1.2).

### 6.3. Other topics

#### 6.3.1. Numerical analysis of POD-based Galerkin approximations

**Participants:** Dominique Chapelle, Asven Gariah, Philippe Moireau, Jacques Sainte-Marie.
In [2], we proposed a numerical analysis of Proper Orthogonal Decomposition (POD) model reductions in which a priori error estimates are expressed in terms of the projection errors that are controlled in the construction of POD bases. These error estimates are derived for generic parabolic evolution PDEs, including with non-linear Lipschitz right-hand sides, and for wave-like equations. A specific projection continuity norm appears in the estimates and — whereas a general uniform continuity bound seems out of reach — we prove that such a bound holds in a variety of Galerkin bases choices. Furthermore, we directly numerically assess this bound — and the effectiveness of the POD approach altogether — for test problems of the type considered in the numerical analysis, and also for more complex equations. Namely, the numerical assessment includes a parabolic equation with super-linear reaction terms, inspired from the FitzHugh-Nagumo electrophysiology model, and a 3D biomechanical heart model. This shows that the effectiveness established for the simpler models is also achieved in the reduced-order simulation of these highly complex systems.

This work is now being continued in order to handle parameter-dependent models, and thence estimation problems.

6.3.2. Sail modeling

Participants: Dominique Chapelle, Daniele Trimarchi, Marina Vîdrascu.

A dynamic Finite Element method — based on non-linear MITC shell finite elements implemented in the MITCNL software — has been proposed and assessed for the analysis of downwind sail-type structures, see [12]. The main purpose was here to investigate the development of wrinkling, a phenomenon commonly observed in practice for such structures. Considering the wrinkling in this type of analysis is of great interest, since wrinkling affects the stress distribution in the fabric. Further developments primarily regard various refinements of the model, in order to represent some even more realistic sail configurations such as with non-isotropic material models, corner reinforced zones, and cable boundary conditions. Of course, another very important perspective — and work in progress, indeed — concerns the use of such sail models coupled with the wind flow in a fluid-structure interaction framework.
6. New Results

6.1. Inverse Problems

6.1.1. Reconstruction of an elastic scatterer immersed in a homogeneous fluid

Participants: Hélène Barucq, Rabia Djellouli, Élodie Estecahandy.

The determination of the shape of an obstacle from its effects on known acoustic or electromagnetic waves is an important problem in many technologies such as sonar, radar, geophysical exploration, medical imaging and nondestructive testing. This inverse obstacle problem (IOP) is difficult to solve, especially from a numerical viewpoint, because it is ill-posed and nonlinear. Its investigation requires as a prerequisite the fundamental understanding of the theory for the associated direct scattering problem, and the mastery of the corresponding numerical solution methods.

In this work, we are interested in retrieving the shape of an elastic obstacle from the knowledge of some scattered far-field patterns, and assuming certain characteristics of the surface of the obstacle. The corresponding direct elasto-acoustic scattering problem consists in the scattering of time-harmonic acoustic waves by an elastic obstacle $\Omega$ embedded in a homogeneous medium $\Omega_f$, that can be formulated as follows:

\[
\begin{align*}
\Delta p + (\omega^2/c_f^2) p &= 0 & \text{in } \Omega_f \\
\nabla \cdot \sigma(u) + \omega^2 \rho_s u &= 0 & \text{in } \Omega_s \\
\omega^2 \rho_f u \cdot n &= \partial p/\partial n + \partial e^i(\omega/c_f) x d/\partial n & \text{on } \Gamma' \\
\sigma(u)n &= -pn - e^i(\omega/c_f) x d_n & \text{on } \Gamma' \\
\lim_{r \to +\infty} r (\partial p/\partial r - i (\omega/c_f) p) &= 0 & \text{on } \Gamma'
\end{align*}
\]

where $p$ is the fluid pressure in $\Omega_f$ whereas $u$ is the displacement field in $\Omega_s$, and $\sigma(u)$ represents the stress tensor of the elastic material.

This boundary value problem has been investigated mathematically and results pertaining to the existence, uniqueness and regularity can be found in [65] and the references therein, among others. We propose a solution methodology based on a regularized Newton-type method for solving the IOP. The proposed method is an extension of the regularized Newton algorithm developed for solving the case where only Helmholtz equation is involved, that is the acoustic case by impenetrable scatterers [55]. The direct elasto-acoustic scattering problem defines an operator $F : \Gamma' \to p_{\infty}$ which maps the boundary $\Gamma'$ of the scatterer $\Omega'$ onto the far-field pattern $p_{\infty}$. Hence, given one or several measured far-field patterns $\tilde{p}_{\infty}(\tilde{x})$, corresponding to one or several given directions $d$ and wavenumbers $k$, one can formulate IOPs as follows:

Find a shape $\Gamma'$ such that $F(\Gamma')(\tilde{x}) = \tilde{p}_{\infty}(\tilde{x}); \quad \tilde{x} \in S^1$.

We propose a solution methodology based on a regularized Newton-type method to solve this inverse obstacle problem. At each Newton iteration, we solve the forward problem using a finite element solver based on discontinuous Galerkin approximations, and equipped with high-order absorbing boundary conditions. We have first characterized the Fréchet derivatives of the scattered field. They are solution to the same boundary value problem as the direct problem with other transmission conditions. This work has been presented both in FACM11 and in WAVES 2011. A paper has been submitted.
6.1.2. Seismic data interpretation using the Hough transform and principal component analysis

Participants: M.-G Orozco-del-Castillo, Carlos Ortiz-Aleman, Roland Martin, Rafael Avila-Carrera, Alejandro Rodriguez-Castellanos.

In [29], two novel image processing techniques are applied to detect and delineate complex salt bodies from seismic exploration profiles: Hough transform and principal component analysis (PCA). It is well recognized by the geophysical community that the lack of resolution and poor structural identification in seismic data recorded at sub-salt plays represent severe technical and economical problems. Under such circumstances, seismic interpretation based only on the human-eye is inaccurate. Additionally, petroleum field development decisions and production planning depend on good-quality seismic images that generally are not feasible in salt tectonics areas. In spite of this, morphological erosion, region growing and, especially, a generalization of the Hough transform (closely related to the Radon transform) are applied to build parabolic shapes that are useful in the idealization and recognition of salt domes from 2D seismic profiles. In a similar way, PCA is also used to identify shapes associated with complex salt bodies in seismic profiles extracted from 3D seismic data. To show the validity of the new set of seismic results, comparisons between both image processing techniques are exhibited. It is remarkable that the main contribution of this work is oriented in providing the seismic interpreters with new semi-automatic computational tools. The novel image processing approaches presented here may be helpful in the identification of diapirs and other complex geological features from seismic images. Conceivably, in the near future, a new branch of seismic attributes could be recognized by geoscientists and engineers based on the encouraging results reported here.

6.1.3. Gravimetry Inversion

Participants: Roland Martin, Dimitri Komatitsch, Mark Jessel, Stéphane Perroud, Vadim Monteiller.

In order to improve the subsoil images in regions which are not well covered by a dense seismic array or can not be well retrieved by using seismic imaging techniques alone (salty dome regions like in the Gulf of Mexico for instance), we have been developing new gravity imaging techniques using supercomputing. In regions like Ghana or the Chicxulub crater located in Yucatan plate (Mexico), 3D sensitivity kernels are calculated for gravity potential data sets measured over 2000 up to 10000 locations randomly distributed in space. The density anomaly computational domain covers a 250 km \times 250 km \times 20 km volume in these two regions. For instance for the Ghana region two resolutions are taken : 1 point each 2 kms in the horizontal plane and 200 m in the vertical direction in the less accurate configuration and one point each 500 m in the horizontal plane and one point each 100 m in the vertical direction for the most accurate configuration. The gravity anomalies are inverted using an optimized least-square method applied to a sensitivity kernel of up to 10^10 elements. The least-square method using a L^2-norm or L^1-norm has been implemented on hybrid multi-CPU/multi-GPU Titan machines at CCRT of the French Nuclear Energy Agency. L^1 norm gives us sharper boundaries of the density structures when compared to L^2-norm solutions but these L^1 solutions are obtained at the expense of one order of magnitude in the number of iterations necessary to the inversion. The L^2-norm gives slightly smoother solutions but is much more faster than L^1 norm by at least one order of magnitude in terms of acceleration. The optimized CPU version has allowed us to reduce drastically the computation time form 5 hours on 512 processors to 25 minutes. Furthermore the multi-GPU version has decreased this computational time around 15 minutes. Our collaboration with geologists of IRD (Mark Jessell and Stephane Perroud) has allowed us to determine a realistic a priori model of Ghana with different resolutions in order to obtain more realistic models after inversion. We are still optimizing the multi-GPU code, with the challenging goal of obtaining results lower than 10 minutes and then obtaining an acceleration factor of at least 4 when compared to the optimized multi-CPU inversion code. By now the code is further optimized and more improvements are already in course in terms of better preconditioning, automatic multi-resolution procedure implementation, gravimetry-seismic joint inversion and extension to the global earth imaging. An international article in a scientific peer-reviewed journal is in preparation.
6.2. Modeling

6.2.1. Local approximate DtN exterior boundary condition

Participants: Hélène Barucq, Rabia Djellouli, Anne-Gaëlle Saint-Guirons.

We investigate analytically the asymptotic behavior of high-order spurious prolate spheroidal modes induced by a second-order local approximate DtN absorbing boundary condition (DtN2) when employed for solving high-frequency acoustic scattering problems. We prove that these reflected modes decay exponentially in the high frequency regime. This theoretical result demonstrates the great potential of the considered absorbing boundary condition for solving efficiently exterior high-frequency Helmholtz problems. In addition, this exponential decay proves the superiority of DtN2 over the widely used Bayliss-Gunserburger-Turkel absorbing boundary condition. This work has been accepted for publication in Progress In Electromagnetics Research B. [19].

6.2.2. Non-reflecting boundary condition on ellipsoidal boundary

Participants: Hélène Barucq, Anne-Gaëlle Saint-Guirons, Sébastien Tordeux.

The modeling of wave propagation problems using finite element methods usually requires the truncation of the computational domain around the scatterer of interest. Absorbing boundary condition are classically considered in order to avoid spurious reflections. In this paper, we investigate some properties of the Dirichlet to Neumann map posed on a spheroidal boundary in the context of the Helmholtz equation. We focus on the impedance coefficients defining the DtN condition and we aim at establishing suitable properties in order to propose an accurate numerical method for their computation. Then, we state the well-posedness of the corresponding mixed problem and propose a variational formulation adapted to a finite element discretization. This work has been submitted.

6.2.3. A new modified equation approach for solving the wave equation

Participants: Cyril Agut, Hélène Barucq, Julien Diaz, Florent Ventimiglia, Roland Martin, Dimitri Komatitsch.

The new method involving $p$-harmonic operator described in section 3.2 has been presented in [13]. We have proved the convergence of the scheme and its stability under a CFL condition. Numerical results in one, two and three-dimensional configurations show that this CFL condition is slightly greater than the CFL condition of the second-order Leap-Frog scheme. We have also studied the penalization parameters involved in the new schemes and their influence of the CFL condition. These results are presented in the PhD thesis of Cyril Agut [11]. In the framework of the PhD thesis of Florent Ventimiglia, we are now considering the extension of this technique to the first order formulation of the elastodynamic equations.

6.2.4. Stability Analysis of an Interior Penalty Discontinuous Galerkin Method for the Wave equation

Participants: Cyril Agut, Hélène Barucq, Julien Diaz.

The Interior Penalty Discontinuous Galerkin Method [42], [38], [61] we use in the IPDGIFEM code requires the introduction of a penalty parameter. Except for regular quadrilateral or cubic meshes, the optimal value of this parameter is not explicitly known. Moreover, the condition number of the resulting stiffness matrix is an increasing function of this parameter, but the precise behaviour has not been explicitied neither. We have carried out a theoretical and numerical study of the CFL condition for quadrilateral and cubic meshes, which is presented in the PhD thesis of Cyril Agut [11]. These results were also presented at the peer-reviewed conference Waves 2011 (Vancouver, Canada, July 2011).

6.2.5. Higher Order Absorbing Boundary Conditions for the Wave Equation

Participants: Hélène Barucq, Julien Diaz, Véronique Duprat.
The numerical simulation of wave propagation is generally performed by truncating the propagation medium. Absorbing boundary conditions are then needed. We construct a new family of absorbing boundary conditions from the factorization of the wave equation formulated as a first order system. Using the method of M.E. Taylor, we show that we can generate an infinite number of boundary conditions which can not be obtained via the Niremberg’s factorization method. The conditions can be applied on arbitrarily-shaped surfaces and involve second-order derivatives. We then propose a reduced formulation of the wave equation using an auxiliary unknown which is defined on the regular surface only. The reduced problem allows one to easily include the boundary conditions inside the variational formulation. The corresponding boundary value problem remains well-posed in suitable Hilbert spaces and we give a demonstration in a framework that is suitable to applications. We then study the long-time behavior of the wave field and we show that it tends to 0 as time tends to infinity. This provides a weak stability result that should be completed in the second part of this work. We have then decided to improve the stability result by performing a quantitative study of the energy. We have then shown that the energy is exponentially decaying if the obstacle is star-shaped and the external boundary is convex. This work has been published as INRIA Research Reports \cite{34}, \cite{35} and two papers are submitted. We have next addressed the issue of enriching these ABCs by representing evanescent and damping waves. This has given rise to a work for the Helmholtz equation and we have shown that the enriched ABCs performed better than standard ABCs. The extension to the acoustic wave has led to new conditions involving fractional derivatives. To the best of our knowledge, it is the first time that fractional derivatives have been used for optimizing the performance of ABCs. These new results have been presented in two seminars (University of Bordeaux I and University of Genova) and in two conferences (FACM11 and WAVES 2011). A paper has been published \cite{17} for the case of evanescent waves and two papers are in preparation. All these results are presented in the PhD thesis of Véronique Duprat \cite{12}.

6.2.6. Numerical methods combining local time stepping and mixed hybrid elements for the terrestrial migration

**Participants:** Caroline Baldassari, Hélène Barucq, Henri Calandra [Expert Engineer, TOTAL], Bertrand Denel [Research Engineer, TOTAL], Julien Diaz, Florent Ventimiglia.

In order to justify the use of our code IPDGFEM for the Reverse Time Migration, we have carried out a performance analysis of the Interior Penalty Discontinuous Galerkin method and of the Spectral Element Method. This analysis, which shows that IPDG performs as well as SEM, has been presented in \cite{14}.

Another aspect of the work concerns the design of local time-stepping algorithms. The local-time stepping strategy proposed in \cite{5} allows for high-order time schemes where the time scheme is adapted to the various space step of the mesh. However, when the mesh contains both low-order and high-order cells, this method not allows for the adaptation of the order of the time-scheme to the order of the cells. We have then presented a new local time-stepping algorithm where both the order of the scheme and the time step vary in the different parts of the mesh. This method has been presented in \cite{14} and at the peer-reviewed conferences Waves 2011 (Vancouver, Canada, July 2011) and DD20 (Domain Decompostion, San Diego, USA, February 2011).

The local-time stepping algorithm is not adapted to handle dissipation terms. A method has been proposed in \cite{59}, but it is based on an Adams-Bashworth scheme and it requires the storage of additional unknowns. We can not use this scheme for the simulation of seismic waves in very large heterogeneous domains due to memory limitation. We are now working on the design of alternative schemes which would not require the introduction of the auxiliary unknowns. This one of the topics of the PhD. thesis of Florent Ventimiglia.

6.2.7. Perfectly Matched Layers for the Shallow Water equations

**Participants:** Hélène Barucq, Julien Diaz, Mounir Tlemcani [Assistant Professor, University of Oran, Algeria].

In \cite{45}, we have proposed a new Perfectly Matched Layer for Shallow Water equations. This layer required the computation of an auxiliary variable in the whole computational domain. We are now considering a new strategy, which only requires the computation of the auxiliary variable inside the layer. Moreover, the new methodology seems to be well-adapted to the non-linear shallow water equations. We are now performing numerical tests to confirm this point.
6.2.8. Multiperforated plates in linear acoustics  
**Participants:** Abderrahmane Bendali, M’Barek Fares, Sophie Laurens, Estelle Piot, Sébastien Tordeux.

Acoustic engineers use approximate heuristic models to deal with multiperforated plates in liners and in combustion chambers of turbo-engines. These models were suffering from a lack of mathematical justifications and were consequently difficult to improve. Performing an asymptotic analysis (the small parameter is the radius of the perforations), we have justified these models and proposed some improvement. Our theoretical results have been compared to numerical simulations performed at CERFACS (M’Barek Fares) and to acoustical experiments realized at ONERA (Estelle Piot). Two papers are in preparation.

6.2.9. Asymptotic modeling in electromagnetism  
**Participants:** François Buret, Gabriel Caloz, Monique Dauge, Patrick Dular, Marc Duruflé, Erwan Faou, Laurent Krähenbühl, Victor Péron, Ronan Perrussel, Clair Poignard, Damien Voyer.

The following results rely on several problematics developed in section 3.2, item Asymptotic modeling. We consider in [21] the equations of electromagnetism set on a domain made of a dielectric and a conductor subdomain in a regime where the conductivity is large. Assuming smoothness for the dielectric–conductor interface, relying on recent works we prove that the solution of the Maxwell equations admits a multiscale asymptotic expansion with profile terms rapidly decaying inside the conductor. This skin effect is measured by introducing a skin depth function that turns out to depend on the mean curvature of the boundary of the conductor. We then confirm these asymptotic results by numerical experiments in various axisymmetric configurations. We also investigate numerically the case of a nonsmooth interface, namely a cylindrical conductor.

We derive new thin layer models in electromagnetism, in [22]. We study the behavior of the electromagnetic field in a biological cell modeled by a medium surrounded by a thin layer and embedded in an ambient medium. We derive approximate transmission conditions in order to replace the membrane by these conditions on the boundary of the interior domain. Our approach is essentially geometric and based on a suitable change of variables in the thin layer. Few notions of differential calculus are given in order to obtain the first-order conditions in a simple way, and numerical simulations validate the theoretical results. Asymptotic transmission conditions at any order are given.

We present a numerical treatment of rounded and sharp corners in the modeling of 2D electrostatic fields in [36]. This work deals with numerical techniques to compute electrostatic fields in devices with rounded corners in 2D situations. The approach leads to the solution of two problems: one on the device where rounded corners are replaced by sharp corners and the other on an unbounded domain representing the shape of the rounded corner after an appropriate rescaling. Both problems are solved using different techniques and numerical results are provided to assess the efficiency and the accuracy of the techniques.

6.2.10. Operator Based Upscaling for Discontinuous Galerkin Methods  
**Participants:** Hélène Barucq, Théophile Chaumont, Julien Diaz, Victor Péron.

Realistic numerical simulations of seismic wave propagation are complicated to handle because they must be performed in strongly heterogeneous media. Two different scales must then be taken into account. Indeed, the medium heterogeneities are very small compared to the characteristic dimensions of the propagation medium. To get accurate numerical solutions, engineers are then forced to use meshes that match the finest scale representing the heterogeneities. Meshing the whole domain with the fine grid leads then to huge linear systems and the computational cost of the numerical method is then very high. It would be thus very interesting to dispose of a numerical method allowing to represent the heterogeneities of the medium accurately while computing on a coarse grid. This is the challenge of multiscale approaches like homogenization or upscaling. In this work, we use an operator-based upscaling method. Operator-based upscaling methods were first developed for elliptic flow problems (see [41]) and then extended to hyperbolic problems (see [62], [73], [72]). Operator-based upscaling method consists in splitting the solution into a coarse and a fine part. The coarse part is defined on a coarse mesh while the fine part is computed on a fine mesh. In order to speed up
calculations, artificial boundary conditions (ABC) are imposed. By enforcing suitable ABCs on the boundary of every cells of the coarse mesh, calculations on the fine grid can be carried out locally. The coarse part is next computed globally on the coarse mesh. Operator-based upscaling methods were so far developed in joint with standard finite element discretisation strategy. In this work, we investigate the idea of combining an operator based upscaling method with discontinuous Galerkin finite element methods (DGFEM). To begin with, we have used the interior penalty method as presented in [42] for elliptic problems and in [61], [60] for the wave equation. This is a quite natural way of addressing this issue because we can use a software package that has been already developed in the team. The first results that we have obtained seem to indicate that an DG operator based upscaling method could be interesting essentially in case of stationary problems. Nevertheless, the numerical analysis of the discretized problem must be continued. This work has been initiated during the internship of Theophile Chaumont-Frelet who was a fourth year engineer student at Rouen INSA. A paper dealing with the case of the Laplace operator will be submitted soon.

6.2.11. Discontinuous Galerkin Methods for Seismic Wave Propagation

Participants: Hélène Barucq, Caroline Baldassari, Lionel Boillot, Marie Bonnasse, Julien Diaz, Jérôme Luquel, Vanessa Mattesi, Florent Ventimiglia.

In the framework of our collaboration with Total, we are implementing a Discontinuous Galerkin formulation of the first order elastodynamic wave equations in the platform Diva which is developed by Total. We consider the formulation proposed in [54] for isotropic media. During her post-doc, Caroline Baldassari has implemented a three dimensional code with Perfectly Matched Layers for this formulation. Jérôme Luquel has implemented the 2D version of this code during his internship. In the framework of the internship of Marie Bonnasse and the PhD thesis of Lionel Boillot, we have extended the formulation to Vertical Transverse Isotropic and Tilted Transverse Isotropic media in both 2D and 3D. The introduction of Absorbing Boundary Conditions or of PML is still an open problem for these types of media. It is one of the topics of the PhD thesis of Lionel Boillot.

The version of the code that we are using assumed that the properties of the media (density, velocity,...) are constant on each cells of the mesh. Discontinuous Galerkin methods allow for considering more general configurations, where these properties vary as polynomial functions inside each cells. Hence, it is not necessary to define the interfaces between the different media before constructing the mesh. The discontinuities are taken into account directly inside each cells. Moreover, we are able to consider smoothly varying media. In the framework of the internship of Vanessa Mattesi, we have implemented polynomial velocities in a Discontinuous Galerkin formulation. We have compared the results obtained with this method to the one obtained with piecewise constant properties. We have observed that the new formulation was more accurate and that it allowed for a simpler construction of the mesh. However, these gains do not counterbalance the increase of the computational induced by the new method. We have then concluded that considering piecewise constant properties was more appropriate to model seismic wave propagation.

6.2.12. Elastic surface waves in crystals

Participants: José Carcione, Fabio Cavallini, Dimitri Komatitsch, Nathalie Favretto-Cristini.

In [25], we present a review of wave propagation at the surface of anisotropic media (crystal symmetries). The physics for media of cubic and hexagonal symmetries has been extensively studied based on analytical and semi-analytical methods. However, some controversies regarding surfaces waves and the use of different notations for the same modes require a review of the research done and a clarification of the terminology. In a companion paper we obtain the full-wave solution for the wave propagation at the surface of media with arbitrary symmetry (including cubic and hexagonal symmetries) using two spectral numerical modeling algorithms.

In [27], we obtain the full-wave solution for the wave propagation at the surface of anisotropic media using two spectral numerical modeling algorithms. The simulations focus on media of cubic and hexagonal symmetries, for which the physics has been reviewed and clarified in a companion paper. Even in the case of homogeneous media, the solution requires the use of numerical methods because the analytical Green’s
function cannot be obtained in the whole space. The algorithms proposed here allow for a general material variability and the description of arbitrary crystal symmetry at each grid point of the numerical mesh. They are based on high-order spectral approximations of the wave field for computing the spatial derivatives. We test the algorithms by comparison to the analytical solution and obtain the wave field at different faces (stress-free surfaces) of apatite, zinc and copper. Finally, we perform simulations in heterogeneous media, where no analytical solution exists in general, showing that the modeling algorithms can handle large impedance variations at the interface.

6.2.13. Application of an elastoplastic spectral-element method to 3D slope stability analysis

Participants: Hom Nath Gharti, Dimitri Komatitsch, Oye Volker, Roland Martin, Jeroen Tromp.

In [26], we implement a spectral-element method for 3D time-independent elastoplastic problems in geomechanics. As a first application, we use the method for slope stability analyses ranging from small to large scales. The implementation employs an element-by-element preconditioned conjugate gradient solver for efficient storage. The program accommodates material heterogeneity and complex topography. Either simple or complex water table profiles may be used to assess effects of hydrostatic pressure. Both surface loading and pseudostatic seismic loading are implemented. In order to simulate elastoplastic behavior of slopes, a Mohr-Coulomb yield criterion is employed using an initial strain method (i.e., a viscoplastic algorithm). For large-scale problems, the software is parallelized based on domain decomposition using MPI (Message Passing Interface). Strong-scaling measurements demonstrate that the parallelized software performs efficiently. We validate our spectral-element results against several other methods, and apply the technique to simulate failure of an earthen embankment and a mountain slope.

6.2.14. Indirect Boundary Element Method applied to Fluid-Solid Interfaces


In [31], scattering of elastic waves in fluid-solid interfaces is investigated. We use the Indirect Boundary Element Method to study this wave propagation phenomenon in 20 models. Three models are analyzed: a first one with an interface between two half-spaces, one fluid on the top part and the other solid in the bottom; a second model including a fluid half-space above a layered solid; and finally, a third model with a fluid layer bounded by two solid half-spaces. The source, represented by Hankel’s function of the second kind, is always applied in the fluid. This indirect formulation can give to the analyst a deep physical insight on the generated diffracted waves because it is closer to the physical reality and can be regarded as a realization of Huygens’ principle. In any event, mathematically it is fully equivalent to the classical Somigliana’s representation theorem. In order to gauge accuracy we test our method by comparing with an analytical solution known as Discrete Wave Number. A near interface pulse generates scattered waves that can be registered by receivers located in the fluid and it is possible to infer wave velocities of solids. Results are presented in both time and frequency domain, where several aspects related to the different wave types that emerge from this kind of problems are pointed out.

6.2.15. Multiperforated plates in linear acoustics

Participants: Mohamed Amara, Sharang Chaudhry, Julien Diaz, Rabia Djellouli, Magdalena Grigoroscuta-Strugaru.

We have designed a new and efficient solution methodology for solving high-frequency Helmholtz problems. The proposed method is a least-squares based technique that employs variable bases of plane waves at the element level of the domain partition. A local wave tracking strategy is adopted for the selection of the basis at the regional/element level. More specifically, for each element of the mesh partition, a basis of plane waves is chosen so that one of the plane waves in the basis is oriented in the direction of the propagation of the field inside the considered element. The determination of the direction of the field inside the mesh partition is formulated as a minimization problem. Since the problem is nonlinear, we apply Newton’s method to determine the minimum. The computation of Jacobians and Hessians that arise in the iterations of the Newton’s method is based on the exact characterization of the Fréchet derivatives of the field with
respect to the propagation directions. Such a characterization is crucial for the stability, fast convergence, and computational efficiency of the Newton algorithm. These results are part of the Master thesis of Sharang Chaudhry (student at CSUN).

6.3. High Performance methods for solving wave equations

6.3.1. Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes

**Participants:** Daniel Peter, Dimitri Komatitsch, Yang Luo, Roland Martin, Nicolas Le Goff, Emanuelle Casarotti, Pietre Le Loher, Federica Magnoni, Qinya Liu, Céline Blitz, Tarje Nissen-Meyer, Piero Basini, Jeroen Tromp.

In [30], we present forward and adjoint spectral-element simulations of coupled acoustic and (an)elastic seismic wave propagation on fully unstructured hexahedral meshes. Simulations benefit from recent advances in hexahedral meshing, load balancing and software optimization. Meshing may be accomplished using a mesh generation tool kit such as CUBIT, and load balancing is facilitated by graph partitioning based on the SCOTCH library. Coupling between fluid and solid regions is incorporated in a straightforward fashion using domain decomposition. Topography, bathymetry and Moho undulations may be readily included in the mesh, and physical dispersion and attenuation associated with anelasticity are accounted for using a series of standard linear solids. Finite-frequency Fréchet derivatives are calculated using adjoint methods in both fluid and solid domains. The software is benchmarked for a layercake model. We present various examples of fully unstructured meshes, snapshots of wavefields and finite-frequency kernels generated by Version 2.0 ‘Sesame’ of our widely used open source spectral-element package SPECFEM3D.

6.3.2. Fluid-solid coupling on a cluster of GPU graphics cards for seismic wave propagation

**Participant:** Dimitri Komatitsch.

In [28], we develop a hybrid multiGPUs and CPUs version of an algorithm to model seismic wave propagation based on the spectral-element method in the case of models of the Earth containing both fluid and solid layers. Thanks to the overlapping of communications between processing nodes on the computer with calculation by means of non-blocking message passing, we obtain excellent weak scalability of this finite-element code on a cluster of 192 GPUs and speedup factors of more than one order of magnitude compared to the same code run on a cluster of traditional CPUs. This enables us to show a new geophysical phenomenon concerning wave propagation of diffracted shear waves in a layer called D” located at the base of the Earth’s mantle, namely that in this layer the transverse and radial components of these waves can undergo a relative shift even in an isotropic Earth model, whereas this observation in real seismological data was interpreted until now as an indication of the presence of anisotropy in this layer.
6. New Results

6.1. Yeast comparative genomics

**Participants:** David James Sherman, Pascal Durrens [correspondent], Tiphaine Martin, Nicolás Loira.

By using the MAGNOME software developments, including the MAGUS system and YAGA software, we have successfully realized a full annotation and analysis of seven new genomes, provided to the Génolevures Consortium by the CEA–Génoscope (Évry). Two distant genomes from the Debaryomycetaceae and mitosporic Saccharomycetales clades of the Saccharomycetales were annotated using previously published Génolevures genomes [5], [9], [10] as references. A further group of five species, comprised of pathogenic and nonpathogenic species, was analyzed with the goal of identifying virulence determinants [13]. By choosing species that are highly related but which differ in the particular traits that are targeted, in this case pathogenicity, we are able to focus on the few hundred genes related to the trait. The approximately 40,000 new genes from these studies were classified into existing Génolevures families as well as branch-specific families. The results from these two studies will be published in the coming year.

6.2. Assembly, annotation and comparison of Oenococcus strains

**Participants:** David James Sherman, Pascal Durrens, Elisabeth Bon [correspondent], Tiphaine Martin, Aurélie Goulielmakis.

*Oenococcus oeni* is part of the natural microflora of wine and related environments, and is the main agent of the malolactic fermentation (MLF), a step of wine making that generally follows alcoholic fermentation (AF) and contributes to wine deacidification, improvement of sensorial properties and microbial stability. The start, duration and achievement of MLF are unpredictable since they depend both on the wine characteristics and on the properties of the *O. oeni* strains. In collaboration with Patrick Lucas’s lab of the ISVV Bordeaux that is currently proceeding with genome sequencing, explorative and comparative genomics, Elisabeth Bon coordinates our efforts into the OENIKITA project (since 2009), a scale switching challenge including highthroughput exploratory and comparative genomics for oenological bacterial starters, and the development of an online web-collaborative multigenomic comparative platform (under development) based on the the Génolevures database architecture and MAGUS / YAGA systems.

**OENI-Genomics axis:** In comparative genomics, we investigated gene repertoire and genomic organization conservation through intra- and inter-species genomic comparisons, which clearly show that the *O. oeni* genome is highly plastic and fast-evolving. Results reveal that the optimal adaptation to wine of a strain mostly depends on the presence of key adaptive loops and polymorphic genes. They also point up the role of horizontal gene transfer and mobile genetic elements in *O. oeni* genome plasticity, and give the first clues of the genetic origin of its oenological aptitudes. As a result of the scaling out challenge, we completed the assembly of 19 fully sequenced *O. oeni* genome variants.

**KITA-Genomics (E. Bon, D. Sherman):** This project that is focused on the sequencing, assembly, exploration and comparison of the *O. kitaharae* genome, has benefited to an international collaboration involving Dr V. Makeev. MAGNOME contributed to the assembly of the genome. The comparison against the *O. oeni* genomic architecture will contribute to shed light on the evolutionary mechanisms which are responsible for the atypically long branch of the genus Oenococcus in phylogenetic trees.
Transcriptomic axis (E. Bon, A. Goulielmakis): Under the supervision of E. Bon, Aurélie Goulielmakis has completed for the ANR DIVOENI a detailed manual annotation of a new reference strain of *O. oeni* and performed comparative transcriptome analysis to identify genes differentially expressed under different culture conditions. We explored and compared how the expression system is solicited when *O. oeni* strains adapted to grow in some niches are placed under stress-exposure conditions. The monitoring of gene expression status between strains, through the definition of a global expression pattern proper to each gene, partially lift the veil on how *O. oeni* genome adapts function to its environment. The weight of genetic background and ecological niche pressure on gene expression flexibility was evaluated, and the *O. oeni* pan-transcriptome architecture characterized. The first guidelines revealed a supra-spatial organization of stress response into activated and repressed larger macro-domains defining functional landmarks and intra-chromosomal territories [16]. Decryption of stress-sensitive gene repertoires promises to be an efficient tool in the conquest of *O. oeni* “domestication” through the identification of molecular markers responsible for different physiological capabilities, and the selection of the best adapted strains.

Gene plasticity modelisation (E. Bon, A. Goulielmakis): A novel axis of research recently emerged under the initiative of E. Bon (pseudOE project) around the detection, characterization and conservation of pseudogenes populations in Oenococcus bacteria. Such topic presents a double interest: phylogenetic at first because it should allow to better estimate the degree of genic/genomic plasticity of these bacteria, and algorithmic then because the pseudogenes are a source of confusion for the automatic prediction of genes. Through a transversal collaboration and a cooperative supervision with the Algorithms for Analysis of Biological Structures Group (P. Ferraro, J. Allali) at LaBRI, Laetitia Bourgeade (PhD, Univ. Bordeaux1) was recruited to develop dedicated methods to improve pseudogenes automatic detection, and therefore gene predictions, and to reconstruct fossil and modern genes evolutionary history.

6.3. Scaling-out

Participants: David James Sherman [correspondant], Pascal Durrens, Tiphaine Martin, Natalia Golenetskaya, Florian Lajus.

The Tsvetok project in MAGNOME targets “scaling out” for data and computation, both to improve capacity for handling large volumes of data and to permit more automatic analysis of projects of the “comparative genomics of related species” type, where a set of genomes is sequenced and analyzed as part of the same process. Natalia Golenetskaya has designed and implemented a NoSQL schema through the identification of standard queries, definition of the appropriate query-oriented storage schema, and mapping of structured values to this schema. This prototype is being tested on an Apache Cassandra ring deployed in MAGNOME’s dedicated computing cluster.

Large-scale data-mining such as that required for comparative genomics is fundamentally *data-parallel*: an initial transformation is applied to every data object of a given type (such as genes or even individual nucleotides), then a statistical machine learning procedure is applied to the transformed data to produce a summary or to learn a classification function. Analyses of this kind are the design goal of the MapReduce paradigm [31]. Using Tsvetok as a generator for Apache Hadoop, Natalia is designing MapReduce solutions for the principal whole-genome and data-mining analyses used by MAGNOME for eukaryote and prokaryote comparative genomics.

6.4. Affinity Proteomics: Standards for affinity binders

Participants: David James Sherman [correspondant], Natalia Golenetskaya.

Last year we successfully completed and released the MIAPAR and PSI-PAR international standards for knowledge representation and data exchange of affinity binder properties, a five-year effort organized as part of the ProteomeBinders and HUPO-PSI consortia. These standards were reported in *Nature Biotechnology* and *Molecular and Cellular Proteomics* to the research community [3] [37], [52] and extend previous work [38], [43]. One long-standing issue is the adoption of these standards by individual researchers in the lab: initial data entry must be simple enough that standards-based reporting can be integrated into the process
of writing the paper. We used an extensive dataset of affinity proteomics data to evaluate “last mile” tools for data entry and initial reporting of affinity proteomics data, and identified places where existing tools need to be adapted to meet these specific needs[ 21 ].

6.5. Inferring metabolic models

Participants: David James Sherman [correspondant], Pascal Durrens, Tiphaine Martin, Nicolás Loira, Anna Zhukova.

In collaboration with Prof Jean-Marc Nicaud’s lab at the INRA Grignon, we developed the first functional genome-scale metabolic model of an oleaginous yeast. Most work in producing genome-scale metabolic models has focused on model organisms, in part due to the cost of obtaining well-annotated genome sequences and sufficiently complete experimental data for refining and verifying the models. However, for many fungal genomes of biotechnological interest, the combination of large-scale sequencing projects and in-depth experimental studies has made it feasible to undertake metabolic network reconstruction for a wider range of organisms.

An excellent representative of this new class of organisms is Yarrowia lipolytica, an oleaginous yeast studied experimentally for its role as a food contaminant and its use in bioremediation and cell factory applications. As one of the hemiascomycetous yeasts completely sequenced in the Génolevures program it enjoys a high quality manual annotation by a network of experts. It is also an ideal subject for studying the role of species-specific expansion of paralogous families, a considerable challenge for eukaryotes in genome-scale metabolic construction. To these ends, we undertook a complete reconstruction of the Y. lipolytica metabolic network.

Methods: A draft model was extrapolated from the S. cerevisiae model iIN800, using in silico methods including enzyme conservation predicted using Génolevures and reaction mapping maintaining compartments. This draft was curated by a group of experts in Y. lipolytica metabolism, and iteratively improved and validated through comparison with experimental data by flux balance analysis. Gap filling, species-specific reactions, and the addition of compartments with the corresponding transport reactions were among the improvements that most affected accuracy.

Results: We produced an accurate functional model for Y. lipolytica, MODEL1111190000 in Biomodels.net, that has been qualitatively validated against gene knockouts.

6.6. Hierarchical modeling with BioRica

Participants: David James Sherman [correspondant], Tiphaine Martin, Alice Garcia, Rodrigo Assar-Cuevas, Nicolás Loira.

A recurring challenge for in silico modeling of cell behavior is that experimentally validated models are so focused in scope that it is difficult to repurpose them. Hierarchical modeling is one way of combining specific models into networks. Effective use of hierarchical models requires both formal definition of the semantics of such composition, and efficient simulation tools for exploring the large space of complex behaviors.

BioRica is a high-level hierarchical modeling framework for models combining continuous and discrete components. By providing a reliable and functional software tool backed by a rigorous semantics, we hope to advance real adoption of hierarchical modeling by the systems biology community. By providing an understandable and mathematically rigorous semantics, this will make it easier for practicing scientists to build practical and functional models of the systems they are studying, and concentrate their efforts on the system rather than on the tool.

Rodrigo Assar formalized two strategies for integrating discrete control with continuous models, coefficient switches that control the parameters of the continuous model, and strong switches that choose different models. This was translated by Alice Garcia into a BioRica specification for hybrid systems that assures integrity of models, allowing composition, reconciliation, and reuse of models with SBML specifications. Rodrigo used this approach to describe two systems: wine fermentation kinetics, and cell fate decisions leading to bone and fat formation[ 11 ]. In the first, known models that describe the responses of yeast cells to different
temperatures, resources and toxins, were reconciled using coefficient switches that gave the best adjustment of the model depending on the initial conditions and fermentation variable. In the second, a combination of accurate models to predict the bone and fat formation in response to activation of pathways such as the Wnt pathway, and changes of conditions affecting these functions such as increments in Homocysteine, were used to analyze the responses to treatments for osteoporosis and other bone mass disorders. Our hope is that this is a first step in obtaining in silico evaluations of medical treatments before testing them in vivo or in vitro.

Maria Lluberes of the University of Puerto Rico visited MAGNONE and we established formal relationships between BioRica models and probabilistic boolean networks.
5. New Results

5.1. Metapopulation

In the framework of metapopulation models we have obtained some new results. We have studied a metapopulation model with $n$ patches. The migration model is with residents and travelers and the epidemic model is of SIS type. In particular, we have proved analytically the conjecture of Arino and van den Driessche [12]. The global behavior of the system is addressed in [10].

We are also currently studying, with our brazilian partners in the framework of the CAPES-COFECUB project (see international program), a metapopulation model for studying the propagation of dengue in the state of Rio. We use data provided by the foundation FIOCRUZ. These data concern different epidemics in RIO between 1986 to 2006. The effects of transportation system are taken into account.

![Transportation system in Rio](image)

5.2. Identification and state estimation

We identify some unknown parameters using the framework of observers. We have used numerical observers to identify the transmission parameters of bilharzia.

We also have used observers for models with unknown inputs to identify the transmission parameters for intra-host models of malaria. We have developed a method for estimating total parasite load in *falciparum* malaria patients using the clinical observations of peripheral parasitaemia. We have applied this method using the data from malariatherapy. The results are illustrated in Figure 4.
Figure 2. Reduced metapopulation model

Figure 3. Model vs data
Figure 4. Estimated total parasite load (red) and parasitaemia measures (blue)
6. New Results

6.1. Theoretical results

6.1.1. Theory of competition for one limited resource

**Participants:** Claude Lobry, Tewfik Sari, Radhouene Fekih-Salem.

In the paper [24], we give a global asymptotic stability result for a mathematical model of competition between several species in a chemostat, by using a new Lyapunov function. The model includes both monotone and non-monotone response functions, distinct removal rates for the species and variable yields, depending on the concentration of substrate.

In the paper [14], we consider the mathematical model of two species microbial competition on a single food resource in a chemostat, when one takes into account species interactions between the two populations of microorganisms and intraspecific interactions between individuals themselves, using strictly monotonic growth functions and distinct dilution rates.

6.1.2. Study of input/output maps of interconnected chemostats

**Participants:** Alain Rapaport, Ihab Haidar.

Patch or island models are popular in ecology, and are a convenient way to study the influence of a spatial structure of a geography on the distribution of the abundance of resources. Coupling such a structure with abiotic/biotic models and studying its input-output properties has been very rarely tackled in the literature. In biotechnology engineering, dead-zones models, that distinguish two sub-domains (a “living” and a “dead” one) are often used for approximating non perfectly mixed tanks. No more sophisticated representation, apart continuous space models (systems of partial differential equations), have been investigated.

We consider an hydric capacity and an nutrient flow that are fixed, and analyze the influence of different structures, having the same total hydric volume, on the output concentrations at steady-state. Three configurations are compared, under the assumption of a monotonic growth rate: perfectly-mixed, serial and parallel with diffusion rate. In each case, we show the uniqueness of a steady-state different to the washout equilibrium and its global asymptotic stability in the positive orthant. We prove the existence of a threshold on the input concentration of nutrient for which the benefits of the serial and parallel configurations over the perfectly-mixed one are reversed. In addition, we show that the dependency of the output concentrations on the diffusion rate can be non-monotonic, and give precise conditions for the diffusion effect to be advantageous [19], [33]. The study encompasses the dead-zone models.

The possibly non-monotonic influence of the diffusion parameter on the output steady state is not intuitive, and leaves further investigations open for understanding or taking benefit of this property for natural ecosystems (such as saturated soils or wetlands) as well as for bioprocesses (such as wastewater treatments). This result can be also of interest for reverse engineering when deciding which among serial or parallel configurations is a better fit for the modelling of chemostat-like ecosystems, providing that one has an estimation of the hydric capacity of the system.

This work is part of the material thesis of I. Haidar [11].

6.1.3. Aggregation models in the chemostat

**Participants:** Claude Lobry, Alain Rapaport, Jérôme Harmand, Tewfik Sari, Radhouene Fekih-Salem.
Bacteria aggregation often occurs in bioprocesses, creating flocks or biofilms (the latter being attached to the tank walls). At a macroscopic level with large populations of aggregated and non-aggregated individuals, a simple way of modelling this phenomenon in the chemostat is to distinguish explicitly two populations: planktonic or free bacteria and attached ones. The main differences between flocks and biofilms rely on the attachment/detachment terms and the effective dilution rate (assumed to be zero or very small for biofilms). Typically, the specific growth rate of free bacteria is expected to be larger than the attached one (that have in average a restricted access to nutrient and use part of their energy to glue together).

Based on former works of the team and the main assumption that attachment and detachment dynamics are much faster than the biological one, we have shown that a significant difference between the specific dilution rates of the free and attached populations can surprisingly lead to bi-stability, even for a single species and monotonic growth rates \[ 30 \], \[ 42 \], \[ 43 \].

A work in progress addresses the case of two species, one of them having a non-monotonic growth rate (due to substrate inhibition) and the ability to form flocks. Without flock, the Competitive Exclusion Principle extended by G. Wolkowicz and her co-authors \[ 47 \] shows the possibility to have one of the two species winning the competition depending on their initial repartition. Here, the presence of flocks may lead on the contrary to a single winner.

### 6.1.4. Neutral community models for microbial ecology

**Participant:** Bart Haegeman.

Hubbell’s neutral model \[ 50 \] describes the dynamics of an ecological community in terms of random birth, death, immigration and speciation events, attributing equivalent characteristics to all species. Despite the absurd simplicity of these assumptions, remarkable agreement between neutral model predictions (e.g., the distribution of the abundance of the species present in the community) and empirical observations has been reported for some, mostly rather diverse, ecological communities.

There is some evidence that also certain aspects of microbial communities can be well described by the neutral model. Highly diverse microbial communities have been difficult to deal with using more traditional modelling approaches from community ecology. The neutrality assumption could lead to an effective global description, without requiring quantitative species data (growth characteristics, interaction strengths, etc). We are actively participating in the development of neutral community models, with a focus on microbial systems.

1. **Effect of speciation process**

It has been argued that the neutral model predictions are rather insensitive to its assumptions. However, we have found that the details of the way new species appear in the community (i.e., the speciation process) do matter, and can drastically change the model predictions. In particular, we have studied the neutral community model with random fission speciation. This speciation model is quite different from the point mutation model usually considered in neutral community model, and is generally believed to be more realistic.

Using a technique from theoretical physics, we have obtained the stationary distribution of species abundances for the random fission model. We have compared our solution with the well-known stationary distribution of species abundances for the point mutation model on empirical data (tree communities in tropical forests) \[ 13 \]. Surprisingly, we found that the point mutation model fits the data better than the random fission model, although the latter is believed to be more realistic.

2. **Comparison with niche models**

Neutral community models challenge more traditional, niche-based models in community ecology. Niche theory states that species can coexist only if they differ sufficiently in their characteristics (for example, their use of available substrates). Neutral theory assumes that all species have approximately equal characteristics. Hence, the two theories describe species coexistence in fundamentally different ways.

We have tried to narrow the gap between the two theories. We have proposed a mathematical model that combines essential features of niche-based and neutral community models \[ 17 \]. It integrates species niches, described as Lotka-Volterra interactions, in the standard neutral community model. The analysis of this model
indicates that the addition of species interactions has a limited effect on the species abundance distribution. We have further clarified this result using a slightly different model that also combines niche and neutral features [16]. For the latter model we have proved that the niche structure does not affect at all the species abundance distribution.

### 6.1.5. Quantifying genetic diversity of bacteria

**Participant:** Bart Haegeman.

With the wide availability of DNA sequencing, microbiologists are now able to rapidly sequence entire bacterial genomes. Comparison of these genomes has revealed a large genetic diversity within bacterial species. For example, one genome of the bacteria *E. coli* has about 4000 different genes, but a set of 10 genomes of *E. coli* has typically over 10000 different genes. Some of these genes are shared by all or almost all of the genomes, but many other genes are only present in one or a few of the genomes. This observation has important implications for the definition of bacterial species and for the description of the functional characteristics of bacteria.

We have been dealing with the problem of how to quantify this observed genetic diversity. Microbiologists have introduced notions like the pan genome of a bacterial species (that is, the set of genes that are present in at least one of the organisms from the species) and the core genome of a bacterial species (that is, the set of genes that are present in every organisms from the species). However, we have argued that both the pan and core genome are difficult to estimate, and should not be used for quantitative purposes [21]. Instead, we have proposed a measure of genetic diversity that has much better estimation properties. It is based on the average number of genes shared by a pair of genomes when sequencing two randomly sampled organisms from the species under consideration. We have applied our estimator on six bacterial species (about 100 sequenced genomes in total). Software for our robust estimation procedure of genetic diversity is freely available, see http://ecotheory.biology.gatech.edu/downloads/genomic-fluidity-scripts.

### 6.1.6. Individual-based modelling

**Participants:** Fabien Campillo, Chloé Deygout, Coralie Fritsch, Marc Joannides, Claude Lobry.

In terms of computational modelling of ecosystems, individual-based models (IBMs) are an interesting path to explore. We can outline two types of IBMs. On the one hand “detailed IBM” attempt to integrate in an ad-hoc way all the knowledge available about an ecosystem. On the other hand, “simplified IBM” are limited to one or several mechanisms to simplify the analysis. The former may be more realistic but are often difficult to analyze. Although the latter are too simplistic in realistic situations they lend themselves to the analysis and numerical analysis. We focus on the latter.

The IBMs offer an interdisciplinary language between biologists, biotechnologists, mathematicians, and computer scientists, to develop models in the form of relatively simple rules. In the case of simplified IBMs it is possible to translate these rules in the form of a branching Markov process with values in a space of measures. Using scaling methods, the IBMs can be approximated by integro-differential equations; using model simplification methods IBMs can be reduced to stochastic or ordinary differential equations. The mathematical interpretation of the IBMs and their analysis is relatively recent and still very few studies exist [48]. The numerical analysis of these models is yet to be built. Under certain conditions, IBMs themselves can be simulated through adapted Monte Carlo procedures.

The MODEMIC project-team develops three studies in the field of IBMs. The first is part of the ANR MODECOC on the modelling of clonal plant growth (see Section 7.5); the second is part of the ANR DISCO on modelling of biofilms (see Section 7.4); the last one is a starting thesis. In all cases, we aim at developing the Monte Carlo simulation of the IBM as well as analyzing their links with integro-differential models. We also seek to make connections with non-IBM models proposed in Section 6.1.8.
In October 2011, Coralie Fritsch started a thesis at the École Doctorale I2E of the University of Montpellier 2, under the supervision of Fabien Campillo, Jérôme Harmand and Marc Joannides. This thesis is supported by a grant of the MESR and a grant of INRA from the MEM Meta-program (Méta-omiques des écosystèmes microbiens). The thesis aims at developing and analyzing individual-based microbial ecosystems models that capture both the spatial, biodiversity and function of these ecosystems. The thesis received the Agreenium label in December.

6.1.7. Hybrid modelling of biofilms in plug-flow reactors

Participants: Fabien Campillo, Chloé Deygout, Annick Lesne, Alain Rapaport.

Within the DISCO project of the SYSCOMM program founded by the ANR, we have proposed a multi-scaled modelling that combines three scales: a microscopic one for the individual bacteria, a mesoscopic or “coarse-grained” one that homogenises at an intermediate scale the quantities relevant to the attachment/detachment process, and a macroscopic one in terms of substrate concentration (see the Section 7.4).

Such an “hybrid” approach allows for modelling and understanding in plug-flow reactors [41] the interplay between

- the formation of the biofilm at a microscopic scale, that starts from a small number of bacteria (thus a stochastic individual based description),
- the limitation of the biofilm, due the carrying capacity of the wall attachment, at a mesoscopic scale,
- the consumption of nutrient along the flow at a macroscopic level, as a solution of a coupled transport-reaction partial differential equation.

The numerical computation of such a model requires a software architecture that allows the simultaneous simulation of stochastic events at the bacteria scale and the continuous evolution (in space an time) of the substrate density.

Experiments on real tubular plug-flow reactors are currently driven at Cemagref HBAN with the perspective of comparison with numerical simulations. After spending one year at Montpellier for deriving and simulating the theoretical model, our post-doctoral fellow C. Deygout is presently participating to the real experiments at Cemagref Antony.

The multi-species case with different bacteria specialized in different environments (poor or rich in nutrient) is a work in progress.

6.1.8. Stochastic modelling for biotechnology

Participants: Fabien Campillo, Marc Joannides, Claude Lobry.

This year we continue to study stochastic models for the chemostat [12], [39], [27]. Starting from the well-known ordinary differential equation systems, we propose first a pure jump process model at the microscopic scale that leads to a stochastic differential equation at the intermediate scale and to an ordinary differential equation at the macroscopic level (fluid limit model). After developing the model, we establish the Fokker-Planck partial differential equation for the diffusion model. This PDE integrates a specific washing-out term. We proposed an ad hoc numerical integration scheme for the simulation of this PDE [39].

In [40], we consider a stochastic version of the basic predator-prey differential equation model. The model, which contains a parameter ω which represents the number of individuals for one unit of prey – if x denotes the quantity of prey in the differential equation model x = 1 means that there are ω individuals in the discontinuous one – is derived from the classical birth and death process. It is shown by the mean of simulations and explained by a mathematical analysis based on results in singular perturbation theory (the so called theory of Canards) that qualitative properties of the model like persistence or extinction are dramatically sensitive to ω. For instance, in our example, if ω = 10^7 we have extinction and if ω = 10^8 we have persistence. This means that we must be very cautious when we use continuous variables in place of jump processes in dynamic population modelling even when we use stochastic differential equations in place of deterministic ones.
6.1.9. Minimal time control of batch bioprocesses

Participants: Denis Dochain, Alain Rapaport.

Minimal time control problems often occur in biotechnology when one has to fill tanks. Typically, the objective to be reached is to have the tank full with a prescribed value of substrate or product concentrations, the tank being filled with a high concentration of nutrient.

When a single reaction occurs, the optimal solution is already known and has been rigorously proved by J. Moreno in 1999 [51] using the Green’s theorem: it consists in a “bang-bang” strategy (fill as fast as possible or do not fill) and possibly a singular arc when the growth function presents an inhibition (i.e. a maximum growth for a precise concentration of nutrient). When impulse controls in addition to regular control are allowed, an extension of this result has been recently proposed with a different technique that do not use the Green’s theorem [4]. This technique has also allowed to solve partially the problems when several species compete for the degradation of the substrate, but when all of them have a monotonic growth.

In the presence of complex non monotonic kinetics, typically characterized by the combination of two non-monotonic growth functions, aimed at emphasizing the presence of two parallel metabolic pathways to transform the limiting substrate into the biomass, the candidate singular arcs are multiple and determining which singular arc is eventually optimal is clearly a crucial issue. The local optimality conditions based on the Pontryagin Maximum Principle allow to characterize the geometric structure of the extremal trajectories, in which there may be singular arcs, but these necessary conditions are not always sufficient for determining which extremals are (globally) optimal. Then one has to compute the cost of each extremal or use global optimization methods such as dynamic programming or Hamilton-Jacobi-Bellman equation. The extremals are traditionally determined numerically by considering shooting methods, but for bang-bang control, it is well known that one may face numerical troubles because the shooting function is in general not smooth. For this problem, we have used an approximation technique first proposed and studied by C. Lobry and his students [53] and later by C. Silva and E. Trélat [52], that consist in adding an artificial control. In [22], we have proposed a new proof of convergence based on differential inclusions arguments that allows to relax the assumption of the uniqueness of the optimal solution for the convergence of the optimal paths of [52]. Then we have shown how to apply numerically this approximation procedure for analyzing the field of extremals on the whole state space. This technique appears to be quite effective for the practical determination of optimal synthesis in the planar case even in presence of multiple singular arcs.

6.1.10. Optimal control of continuous bioprocesses

Participants: Jérôme Harmand, Alain Rapaport, José Fernandez, Walid Bouhafs, Amel Ghouali.

In continuous bioprocesses, a usual objective is to stabilize the output of the bioreactors about a desired steady state (in wastewater industry, this value is typically chosen under the norm of authorized discharge). It happens more and more frequently that transient trajectories are expected also to maximize a product of interest.

We have begun to study the maximization of the gaseous production of methane in anaerobic processes over a given period of time on specific problems. For the moment we have proved that the optimal trajectory consists in approaching a unique singular arc as fast as possible when only one limiting substrate has to be converted, but the problem is still open when involving several substrates. These works are part of the PhD work of A. Ghouali and W. Bouhafs.

Reference points in batch processes can be mimicked by a series of continuously stirred bioreactors in series at steady state (see applications 6.2.2 and 7.2). We study the minimal time problem to drive the nutrients concentrations of a cascade of chemostats. The control variable is the dilution rates of each tank, under the constraint that each dilution rate is bounded by the one of the previous tank, that makes the system not locally controllable. For the particular case of two tanks with total mass at steady state, the planar feedback synthesis has been found but the problem is still under investigation for the general case.

6.1.11. Minimal time bioremediation of natural resources

Participants: Jérôme Harmand, Alain Rapaport, Antoine Rousseau.
In biological wastewater treatment (batch or continuous bioprocesses), one has always to separate biomass from the purified liquid phase at the output of the tanks, that is not possible when tanks are rather natural reservoirs such as lakes or water tables.

We have proposed a new operation strategy that consists in treating with the help of a bioreactor aside. No bacteria are introduced in the reservoir but water is pumped and treated by microorganisms in a smaller tank, and treated water returns to the reservoir after being separated from the biomass. Consequently, there is no need of a separation operation for the reservoir.

The minimal time control problem consists in controlling the flow rate for having the substrate concentration of the whole reservoir below a given reference value as fast as possible.

Last year, we have determined analytical expressions of optimal feedback strategies for a general class of growth functions under the assumptions that the volume of the bioreactor is much smaller than the reservoir one, and that the spatial repartition of the concentration of the pollutant in the reservoir can be modelled by simple spatial representations: either perfectly mixed or discrete one directional gradient [44], [15]. This year, we have studied more realistic spatial motifs:

- dead-zones: we have shown that the optimal synthesis is identical to the perfectly mixed case, even though the time to reach the target is larger [31].
- two parallel zones, allowing to control the repartition of the flow rate between the two zones. Without diffusion between the zones, the optimal solution is almost straightforward and under investigation in presence of lateral diffusion.

This work is mainly achieved in cooperation with Chilean researchers and PhD students within the associated team DYMECOS.

### 6.2. Applications

#### 6.2.1. Modelling and control of Anaerobic Digestion processes

**Participants:** Amine Charfi, Radhouene Fekih-Salem, Jérôme Harmand, Boumediene Benyahia, Tewfik Sari.

We consider the AM2 or AMOCO model developed in [46] and extend both the model in itself and its analysis to the following cases:

- Depending on the AM2 model parameters, the steady states were analytically characterized and their stability were analyzed. Following this study, it was shown that the overloading tolerance, a parameter proposed in [49] to on-line monitoring anaerobic processes, may be not adapted under certain operating conditions and even lead to bad operating decisions.

- Within the framework of the PhD theses of Amine Charfi and Boumediene Benyahia, we have included the fouling dynamics of membranes into the AM2 and we have analyzed the resulting model (called the AM2b).

- We actually work towards two directions: (i) we are extending these results in including into the AM2 an additional process, i.e. the hydrolysis step in order to study bioprocesses treating solid waste (the resulting model being called the AM3); (ii) we try to find links between complex models such as the ADM1 model and simple models such as the AM2b or the AM3.

#### 6.2.2. Modelling and control of cascade biosystems to mimic batch wine making processes

**Participants:** Jérôme Harmand, Alain Rapaport, José Fernandez.
An experimental setup of four tanks connected in series has been designed by the research unit SPO (Montpellier) for studying four physiological stages of yeast as steady state. The manipulated variables are the flow rates $Q_i$ of each tank with the constraint $Q_i \geq Q_{i-1} \geq 0$, and the objective is to reach simultaneously four set-points in the four tanks. We are studying two kinds of control strategies:

- a linearizing feedback law that drives exponentially the dynamics to the target. This is not the fastest strategy but is has good robustness properties. Nevertheless, the inputs constraint imposes to use saturation functions that provide satisfactory convergence in simulations but that is hard to prove mathematically.
- a minimal time feedback. Due to lack of local controllability imposed by the constraint on the inputs, the optimal synthesis is not smooth with the presence of “barriers” (see Section 6.1.10).

Those feedback laws will be implemented and tested on real pilot plant at SPO lab in the scope of the European project CAFE described in Section 7.2.

6.2.3. Modelling and simulating terrestrial plant ecological dynamics

**Participants:** Fabien Campillo, Ihab Haidar.

This study is part of the ANR Syscomm MODECOL that is done in collaboration particularly with the University of Rennes I, the University of La Rochelle and INRIA. This is the second year of the three years program. We propose a stochastic individual-based model for clonal plant dynamics in continuous time and space, focusing on the effects of the network structure of the plants on the reproductive strategy of ramets. This model is coupled with an explicit advection-diffusion dynamics for resources. We develop a partially exact simulation scheme of the model; the capacity of the model to reproduce specific features of clonal plants, such as their efficiency to forage resources over the field, is numerically studied. Next, we propose a large population approximation of the model for phalanx-type populations, taking the form of an advection-diffusion PDE for population densities, where the influence of the local graph structure of the plant takes the form of a nonlinear dependence in the gradient of resources. This year we improved the simulation code that was proposed last year and made three communications in international conferences [26] [36] [35].

6.2.4. Modelling and inferring agricultural dynamics

**Participants:** Fabien Campillo, Angelo Raherinirina.

The International Laboratory LIRMA supports this work that is done in collaboration with the University of Fianarantsoa in Madagascar and with Dominique Hervé (IRD, Fianarantsoa, Madagascar). The aim is to study the dynamics of agricultural plots on the edge of primary forest. In [38] we propose a Markov chain model where the transition matrix is estimated both by maximum likelihood and Bayesian approaches. We also test if the Markov chain model is adapted to this problem. In an ongoing work we develop semi-Markov models for an extended data set.

6.2.5. Modelling and simulating microbial ecosystems in soils

**Participants:** Ihab Haidar, Jérôme Harmand, Alain Rapaport.

The team studies simple representations of the spatial inhomogeneity for bioprocesses, in terms of networks of interconnected compartments. Each compartment is modelled as a perfectly mixed bioreactor. Simulation of such networks is performed with the software developed by the VITELBIO project (see 5.1). Comparisons of simple structures (cascade of bioreactors) with numerical simulators based on REV (Representative Elementary Volumes) have been performed. The objective is to understand the role of the topology of the network on the biological functions of the overall system. For the moment simple configurations with two to four nodes and one single species have been investigated.

In addition, we study in simulation how software implementations of transport reactions models such as MIN3P, can mimic such simple configurations. Transport reactions models are often used in the community of soil functioning modelling by simulation. Here we compare their numerical computation with the exact solutions that can be approximated with a good accuracy using classical o.d.e. solvers. We found that one has to be careful when the trajectories are close to a non-hyperbolic equilibrium [45], [18].
6.2.6. Numerical optimisation in non perfectly mixed tanks

**Participants:** Jérôme Harmand, Alain Rapaport, Antoine Rousseau.

The optimisation of series of bioreactors in terms of minimizing the total residence time have been already investigated in the literature. In such models, the space has a one-dimensional representation. There is comparatively much less work on 2D or 3D space models. We consider different shapes of reactors of the same volume, and simulate with the help of multi-physics numerical software the coupling of the hydrodynamics laws in 3D or 2D (under cylindrical symmetry), solved numerically by the Navier-Stokes equations, with the system of differential equations of biotic/abiotic concentrations [34].

We are studying the influence of the shape of the domain and possible obstacles on the output concentration at steady state.

With A. Rousseau (EPI MOISE, INRIA Rhône-Alpes) we study with simulation of 2D Navier-Stokes equations the benefits of having several pumping points and how to control the pumps speed to minimize the treatment duration for the bioremediation of natural reservoirs presented in Section 6.1.11. We evaluate the feedback strategies that are derived to be optimal for simple models of o.d.e. in the more realistic framework of hydrodynamics simulation.

6.2.7. Individual-based models for the bacterial degradation of the cellulose

**Participants:** Fabien Campillo, Chloé Deygout, Marc Joannides.

We propose an individual-based model for the degradation of one cellulose bead (dozens of micrometers in diameter) by cellulolytic bacteria. Our aim is to determine the macroscopic degradation behavior. The initial stages of the degradation process may involve a very limited number of bacteria that cannot be properly modelled by classical models based on deterministic equations. In the present work we only consider a two-dimensional model for the degradation of a cellulose disc.

6.2.8. Non-linear filtering for the chemostat

**Participants:** Boumediene Benyahia, Fabien Campillo, Jérôme Harmand.

We propose numerical non-linear filtering approaches for the identification of non-observed components of dynamical systems in the context of the chemostat. This recently started study relies on the work of the project-team in the stochastic modelling of the chemostat. In a preliminary work we consider the bootstrap particle filter.
6. New Results

6.1. Mathematical Modelling of the Ocean Dynamics

6.1.1. Beyond the traditional approximation on the Coriolis force

Participant: Antoine Rousseau.

Formerly, A. Rousseau has performed some theoretical and numerical studies around the derivation of quasi-hydrostatic models. With C. Lucas, he proved that it is sometimes necessary to take into account the cosine part of the Coriolis force (which is usually neglected, leading to the so-called Traditional Approximation). They have also shown that the non-traditional terms do not raise any additional mathematical difficulty in the primitive equations: well-posedness for both weak and strong solutions.

In 2011, A. Rousseau and J. McWilliams (UCLA) proposed in [62] a mathematical justification of the tilt of convective plumes in the quasi-geostrophic regime, thanks to the account of the complete Coriolis force in the so-called quasi-hydrostatic quasi-geostrophic (QHQG) model. The new model has been presented in several international conferences [26], [49].

6.1.2. Coupling Methods for Oceanic and Atmospheric Models

Participants: Eric Blayo, David Cherel, Laurent Debreu, Antoine Rousseau, Manel Tayachi.

6.1.2.1. Interface conditions for coupling ocean models

Many physical situations require coupling two models with not only different resolutions, but also different physics. Such a coupling can be studied within the framework of global-in-time Schwarz methods. However, the efficiency of these iterative algorithms is strongly dependent on interface conditions. As a first step towards coupling a regional scale primitive equations ocean model with a local Navier-Stokes model, a study on the derivation of interface conditions for 2-D $x-z$ Navier-Stokes equations is currently underway (D. Cherel’s PhD thesis). It has been shown theoretically that several usual conditions lead to divergent algorithms, and that a convergent algorithm is obtained when using transmission conditions given by a variational calculation.

Moreover the impact of two numerical schemes (a projection method, and a direct method [67]) on the implementation of the domain decomposition method has been discussed.

Using the direct method in a 2D $x-z$ Navier-Stokes model, D. Cherel has implemented a Schwarz-based domain decomposition method, for which he used the so-called transmission boundary conditions that mix the velocity and pressure variables on an Arakawa-C grid. The numerical results confirm the rate of convergence that has been obtained theoretically, thanks to a Fourier analysis of the semi-discretized problem. A paper is in preparation.

6.1.2.2. Coupling dimensionally heterogeneous models

The coupling of different types of models is gaining more and more attention recently. This is due, in particular, to the needs of more global models encompassing different disciplines (e.g. multi-physics) and different approaches (e.g. multi-scale, nesting). Also, the possibility to assemble different modeling units inside a friendly modelling software platform is an attractive solution compared to developing more and more global complex models. More specifically one may want to couple 1D to 2D or 3D models, such as Shallow Water and Navier Stokes models: this is the framework of our partnership with EDF in the project MECSICO. In her PhD, M. Tayachi is aimed to build a theoretical and numerical framework to couple 1D, 2D and 3D models for river flows.

This year, she obtained both numerical and theoretical results on a Laplace equation in a domain that suggests a domain decomposition method with two sub-domains that do not have the same space dimension (see Figure 1). A paper is in preparation.
6.1.3. Numerical schemes for ocean modelling

Participants: Laurent Debreu, Jeremie Demange.

Reducing the traditional errors in terrain-following vertical coordinate ocean models (or sigma models) has been a focus of interest for the last two decades. The objective is to use this class of model in regional domains which include not only the continental shelf, but the slope and deep ocean as well. Two general types of error have been identified: 1) the pressure-gradient error and 2) spurious diapycnal diffusion associated with steepness of the vertical coordinate. In a recent paper [78], we have studied the problem of diapycnal mixing. The solution to this problem requires a specifically designed advection scheme. We propose and validate a new scheme, where diffusion is split from advection and is represented by a rotated biharmonic diffusion scheme with flow-dependent hyperdiffusivity satisfying the Peclet constraint.

In 2011, in collaboration with F. Lemarié at UCLA, this work has been extended in order to render the biharmonic diffusion operator scheme unconditionally stable (paper submitted to ocean modelling). This is particularly needed when the slopes between coordinates lines and isopycnals surfaces are important so that the rotation of the biharmonic leads to strong stability condition along the vertical coordinate where the grid size is relatively small. This work also extends more classical results on the stability of laplacian diffusion with mixed derivatives.

In his Ph'D, Jeremie Demange begins a work on advection-diffusion schemes for ocean models (Supervisors: L. Debreu, P. Marchesiello (IRD)). His work will focus on the link between tracers (temperature and salinity) and momentum advection and diffusion in the non hyperbolic system of equations typically used in ocean models (the so called primitive equations with hydrostatic and Boussinesq assumptions).

Salinity at 1000 m in the Southwest Pacific ocean is shown in figure 2. The use of traditional upwind biased schemes (middle) exhibits a strong drift in the salinity field in comparison with climatology (left). The introduction of high order diffusion rotated along geopotential surfaces prevents this drift while maintaining high resolution features (right).

6.2. Development of New Methods for Data Assimilation

6.2.1. Variational Data Assimilation with Control of Model Error

Participants: Bénédicte Lemieux-Dudon, Arthur Vidard.
One of the main limitations of the current operational variational data assimilation techniques is that they assume the model to be perfect mainly because of computing cost issues. Numerous researches have been carried out to reduce the cost of controlling model errors by controlling the correction term only in certain privileged directions or by controlling only the systematic and time correlated part of the error.

Both the above methods consider the model errors as a forcing term in the model equations. Trémolet (2006) describes another approach where the full state vector (4D field: 3D spatial + time) is controlled. Because of computing cost one cannot obviously control the model state at each time step. Therefore, the assimilation window is split into sub-windows, and only the initial conditions of each sub-window are controlled, the junctions between each sub-window being penalized. One interesting property is that, in this case, the computation of the gradients, for the different sub-windows, are independent and therefore can be done in parallel.

We are implementing this method in a realistic Oceanic framework using OPAVAR/ NEMOVAR as part of the VODA ANR project. An extensive documentation is being produced and should be available along with a first prototype early 2012.

**6.2.2. Variational Data Assimilation and Control of the Boundary Conditions**

**Participant:** Eugène Kazantsev.

A variational data assimilation technique applied to the identification of the optimal discretization of interpolation operators and derivatives in the nodes adjacent to the boundary of the domain is discussed in the context of the shallow water model. It was shown in [8] that control of approximation of boundary derivatives and interpolations can increase the model’s accuracy in boundary regions and improve the solution in general. On the other hand, optimal schemes obtained in this way may not approximate derivatives in a common sense. This may lead to another model physics, violating, for example, impermeability boundary condition.

Experiments with a full non-linear shallow water model in [7] show that controlling the discretization of operators near a rigid boundary can bring the model solution close to observations as in the assimilation window and beyond the window. This type of control allows also to improve climatic variability of the model. These properties have been studied in two different configurations: an academic case of assimilation of artificially generated observational data in a square box configuration and assimilation of real observations in a model of the Black sea [30].
The sensitivity of the shallow water model in the previously described configurations has been studied in detail in [9]. It is shown in both experiments that the boundary conditions near a rigid boundary influence the solution higher than the initial conditions. This fact points out the necessity to identify optimal boundary approximation during a model development.

In order to illustrate the influence of optimal discretization of operators near the boundary we compare this influence with now classical data assimilation for identification of the optimal initial conditions of the model. The norm of the difference between the model solution and real observational data is plotted in figure 3. Observed sea surface elevation of the Black sea was assimilated during 50 days (May–June 1992) to identify optimal initial and boundary conditions. After that, models have been integrated forward for 500 days and their solutions have been compared with data. One can see that, starting from optimal initial point, the model remains close to observations during less than 100 days while optimal optimal discretization of operators allows the model to be always closer than the model with default parameters.

Adjoint models, necessary to variational data assimilation have been produced by the TAPENADE software, developed by the TROPICS team.

![Figure 3. Evolution of the difference "model–observations".](image)

### 6.2.3. A Nudging-Based Data Assimilation Method: the Back and Forth Nudging

**Participants:** Didier Auroux, Jacques Blum, Maëlle Nodet.

The Back and Forth Nudging (BFN) algorithm (see [63]) has been recently introduced for simplicity reasons, as it does not require any linearization, nor adjoint equation, or minimization process in comparison with variational schemes. Nevertheless it provides a new estimation of the initial condition at each iteration.

Previous theoretical results [65] showed that BFN was often ill-posed for viscous partial differential equations. To overcome this problem, we proposed a new version of the algorithm, which we called the Diffusive BFN [2], and which showed very promising results on one-dimensional viscous equations. Experiments on more sophisticated geophysical models, such as Shallow-Water equations and NEMO ocean model are still in progress, in collaboration with University of Nice.

### 6.2.4. Variational Data Assimilation for locally nested models.

**Participants:** Eric Blayo, Laurent Debreu, François-Xavier Le Dimet, Emilie Neveu.
The objectives are to study the mathematical formulation of variational data assimilation for locally nested models and to conduct numerical experiments for validation.

The state equations of the optimality system have been written for the general case of two embedded grids, for which several kinds of control (initial conditions, boundary conditions) have been proposed. Both one way and two way interactions have been studied. This last year, we worked on integration of non linear grid interactions in the algorithm. Additionally, the problem of specification of background error covariances matrices has been studied (see [85]).

In the ANR MSDAG project and Emilie Neveu’s PhD, we continue to work on the subject. Our main interest is on the use of multiscale optimization methods for data assimilation. The idea is to apply a multigrid algorithm to the solution of the optimization problem. The work includes the analysis of the ellipticity of the optimization problem [12], the comparison of different multigrid methods (Gauss-Newton multigrid method and Full Approximation Scheme) and specific developments for highly non linear problems. To extend previous work on Burgers equation, the Full Approximation Scheme (FAS) and the Newton Multigrid algorithm have been compared in a more complex shallow water model. The results shows good performance of the FAS and also put more interest in the design of the background error covariance matrix.

6.3. Data Assimilation for Ocean Models

6.3.1. Development of a Variational Data Assimilation System for OPA9/NEMO


We are heavily involved in the development of NEMOVAR (Variational assimilation for NEMO). From 2006, we built a working group (coordinated by A. Vidard) in order to bring together various NEMOVAR user-groups with diverse scientific interests (ranging from singular vector and sensitivity studies to specific issues in variational assimilation), and to get technical and scientific support from Inria Sophia (Automatic adjoint derivation, TROPICS project-team) and ECMWF (Parallelization). This project aimed at avoiding duplication of effort, and at developing a common NEMOVAR platform. It has led to the creation of the VODA (Variational Ocean Data Assimilation for multi scales applications) ANR project.

The project aims at delivering a common NEMOVAR platform based on NEMO platform for 3D and 4D variational assimilation. Following 2009-10 VODA activities, a fully parallel version of NEMOTAM (Tangent and Adjoint Model for NEMO) is now available for the community in the standard NEMO version. This version is based on the released 3.0 version of NEMO. Two upgrades were done to follow NEMO standard development race. As a consequence, NEMOVAR is also available for NEMO version 3.2, 3.2.1 and 3.2.2 both offer fully parallel features. The local group has developed a python interface engine (PIANO) to perform test and run on NEMO. A constant support to NEMOVAR project is ensured to deliver a focused response on dedicated issue (internal and external interaction).

We are also investigating variational data assimilation methods applied to high resolution ocean numerical models. This part of the project is now well advanced and encouraging preliminary results are available on an idealized numerical configuration of an oceanic basin (see Figure 4).

A new topic has been explored this year in the framework of VODA: data assimilation in a framework of nested models. It makes full use of the AGRIF capabilities of NEMO and follows previous work done on a toy model during the PhD thesis of E. Simon. Some early results are available with a global ocean 2 degrees configuration including a 1/2 degree zoom on the Agulhas region (see Figure 5).

As a side project we collaborate with Mercator-Ocean in order to use the adjoint to perform sensitivity analysis with the fourth of a degree global model used for the reanalysis. This collaboration that includes both heavy software developments and challenging scientific investigation, has been going on for 2 years now and is producing interesting results for both part that still need to be published.

Apart from the VODA ANR project, the NEMOVAR working group gets additional financial support by LEFE-Assimilation and the Mercator National Programs.
Figure 4. Surface relative vorticity of a 1/24th of a degree NEMO configuration

Figure 5. Temperature increment at 200m around the zoom on the Agulhas region
6.3.2. Variational data assimilation into highly nonlinear ocean models

Participants: Pierre-Antoine Bouttier, Eric Blayo, Jacques Verron.

The purpose of this study is to explore the behaviour of variational data assimilation methods in a non-linear ocean model. In an eddy-permitting or eddy-resolving ocean model, controlling mesoscale eddies activity is crucial for data assimilation methods. Our goal is to highlight the impact of these non-linearities on the assimilation system. To illustrate this, test experiments are performed with a double-gyre NEMO configuration at different resolutions (1/4°, 1/12°) which mimics Gulfstream-like behaviour in term of eddy system, and an incremental 4D-VAR formulation for the assimilation system.

First, we are mainly interested in observing the impact of the length of assimilation window on the quality of the analyzed trajectory. For that, we are doing twin experiments with 1/4° model, using simulated altimeter data, for different lengths of assimilation window. Helped by diagnoses on error scales, we also attempt to link the non-linear phenomena and error structures observed after assimilation quantitavely and qualitatively.

Then, by increasing the model resolution (and consequently mesoscale eddy activity), we bring to light the sensitivity of our assimilation system to non-linearity by repeating the same experiments on the length of assimilation window, and the same diagnoses about error structures.

6.3.3. Assimilation of Lagrangian Data

Participants: Claire Chauvin, Maëlle Nodet, Arthur Vidard.

When an observation is given at a sequence of positions along the fluid flow, then it can be defined as Lagrangian, from a mathematical point of view. From this sequence of positions (for instance the profiling drifting floats of Argo program), one can deduce important information on the stream that transports the drifters. Such an information has not yet been exploited in an operational framework, although previous works [82] have shown the interest of assimilating this new type of data.

A task of the ANR VODA has thus been defined in order to develop the tools for the variational assimilation of Lagrangian data in the context of NEMOVAR. C. Chauvin is an engineer working on this task. She first constructed the observation operator, which requires the interpolation of the velocity at any point of the domain. This interpolation operator is not linear for the general grids used in NEMO, implying heavy tangent and adjoint operators.

Tangent and adjoint procedures associated to this interpolation method have been developed, as well as the tests of these procedures on the main test configurations GYRE and ORCA2. Their implementation in NEMOVAR has required a specific application, in order to be consistent with the conventions and data structures already present in NEMOVAR. We also performed extensive numerical experiments to assess the impact of Lagrangian data assimilation, and its complementarity with other types of data, and we prepare an article to sum up the results. Preliminary results have been presented at EGU [34].

6.4. Assimilation of Image Data

6.4.1. Direct assimilation of sequences of images


At the present time the observation of Earth from space is done by more than thirty satellites. These platforms provide two kinds of observational information:

- Eulerian information as radiance measurements: the radiative properties of the earth and its fluid envelopes. These data can be plugged into numerical models by solving some inverse problems.
- Lagrangian information: the movement of fronts and vortices give information on the dynamics of the fluid. Presently this information is scarcely used in meteorology by following small cumulus clouds and using them as Lagrangian tracers, but the selection of these clouds must be done by hand and the altitude of the selected clouds must be known. This is done by using the temperature of the top of the cloud.
MOISE was the leader of the ANR ADDISA project dedicated to the assimilation of images, and is a member of its current follow-up GeoFluids (along with EPI FLUMINANCE and CLIME, and LMD, IFREMER and Météo-France).

During the ADDISA project we developed Direct Image Sequences Assimilation (DISA) and proposed a new scheme for the regularization of optical flow problems [86], [90]. Thanks to the nonlinear brightness assumption, we proposed an algorithm to estimate the motion between two images, based on the minimization of a nonlinear cost function [45]. We proved its efficiency and robustness on simulated and experimental geophysical flows [64]. As part of GeoFluids, we are investigating new ways to define distance between a couple of images. One idea was to define this distance as the norm of the apparent motion between two images. This has been done thanks to optical flow methods which turned out to need a specific parametrization for each couple of images. Another idea, currently under investigation, consists in comparing main structures within each image. This can be done using, for example, a wavelet representation of images. We are also part of TOMMI, another ANR project started mid 2011, where we are investigating the possibility to use optimal transportation based distances for images assimilation.

6.4.2. Assimilation of ocean images

**Participants:** Vincent Chabot, Maëlle Nodet, Nicolas Papadakis, Arthur Vidard.

In addition with the direct assimilation approach previously described, a particular attention has been given to the cloud occlusion and the representation of the observation errors in the context of ocean image data. Such works will be intensified with the post-doctorate Alexandros Makris that will start his activities in December. The assimilation of images (SST and chlorophyll) provided by geostationary satellites is also studied with the oceanographers of the Laboratoire des Écoulements Géophysiques et Industriels. The objective is here to take benefit from the correlation that exists between image gradients and ocean flow discontinuities that can be exhibited through the computation of Lyapunov coefficients and vectors from numerical ocean models [70].

6.5. Quantifying Uncertainty

6.5.1. Propagation of uncertainties

**Participants:** François-Xavier Le Dimet, Victor Shutyaev.

Basically, geophysical models are suffering of two types of errors:

- errors in the model itself due to approximations of physical processes and their subgrid parametrization and also errors linked to the necessary numerical discretization;
- errors in the observation because of errors of measurements and also errors due to sampling. For instance, many remote sensings observe only radiances, which are transformed into the state variables thanks to complex processes like the resolution of an inverse problem. This is, of course, a source of errors.

Estimating the propagation of errors is an important and costly (in term of computing resources) task for two reasons:

- the quality of the forecast must be estimated
- the estimation of the statistics of errors has to be included in the analysis to have an adequate norm, based on these statistics, on the forecast and also on the observation.

In the variational framework, models, observations, statistics are linked into the optimality system which can be considered as a “generalized” model containing all the available estimation. The estimation of error covariances are estimated both from the second order analysis and the Hessian of the cost function. Numerical experiments have been carried out on a non-linear model [16]. We expect to extent the numerical experiments to a semi-operational model in cooperation with ECMWF.

6.5.2. Sensitivity analysis for West African monsoon

**Participants:** Anestis Antoniadis, Céline Helbert, Clémentine Prieur, Laurence Viry.
6.5.2.1. Geophysical context

The West African monsoon is the major atmospheric phenomenon which drives the rainfall regime in Western Africa. Therefore, this is the main phenomenon in water resources over the African continent from the equatorial zone to the sub-Saharan one. Obviously, it has a major impact on agricultural activities and thus on the population itself. The causes of inter-annual spatio-temporal variability of monsoon rainfall have not yet been univocally determined. Spatio-temporal changes on the sea surface temperature (SST) within the Guinea Gulf and Saharian and Sub-Saharan Albedo are identified by a considerable body of evidences as major factors to explain it.

The aim of this study is to simulate the rainfall by a regional atmospheric model (RAM) and to analyze its sensitivity to the variability of these inputs parameters. Once precipitations from RAM are compared to several precipitation data sets we can observe that the RAM simulates the West African monsoon reasonably.

6.5.2.2. Statistical methodology

As mentioned in the previous paragraph, our main goal is to perform a sensitivity analysis for the West African monsoon. Each simulation of the regional atmospheric model (RAM) is time consuming, and we first have to think about a simplified model. We deal here with spatio-temporal dynamics, for which we have to develop functional efficient statistical tools. In our context indeed, both inputs (albedo, SST) and outputs (precipitations) are considered as time and space indexed stochastic processes. A first step consists in proposing a functional modeling for both precipitation and sea surface temperatures, based on a new filtering method. For each spatial grid point in the Gulf of Guinea and each year of observation, the sea surface temperature is measured during the active period on a temporal grid. A Karhunen-Loève decomposition is then performed at each location on the spatial grid [91]. The estimation of the time dependent eigenvalues at different spatial locations generates great amounts of high-dimensional data. Clustering algorithms become then crucial in reducing the dimensionality of such data.

Thanks to the functional clustering performed on the first principal component at each point, we have defined specific subregions in the Gulf of Guinea. On each subregion, we then choose a referent point for which we keep a prescribed number of principal components which define the basis functions. The sea surface temperature at any point in this subregion is modeled by the projection on this truncated basis. The spatial dependence is described by the coefficients of the projection. The same approach is used for precipitation. Hence for both precipitation and sea surface temperatures, we obtain a decomposition where the basis functions are functions depending on time and whose coefficients are spatially indexed and time independent.

Then, the most straightforward way to model the dependence of precipitation on sea surface temperatures is through a multivariate response linear regression model with the output (precipitation) spatially indexed coefficients in the above decomposition and the input (SST) spatially indexed coefficients being predictors. A naive approach consists in regressing each response onto the predictors separately; however it is unlikely to produce satisfactory results, as such methods often lead to high variability and over-fitting. Indeed the dimensions of both predictors and responses are large (compared to the sample size). Thus in addition to assuming that only a subset of predictors enter the model, it is also reasonable to assume that a predictor may affect only some but not all responses. By the way we take into account the complex and spatio-temporal dynamics. This work has been published in [1].

6.5.2.3. Distributed Interactive Engineering Toolbox

An important point in the study described above is that the numerical storage and processing of model inputs/outputs requires considerable computation resources. They were performed in a grid computing environment with a middleware (DIET) which takes into account the scheduling of a huge number of
computation requests, the data-management and gives a transparent access to a distributed and heterogeneous platform on the regional Grid CIMENT (http://ciment.ujf-grenoble.fr/).

Thus, a different DIET module was improved through this application. An automatic support of a data grid software (http://www.irods.org) through DIET and a new web interface designed for MAR was provided to physicians.

These works involve also partners from the INRIA project/team GRAAL for the computational approach, and from the Laboratory of Glaciology and Geophysical Environment (LGGE) for the use and interpretation of the regional atmospheric model (RAM).

6.5.3. Tracking for mesoscale convective systems

Participants: Anestis Antoniadis, Céline Helbert, Clémentine Prieur, Laurence Viry, Roukaya Keinj.

6.5.3.1. Scientific context

In this section, we are still concerned with the monsoon phenomenon in western Africa and more generally with the impact of climate change. What we propose in this study is to focus on the analysis of rainfall system monitoring provided by satellite remote sensing. The available data are micro-wave and IR satellite data. Such data allow characterizing the behaviour of the mesoscale convective systems. We wish to develop stochastic tracking models, allowing for simulating rainfall scenarios with uncertainties assessment.

6.5.3.2. Stochastic approach

The chosen approach for tracking these convective systems and estimating the rainfall intensities is a stochastic one. The stochastic modeling approach is promising as it allows developing models for which confidence in the estimates and predictions can be evaluated. The stochastic model will be used for hydro-climatic applications in West Africa. The first part of the work will consist in implementing a model developed in [88] on a test set to evaluate its performances, our ability to infer the parameters, and the meaning of these parameters. Once the model well fitted on toy cases, this algorithm should be run on our data set, and compared with previous results by [80] or by [79]. The model developed by [88] is a continuous time stochastic model to multiple target tracking, which allows in addition to birth and death, splitting and merging of the targets. The location of a target is assumed to behave like a Gaussian Process when it is observable. Targets are allowed to go undetected. Then, a Markov Chain State Model decides when the births, death, splitting or merging of targets arise. The tracking estimate maximizes the conditional density of the unknown variables given the data. The problem of quantifying the confidence in the estimate is also addressed. Roukaya Keinj started working on this topic with a two years postdoctoral position in November 2011.

6.5.4. Sensitivity analysis for forecasting ocean models

Participants: Eric Blayo, Maëlle Nodet, Clémentine Prieur, Gaëlle Chastaing, Alexandre Janon, Jean-Yves Tissot.

6.5.4.1. Scientific context

 Forecasting ocean systems require complex models, which sometimes need to be coupled, and which make use of data assimilation. The objective of this project is, for a given output of such a system, to identify the most influential parameters, and to evaluate the effect of uncertainty in input parameters on model output. Existing stochastic tools are not well suited for high dimension problems (in particular time-dependent problems), while deterministic tools are fully applicable but only provide limited information. So the challenge is to gather expertise on one hand on numerical approximation and control of Partial Differential Equations, and on the other hand on stochastic methods for sensitivity analysis, in order to develop and design innovative stochastic solutions to study high dimension models and to propose new hybrid approaches combining the stochastic and deterministic methods.
6.5.4.2. Estimating sensitivity indices

A first task is to develop tools for estimated sensitivity indices. Among various tools a particular attention was paid to FAST and its derivatives. In [89], the authors present a general way to correct a positive bias which occurs in all the estimators in random balance design method (RBD) and in its hybrid version, RBD-FAST. Both these techniques derive from Fourier amplitude sensitivity test (FAST) and, as a consequence, are faced with most of its inherent issues. And up to now, one of these, the well-known problem of interferences, has always been ignored in RBD. After presenting in which way interferences lead to a positive bias in the estimator of first-order sensitivity indices in RBD, the authors explain how to overcome this issue. They then extend the bias correction method to the estimation of sensitivity indices of any order in RBD-FAST. They also give an economical strategy to estimate all the first-order and second-order sensitivity indices using RBD-FAST.

6.5.4.3. Intrusive sensitivity analysis, reduced models

Another point developed in the team for sensitivity analysis is model reduction. To be more precise regarding model reduction, the aim is to reduce the number of unknown variables (to be computed by the model), using a well chosen basis. Instead of discretizing the model over a huge grid (with millions of points), the state vector of the model is projected on the subspace spanned by this basis (of a far lesser dimension). The choice of the basis is of course crucial and implies the success or failure of the reduced model. Various model reduction methods offer various choices of basis functions. A well-known method is called proper orthogonal decomposition” or principal component analysis”. More recent and sophisticated methods also exist and may be studied, depending on the needs raised by the theoretical study. Model reduction is a natural way to overcome difficulties due to huge computational times due to discretizations on fine grids. In [61], the authors present a reduced basis offline/online procedure for viscous Burgers initial boundary value problem, enabling efficient approximate computation of the solutions of this equation for parametrized viscosity and initial and boundary value data. This procedure comes with a fast-evaluated rigorous error bound certifying the approximation procedure. The numerical experiments in the paper show significant computational savings, as well as efficiency of the error bound. The present preprint is under review. When a metamodel is used (for example reduced basis metamodel, but also kriging, regression, ...) for estimating sensitivity indices by Monte Carlo type estimation, a twofold error appears : a sampling error and a metamodel error. Deriving confidence intervals taking into account these two sources of uncertainties is of great interest. We obtained results particularly well fitted for reduced basis metamodels [61]. Alexandre Janon obtained a best poster award on the topic [40]. An ongoing work deals also with asymptotic confidence intervals in the double limit where the sample size goes to infinity and the metamodel converges to the true model. Implementations have to be conducted on more general models such as Shallow-Water models.

6.5.4.4. Sensitivity analysis with dependent inputs

An important challenge for stochastic sensitivity analysis is to develop methodologies which work for dependent inputs. For the moment, there does not exist conclusive results in that direction. Our aim is to define an analogue of Hoeffding decomposition [75] in the case where input parameters are correlated. A PhD started in October 2010 on this topic (Gaëlle Chastaing). We obtained first results which should be submitted soon, deriving a general functional ANOVA for dependent inputs, allowing defining new variance based sensitivity indices for correlated inputs.

6.5.5. Quantification of uncertainty with Multi-fidelity computer experiments

Participants: Federico Zertuche, Céline Helbert, Anestis Antoniadis.

Propagation of uncertainties through computer codes is a hard task when dealing with heavy industrial simulators. Confidence intervals announced on predictions are often huge because of the lack of data. The context of the study here is the case of simulations when multiple levels of analysis (fast and slow) are available. In most cases the fast (but less trustworthy) and the slow (but more accurate) response values can be obtained independently. Thus, we can learn more about the response by additionally measuring the cheap function(s) on a large number of x ‘s. In most cases the relationship between cheap and expensive responses is modeled by an autoregressive Gaussian regression. This method is a natural extension of the kriging method
in the sense that to build the surrogate one performs a Gaussian regression for the cheap data and one for the difference vector defined by the autoregressive relationship. The prediction error depends on the prediction error of the cheap and expensive surrogates. We can observe that this modeling greatly improves the traditional kriging method when the actual relationship between the cheap and expensive responses is somewhat linear. On another hand, this approach gives worse results when the relation between cheap and expensive is far to be linear. Therefore some improvements must be made on the models to take into account a more precise link between the two levels fidelity of the responses. Some other additional tasks concern the associated numerical designs (must the designs be absolutely nested ?) and the allocation of resources between low and fast runs. The work is currently the object of the thesis of Federico Zertuche that has just begun in October 2011.

6.5.6. Impact of the thermodynamics and chemical kinetics parameters at different scales for the models of CO2 storage in geological media

Participants: Céline Helbert.

In collaboration with Bernard Guy and Joharivola Raveloson (Ecole des Mines de Saint-Etienne) we study the water-gas-rock interactions in the case of CO2 storage in geological environment. The focus is on the scale of observation of geochemical phenomena while taking into account the heterogeneity of the reservoir. This heterogeneity at small and large scale helps to maintain a local variability of the chemical composition of the fluid and influence reaction rates at the pore as well as at the reservoir scale. We propose to evaluate the geostatistical characteristics of local variability thanks to simulations of reactive transport on a small scale in which parameters (namely the equilibrium constants log K and the rate constant k) are perturbed to represent local processes. This contribution is the following of a precedent study of the impact of the reservoir uncertainties on the CO2 storage [72].

6.6. Inverse methods for Glaciology

6.6.1. Dating ice matrix and gas bubbles

Participants: Eric Blayo, Bénédicte Lemieux-Dudon, Habib Toye Mahamadou Kele.

Dating ice matrix and gas bubbles of ice-cores is essential to study paleoclimates. Inverse modelling implemented on 1D ice flow models is being applied for a few years to construct the ice chronology of several deep ice cores. Such a method based on a Monte Carlo algorithm was implemented under the assumption of perfect ice flow models, for one core at a time, and without including the inverse modelling of the densification models, which enables to construct the gas age scale. This approach faces three issues: 1) frequent discrepancies between core chronologies (lack of stratigraphic links between cores as data constraints), 2) frequent failure to verify relevant data constraints (perfect model assumption is too strong), and 3) frequent inconsistency between gas and ice age scales.

A new approach was proposed during the B. Lemieux-Dudon’s PhD to circumvent these restrictions. It introduces the model error in terms of correction functions on three key quantities from which one can calculate both the ice and gas age scales: a) the accumulation rate, b) the total thinning function, and c) the close-off depth in meters of ice equivalent (i.e. depth below the ice-sheet surface where the atmosphere is trapped). A variational formulation of the inverse problem is constructed. It includes several ice cores with background scenarios and paleo data as constraints, among which:

- stratigraphic links between pair of ice cores (methane, tephra, cosmogenic isotopes, etc.) to derive consistent dating between cores,
- ice and gas age markers, as well as delta-depth data (i.e., in situ depth interval between gas and ice of the same age), which enable to optimize the gas and ice age scales simultaneously.

The cost function includes covariance error matrices, and confidence intervals of the solution can be assessed. This method was applied to derive simultaneously a common age scale for the North Grip core and for the two EPICA cores (DML and DC). [76].
This method arouses some interest in the glaciological and paleo community ([68], [87]). Some further developments are however mandatory to ensure the robustness of the dating solution: (i) code optimization, (ii) diagnostics on the assimilation system, and (iii) calibration of the background error covariance matrices.

H. Toye Mahamadou Kele (a 2-year INRIA young engineer contract) joined the MOISE team to modify the code. During the first year, he implemented a shared memory parallelization of the code, and he currently works on the calibration of the covariance error matrices by implementing a posteriori diagnostics.

6.6.2. Inverse methods for large scale ice-sheet models

Participants: Bertrand Bonan, Maëlle Nodet, Catherine Ritz.

In collaboration with C. Ritz (CNRS, Laboratoire de Glaciologie et Geophysique de l’Environnement (LGGE), Grenoble), we aim to develop adjoint methods for ice cap models.

In the framework of global warming, the evolution of sea level is a major but ill-known phenomenon. It is difficult to validate the models which are used to predict the sea level elevation, because observations are heterogeneous and sparse.

Data acquisition in polar glaciology is difficult and expensive. Satellite data have a good spatial coverage, but they allow only indirect observation of the interesting data. We wish to make the most of all available data and evaluate where/when/what we have to add new observations. Sensitivity analysis, and in particular the adjoint method, allows to identify the most influential parameters and variables and can help to design the observation network.

The ANR project ADAGe started one year ago on this subject, and B. Bonan started his PhD in September 2010. During his master internship, he implemented the adjoint code of a simplified ice-sheet flow-line model, Winnie, developed by C. Ritz at LGGE. We then performed twin experiments of data assimilation. These preliminary results have been presented at two international conferences ([32], [33]).

We then implemented Ensemble Kalman Filter (EnKF) on Winnie, which we would like to compare to variational assimilation methods. Coding and testing for the EnKF are still in progress.

6.6.3. Inverse methods for full-Stokes glaciology models

Participants: Olivier Gagliardini, Maëlle Nodet, Catherine Ritz.

We are also interested in inverse modelling for another class of glaciology models, called full-Stokes models. Such a model is developed by LGGE and CSC in Finland, called Elmer/Ice. Contrary to large scale models, Elmer/Ice is based on the full Stokes equations, and no assumptions regarding aspect ratio are made, so that this model is well adapted to high resolution small scale modelling, such as glaciers (and more recently the whole Greenland ice-sheet).

In collaboration with O. Gagliardini, F. Gillet-Chaulet and C. Ritz (Laboratoire de Glaciologie et Géophysique de l’Environnement (LGGE), Grenoble), we investigated a new method to solve inverse problems for a Full-Stokes model of Groenland, which consisted in solving iteratively a sequence of Neumann and Dirichlet problems within a gradient descent algorithm. We also compared this method to an approximate variational algorithm, using the fact that the full Stokes equations are almost self-adjoint. These results have been submitted for publication and presented at EGU and AGU ([36], [37]).

With O. Gagliardini, F. Gillet-Chaulet and M. Jay-Allemand (LGGE), we also implemented these methods to study a complex phenomenon: the surge of Variegated glacier (Alaska). This glacier is indeed known to surge periodically, that is to accelerate suddenly during 1-2 years, and then come back to a quiescent phase during 10-20 years. Inverse modelling allowed us to infer changes in basal conditions from surface velocities, and to come to a better understanding of the surge phenomenon, as seen in Figure 6. These results have been published ([6] and presented at AGU ([42]).

6.7. Image processing

6.7.1. Segmentation and assimilation of medical images

Participant: Nicolas Papadakis.
Figure 6. Distribution of the basal friction parameter along the central flow line for the 25 dates of measurements. Dotted curves indicate where measured surface velocities are missing. In the legend, \(W Y_1-Y_2\) denotes the mean velocity for the winter from year \(Y_1\) to year \(Y_2\), \(S Y\) denotes the mean velocity for the summer of year \(Y\) and \(Surge Y-M_1-M_2\) the mean velocity for year \(Y\) from month \(M_1\) to month \(M_2\).

In collaboration with the Inria team MC2 of the Bordeaux-Sud-Ouest center, we investigate the application of image assimilation to medical issues. The objective is here to use MRI images in order to monitor EDP models dealing with tumor growth in lungs or brains. Using such images, we would like to define a patient specific process allowing to calibrate the numerical model with respect to the observed tumor. First works based on convex relaxation of the binary segmentation problem \([81]\) have been realized in this direction by proposing a 3D segmentation method dedicated to glioblastomas from a set of MRI brain images. The obtained automatic segmentation results are very close to specialist manual segmentations (errors of 5\%) and will be used as pseudo-observations for an assimilation system based on the numerical model describing the tumor growth. The final issue will be to define an observation operator linking images with the model in order to realize a direct assimilation.

### 6.7.2. Optimal transport

**Participants:** Maëlle Nodet, Nicolas Papadakis, Arthur Vidard.

A new activity on optimal transport has been started in collaboration with the EDP and MGMI teams of the Laboratoire Jean Kuntzmann, Grenoble and the MAP5 Laboratory, Paris, through a project funded by the ANR white program. The purpose is to define metric beween images involving the so-called Wasserstein distance. Such metric would be of particular interest in order to introduce pertinent observation operators for assimilating image data \([64]\). Other applications including image morphing and histogram equalization are also studied for image processing purposes. First results has been obtained by adding spatial regularity to the transport map computed by the Benamou-Brenier algorithm \([66]\).

### 6.7.3. Computer vision

**Participant:** Nicolas Papadakis.
In collaboration with Vicent Caselles (Pompeu Fabra University, Barcelona, Spain), different image processing works have been finalized: 3D reconstruction and novel view synthesis for soccer replays [13], stereo inpainting for 3D movie post-production [29]. Other works dedicated to object tracking in image sequences have also been proposed with Aurelie Bugeau (LABRI, Bordeaux) [14]. A main attention is now given to the problem of histogram equalization of different images [15]. Our aim is now to include spatial information on color repartition during the histogram transfer for inpainting and shadow removal purposes. A journal paper dealing with this issue has been recently accepted for publication.

6.8. Multivariate risk indicators

Participant: Clémentine Prieur.

In collaboration with Véronique Maume-Deschamps, Elena Di Bernardino (ISFA, Lyon 1) and Peggy Cenac (Université de Bourgogne), we are interested in defining and estimating new multivariate risk indicators. This is a major issue with many applications (environmental, insurance, ...). Two papers were accepted for publication and one other is submitted. The submitted one deals with the estimation of bivariate tails [56]. In [69] we propose to minimize multivariate risk indicators by using a Kiefer-Wolfowitz approach to the mirror stochastic algorithm. In [4] we present an estimation procedure for multivariate risk indicators making use of a plug-in estimation of level sets of bivariate cumulative distribution functions.

6.9. Stochastic Downscaling Method

Participant: Antoine Rousseau.

In collaboration with TOSCA (Inria Sophia-Antipolis), LMD (Ecole Polytechnique) and CETE (Clermont-Ferrand), we investigate a new method for the numerical simulation of the wind at small scales. In this work, we consider a new approach for the downscaling in CFD. The local model that we propose is inspired from S.B. Pope’s previous works on turbulence. We investigated a new numerical simulation method for the downscaling in CFD, with a strong orientation in applications to meteorology, particularly for the simulation of wind at small scales. The local model that we propose consists in modeling the fundamental equations of fluid motion by a stochastic Lagrangian model describing the behaviour of a fluid particle. Because of the both Lagrangian and stochastic nature of our model, it is discretized thanks to an interacting particle system, combining a time Euler scheme for stochastic differential equations and a Monte–Carlo approximation method. This model called SDM (Stochastic Downscaling Method) is adapted from previous works introduced by S.B. Pope [84] (see http://sdm.gforge.inria.fr/Accueil/index.en.php).

This year, we worked on the comparison of the SDM model (endowed with a physical geostrophic forcing and a wall log law) with simulations obtained with a LES method (Méso-NH code) for the atmospheric boundary layer (from 0 to 750 meters in the vertical direction), in the neutral case. This work allowed to deeply understand the contribution of each elements of the Lagrangian model in terms of the turbulence production and dissipation, we analyzed the returns of various closure parametrization approaches, including viscosity turbulent approach. We also investigated anisotropic effect, with the introduction of GLM model in SDM (see [84]), in particular the isotropic relaxation case. We gave our conclusions as a part of the final report for ADEME [58]. A paper is in preparation.

6.10. Mathematical modelling for CFD-environment coupled systems

Participant: Antoine Rousseau.

6.10.1. Minimal-time bioremediation of natural water resources

The objective of this work is to provide efficient strategies for the bioremediation of natural water resources. Based on a previous paper [74] that deals with an homogeneous resource in space (with a system of ODEs), we implement a coupled ODE-PDE system that accounts for the spatial non-homogeneity of pollution in natural resources. The main idea is to implement a Navier-Stokes model in the resource (such as a lake), with boundary conditions that correspond to the corresponding optimal discharge of a (small) bioreactor. A first mathematical model has been introduced (see [48]) and a journal paper is ready to be submitted.
6.10.2. Mathematical modelling for the confinement of lagoons

This work deals with the concept of confinement of paralic ecosystems. It is based on the recent paper [73] that presents a modelling procedure in order to compute the confinement field of a lagoon. In [59], A. Rousseau and E. Frénod (INRIA CALVI), improve the existing model in order to account for tide oscillations in any kind of geometry such as a non-rectangular lagoons with a non-flat bottom. The new model, that relies on PDEs rather than ODEs, is then implemented thanks to the finite element method. Numerical results confirm the feasibility of confinement studies thanks to the introduced model.
5. New Results

5.1. Statistical analysis of spike trains

Modern advances in neurophysiology techniques, such as two-photons imaging of calcium signals or micro-electrode arrays (MEA) electro-physiology, have made it possible to observe simultaneously the activity of large assemblies of neurons. Such experimental recordings provide a great opportunity to unravel the underlying interactions of neural assemblies and to understand how neural populations dynamically encode information. The goal of the present project is to propose to the neuroscientists community statistical methods and numerical tools to analysing the statistics of action potentials (spike trains) obtained from MEA recordings. Our work is grounded on one hand on theoretical results on Gibbs distributions in neural networks and the other hand on a C/C++ library of algorithms developed jointly with the CORTEX INRIA team, freely available at http://enas.gforge.inria.fr/. We have collaborations with several labs specialized in MEA recording from the retina: Centro Interdisciplinario de Neurociencia de Valparaiso, Universidad de Valparaiso, Chile http://www.cinv.cl/; Department of Molecular Biology and Princeton Neuroscience Institute, Princeton University, USA http://www.princeton.edu/neuroscience/; Institut de la Vision, Paris http://www.institut-vision.org/.

5.1.1. A discrete time neural network model with spiking neurons. Dynamics with noise.

Participant: Bruno Cessac [correspondent].

We provide rigorous and exact results characterizing the statistics of spike trains in a network of leaky Integrate-and-Fire neurons, where time is discrete and where neurons are submitted to noise, without restriction on the synaptic weights. We show the existence and uniqueness of an invariant measure of Gibbs type and discuss its properties. We also discuss Markovian approximations and relate them to the approaches currently used in computational neuroscience to analyse experimental spike trains statistics. This work has appeared in Journal of Mathematical Biology[17].

5.1.2. Statistics of spike trains in conductance-based neural networks: Rigorous results

Participant: Bruno Cessac [correspondent].

We consider a conductance-based neural network inspired by the generalized Integrate and Fire model introduced by Rudolph and Destexhe in 1996. We show the existence and uniqueness of a unique Gibbs distribution characterizing spike train statistics. The corresponding Gibbs potential is explicitly computed. These results hold in the presence of a time-dependent stimulus and apply therefore to non-stationary dynamics. This establishes a rigorous ground for the current investigations attempting to characterize real spike trains data with Gibbs distributions, such as the Ising-like distribution, using the maximal entropy principle. This work has appeared in Journal of Mathematical Neuroscience[18].


Participants: Bruno Cessac [correspondent], Adrian Palacios [Centro de Neurociencia, Valparaiso, Chile].

This work focuses on methods from statistical physics and probability theory allowing the analysis of spike trains in neural networks. Taking as an example the retina we present recent works attempting to understand how retina ganglion cells encode the information transmitted to the visual cortex via the optical nerve, by analyzing their spike train statistics. We compare the maximal entropy models used in the literature of retina spike train analysis to rigorous results establishing the exact form of spike train statistics in conductance-based Integrate-and-Fire neural networks. This work is submitted in “Mathematical Problems in Computational Biology and Biomedicine” Springer, [54].
5.1.4. Gibbs distribution analysis of temporal correlations structure in retina ganglion cells

Participants: Michael Berry II [Department of Molecular Biology, Princeton University, USA], Bruno Cessac [correspondent], Olivier Marre, Adrian Palacios [Centro de Neurociencia, Valparaiso, Chile], Juan-Carlos Vasquez.

We present a method to estimate Gibbs distributions with spatio-temporal constraints on spike trains statistics. We apply this method to spike trains recorded from ganglion cells of the salamander retina, in response to natural movies. Our analysis, restricted to a few neurons, performs more accurately than pairwise synchronization models (Ising) or the 1-time step Markov models (Marre et al. (2009)) to describe the statistics of spatio-temporal spike patterns and emphasizes the role of higher order spatio-temporal interactions. This work has been accepted in Journal of Physiology, Paris [28] (in press).

5.1.5. A Markovian event-based framework for stochastic spiking neural networks

Participants: Olivier Faugeras, Jonathan Touboul.

In spiking neural networks, the information is conveyed by the spike times, that depend on the intrinsic dynamics of each neuron, the input they receive and on the connections between neurons. In this article we study the Markovian nature of the sequence of spike times in stochastic neural networks, and in particular the ability to deduce from a spike train the next spike time, and therefore produce a description of the network activity only based on the spike times regardless of the membrane potential process. To study this question in a rigorous manner, we introduce and study an event-based description of networks of noisy integrate-and-fire neurons, i.e. that is based on the computation of the membrane potential process. We show that the firing times of the neurons in the networks constitute a Markov chain, whose transition probability is related to the probability distribution of the interspike interval of the neurons in the network. In the cases where the Markovian model can be developed, the transition probability is explicitly derived in such classical cases of neural networks as the linear integrate-and-fire models with excitatory and inhibitory interactions, for different types of synapses, possibly featuring noisy synaptic integration, transmission delays and absolute and relative refractory period. This covers most of the cases that have been investigated in the event-based description of spiking deterministic neural networks.

This work has appeared in the Journal of Computational Neuroscience [26].

5.2. Coding by spikes

Our goal here is a better understanding of the extent to which computing and modeling with spiking neuron networks might be biologically plausible and computationally efficient. Based on a thorough characterization of the main constraints on spiking neural networks dynamics this has led us to propose new algorithms to infer the structure of the network from its spike trains and to propose an FPGA implementation of spiking neural networks.

5.2.1. Reverse-engineering of spiking neural networks parameters

Participants: Bruno Cessac [correspondent], Horacio Rostro-Gonzalez, Thierry Viéville [Cortex].

We consider the deterministic evolution of a time-discretized spiking network of neurons with connection weights having delays, modeled as a discretized neural network of the generalized integrate and fire (gIF) type. The purpose is to study a class of algorithmic methods allowing to calculate the proper parameters (synaptic weights) to reproduce exactly a given spike train generated by an hidden (unknown) neural network. This problem is linear (L) if the membrane potentials are observed and LP (Linear-Programming) if only spike times are observed, in the context of gIF models. The L or LP adjustment mechanism is local to each unit and has the same structure as an "Hebbian" rule. This paradigm is easily generalizable to the design of input-output spike train transformations. This means that we have a practical method to "program" a spiking network, i.e. find a set of parameters allowing us to exactly reproduce the network output, given an input.

This work has been submitted in the Journal of Neural Engineering, 2011[25].
5.2.2. Development of FPGA-based efficient reconfigurable architectures for spiking neural networks

Participants: Bruno Cessac, Bernard Girau [INRIA Cortex], Horacio Rostro-Gonzalez, Cesar Torres-Huitzil [Information Technology Department, Polytechnic University of Victoria (UPV), Tamaulipas, Mexico], Thierry Viéville [Cortex, correspondent].

Spiking neural networks are able to perform very powerful computations with precise timed spikes. We are developing an FPGA (Field Programmable Gate Array) reconfigurable platform that enables the simulation of in silico models of spiking neural networks. Since the model is directly mapped into a FPGA device, the neural processing is accelerated and the time consumption reduced. We use VHDL and Handel-C to design the reconfigurable architecture of a discrete time Integrate-and-Fire model coded in CUDA, running on GPU. This work has been accepted in Journal of Physiology, Paris [24].

5.2.3. Towards biologically inspired image coders

Participants: Marc Antonini [Laboratoire I3S, Sophia Antipolis, France], Pierre Kornprobst, Khaled Masmoudi [Laboratoire I3S, Sophia Antipolis, France].

In [51] we presented a novel bio-inspired and dynamic coding scheme for static images. Our coder aims at reproducing the main steps of the visual stimulus processing in the mammalians retina taking into account its time behavior. The main novelty of this work is to show how to exploit the time behavior of the retina cells to ensure, in a simple way, scalability and bit allocation. To do so, our main source of inspiration has been the biologically plausible retina model Virtual Retina described in Section 4.1. Following a similar structure, our model has two stages. The first stage is an image transform which is performed by the outer layers in the retina. Here it is modelled by filtering the image with a bank of difference of Gaussians with time-delays. The second stage is a time-dependent analog-to-digital conversion which is performed by the inner layers in the retina. Thanks to its conception, our coder enables scalability and bit allocation across time. Also, compared to the JPEG standards, our decoded images do not show annoying artefacts such as ringing and block effects. As a whole, this article shows how to capture the main properties of a biological system, here the retina, in order to the design a new efficient coder.

5.3. Mean field methods

5.3.1. Noise-induced behaviors in neural mean field dynamics

Participants: Jonathan Touboul, Geoffroy Hermann, Olivier Faugeras.

The collective behavior of cortical neurons is strongly affected by the presence of noise at the level of individual cells. In order to study these phenomena in large-scale assemblies of neurons, we consider networks of firing-rate neurons with linear intrinsic dynamics and nonlinear coupling, belonging to a few types of cell populations and receiving noisy currents. Asymptotic equations as the number of neurons tends to infinity (mean field equations) are rigorously derived based on a probabilistic approach. These equations are implicit on the probability distribution of the solutions which generally makes their direct analysis difficult. However, in our case, the solutions are Gaussian, and their moments satisfy a closed system of nonlinear ordinary differential equations (ODEs), which are much easier to study than the original stochastic network equations, and the statistics of the empirical process uniformly converge towards the solutions of these ODEs. Based on this description, we analytically and numerically study the influence of noise on the collective behaviors, and compare these asymptotic regimes to simulations of the network. We observe that the mean field equations provide an accurate description of the solutions of the network equations for network sizes as small as a few hundreds of neurons. In particular, we observe that the level of noise in the system qualitatively modifies its collective behavior, producing for instance synchronized oscillations of the whole network, desynchronization of oscillating regimes, and stabilization or destabilization of stationary solutions. These results shed a new light on the role of noise in shaping collective dynamics of neurons, and gives us clues for understanding similar phenomena observed in biological networks.
5.3.2. Mean Field description of and propagation of chaos in recurrent multipopulation networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons

Participants: Javier Baladron, Diego Fasoli, Olivier Faugeras, Jonathan Touboul.

We derive the mean-field equations arising as the limit of a network of interacting spiking neurons, as the number of neurons goes to infinity. The neurons belong to a fixed number of populations and are represented either by the Hodgkin-Huxley model or by one of its simplified version, the Fitzhugh-Nagumo model. The synapses between neurons are either electrical or chemical. The network is assumed to be fully connected. The maximum conductances vary randomly. Under the condition that all neurons initial conditions are drawn independently from the same law that depends only on the population they belong to, we prove that a propagation of chaos phenomenon takes places, namely that in the mean-field limit, any finite number of neurons become independent and, within each population, have the same probability distribution. This probability distribution is solution of a set of implicit equations, either nonlinear stochastic differential equations resembling the McKean-Vlasov equations, or non-local partial differential equations resembling the McKean-Vlasov-Fokker-Planck equations. We prove the well-posedness of these equations, i.e. the existence and uniqueness of a solution. We also show the results of some preliminary numerical experiments that indicate that the mean-field equations are a good representation of the mean activity of a infinite size network, even for modest sizes. These experiment also indicate that the McKean-Vlasov-Fokker-Planck equations may be a good way to understand the mean-field dynamics through, e.g., a bifurcation analysis.

This work has been submitted for publication in the Journal of Mathematical Neuroscience [55].

5.3.3. Three applications of GPU computing in neuroscience

Participants: Javier Baladron, Olivier Faugeras.

GPUs are low cost highly parallel devices that are now not only used for graphics but also for numerical simulation. We present three applications of a computer system with multiple GPUs to the domain of theoretical neuroscience. The first application is to a continuous model of the primary visual area, the second to the simulation of a stochastic neural network, and the third to the computation of the probability distribution on the possible states of a network. In all three cases we show that the speed-up obtained by the use of GPUs has considerably helped answering a scientific or technological question.

This work has been accepted for publication in Computing in Science and Engineering [63].

5.4. Neural Fields

5.4.1. Modelling the dynamics of contextual motion integration in the primate

Participants: Heiko Neumann [Institute of Neural Information Processing, Ulm University, Ulm, Germany], Pierre Kornprobst, Guillaume Masson [Institut de Neurosciences de la Timone, UMR 6193, CNRS, Marseille, France], Emilien Tlapale.

The dynamics of motion integration show striking similarities when observed at neuronal, psychophysical, and oculomotor levels. Based on the inter-relation and complementary insights given by those dynamics, our goal is to investigate how basic mechanisms of dynamical cortical processing can be incorporated in a dynamical model to solve several aspects of 2D motion integration and segmentation.

Thanks to Emilien Tlapale PhD [13] (see also [16]), we have obtained the following results:

- We proposed a recurrent model of motion integration. Proposing a simple readout mechanism, we reproduced not only motion perception but also the dynamics of smooth pursuit eye movements on various line figures and gratings viewed through different apertures. Our model can also solve various contextual problems where extrinsic junctions should be eliminated, without relying on complex junction detectors or depth computation [71]. Finally, we have also shown how our model can be rewritten in the neural fields formalism (see [52] and the Software MotionLib), which has opened new perspectives as detailed in Section 5.4.2.
We confronted our results to artificial and biological vision. To formalize the comparison against visual performance, we proposed a new evaluation methodology based on human visual performance by establishing a database of image sequences taken from biology and psychophysics literature [70], [69], [67]. We compared our results against the state of the art computer vision approaches and we found that our model also gives results comparable to recent computer vision approaches of motion estimation.

5.4.2. Neural fields models for motion integration: Characterising the dynamics of multi-stable visual motion stimuli

Participants: Olivier Faugeras, Pierre Kornprobst, Guillaume Masson [Institut de Neurosciences de la Timone, UMR 6193, CNRS, Marseille, France], Andrew Meso [Institut de Neurosciences de la Timone, UMR 6193, CNRS, Marseille, France], James Rankin, Emilien Tlapale, Romain Veltz.

In [57] we investigated the temporal dynamics of the neural processing of a multi-stable visual motion stimulus with two complementary approaches: psychophysical experiments and mathematical modeling. The so-called “barber pole” stimulus is considered with an aperture configuration that supports horizontal (H), diagonal (D) or vertical (V) perceived directions for the same input. The phenomenon demonstrates an interesting variable and dynamic competition for perceptual dominance between underlying neural representations of the three directions. We probe the early processing from stimulus presentation to initial perceived direction (before perceptual reversals). Starting from a simplified neural fields model inspired from [13], we constructed a model of the necessary motion integration that shows a shift in perceptual dominance from D to either H or V with increasing duration. Further, the timing of this shift is shown to be controlled by a stimulus gain parameter analogous to contrast. In psychophysics experiments with concurrent eye movement recordings, observers report their perceived direction of motion for presentation durations between 0.1s and 0.5s. There is a also consistent transition in perceptual dominance from D to H/V as duration is increased. This trend, seen in both perceived direction decisions and eye movement patterns, is consistent with previous experiments using similar stimuli with an aperture configured for two (D/H) rather than three (D/H/V) states. The basic dynamic properties of the early transition from D to H/V are well predicted by the model. The experimental work additionally reveals asymmetric data patterns that guide adjustments to the model’s input equations. Observers have an H bias over V, which is also reflected in faster reaction times for H. In order to capture the bias between H and V a separate weighting is attributed to the local input corresponding to each state. The work presented forms a solid foundation for future experimental and modelling work investigating the longer term dynamics for which perceptual reversals are known to occur.

5.4.3. Analysis of a hyperbolic geometric model for visual texture perception

Participants: Pascal Chossat, Grégory Faye, Olivier Faugeras.

We study the neural field equations introduced by Chossat and Faugeras in [64] to model the representation and the processing of image edges and textures in the hypercolumns of the cortical area V1. The key entity, the structure tensor, intrinsically lives in a non-Euclidean, in effect hyperbolic, space. Its spatio-temporal behaviour is governed by nonlinear integro-differential equations defined on the Poincaré disc model of the two-dimensional hyperbolic space. Using methods from the theory of functional analysis we show the existence and uniqueness of a solution of these equations. In the case of stationary, i.e. time independent, solutions we perform a stability analysis which yields important results on their behavior. We also present an original study, based on non-Euclidean, hyperbolic, analysis, of a spatially localised bump solution in a limiting case. We illustrate our theoretical results with numerical simulations.

This work has been published in the Journal of Mathematical Neuroscience [21].

5.4.4. Bifurcation of Hyperbolic Planforms

Participants: Pascal Chossat, Grégory Faye, Olivier Faugeras.
Motivated by a model for the perception of textures by the visual cortex in primates, we analyze the bifurcation of periodic patterns for nonlinear equations describing the state of a system defined on the space of structure tensors, when these equations are further invariant with respect to the isometries of this space. We show that the problem reduces to a bifurcation problem in the hyperbolic plane $D$ (Poincaré disc). We make use of the concept of a periodic lattice in $D$ to further reduce the problem to one on a compact Riemann surface $D/\Gamma$, where $\Gamma$ is a cocompact, torsion-free Fuchsian group. The knowledge of the symmetry group of this surface allows us to use the machinery of equivariant bifurcation theory. Solutions which generically bifurcate are called “H-planforms”, by analogy with the “planforms” introduced for pattern formation in Euclidean space. This concept is applied to the case of an octagonal periodic pattern, where we are able to classify all possible H-planforms satisfying the hypotheses of the Equivariant Branching Lemma. These patterns are, however, not straightforward to compute, even numerically, and in the last section we describe a method for computation illustrated with a selection of images of octagonal H-planforms.

This work has been published in the Journal of Nonlinear Science [19].

5.4.5. Bifurcation diagrams and heteroclinic networks of octagonal H-planforms

Participants: Grégory Faye, Pascal Chossat [correspondent].

This paper completes the classification of bifurcation diagrams for H-planforms in the Poincaré disc $D$ whose fundamental domain is a regular octagon. An H-planform is a steady solution of a PDE or integro-differential equation in $D$, which is invariant under the action of a lattice subgroup $\Gamma$ of $U(1,1)$, the group of isometries of $D$. In our case $\Gamma$ generates a tiling of $D$ with regular octagons. This problem was introduced in [19] as an example of spontaneous pattern formation in a model of image feature detection by the visual cortex where the features are assumed to be represented in the space of structure tensors. Under “generic” assumptions the bifurcation problem reduces to an ODE which is invariant by an irreducible representation of the group of automorphisms $\mathcal{G}$ of the compact Riemann surface $D/\Gamma$. The irreducible representations of $\mathcal{G}$ have dimension one, two, three and four. The bifurcation diagrams for the representations of dimension less than four have already been described and correspond to already well known group actions. In the present work we compute the bifurcation diagrams for the remaining three irreducible representations of dimension four, thus completing the classification. In one of these cases, there is generic bifurcation of a heteroclinic network connecting equilibria with two different orbit types.

This work has been accepted for publication in the Journal of Nonlinear Science [22].

5.4.6. Hopf bifurcation curves in neural field networks with space-dependent delays

Participant: Romain Veltz.

We give an analytical parametrization of the curves of purely imaginary eigenvalues in the delay-parameter plane of the linearized neural field network equations with space-dependent delays. In order to determine if the rightmost eigenvalue is purely imaginary, we have to compute a finite number of such curves; the number of curves is bounded by a constant for which we give an expression. The Hopf bifurcation curve lies on these curves.

This work has appeared in the Comptes Rendus Mathématiques de l’Académie des Sciences [30].

5.4.7. Stability of the stationary solutions of neural field equations with propagation delays

Participants: Olivier Faugeras, Romain Veltz.

We consider neural field equations with space-dependent delays. Neural fields are continuous assemblies of mesoscopic models arising when modeling macroscopic parts of the brain. They are modeled by nonlinear integro-differential equations. We rigorously prove, for the first time to our knowledge, sufficient conditions for the stability of their stationary solutions. We use two methods 1) the computation of the eigenvalues of the linear operator defined by the linearized equations and 2) the formulation of the problem as a fixed point problem. The first method involves tools of functional analysis and yields a new estimate of the semigroup of the previous linear operator using the eigenvalues of its infinitesimal generator. It yields a sufficient condition for stability which is independent of the characteristics of the delays. The second method allows us to find
new sufficient conditions for the stability of stationary solutions which depend upon the values of the delays. These conditions are very easy to evaluate numerically. We illustrate the conservativeness of the bounds with a comparison with numerical simulation.

This work has appeared in the Journal of Mathematical Neuroscience [29].

5.4.8. Neural Mass Activity, Bifurcations and Epilepsy

Participants: Patrick Chauvel [INSERM U751, Marseille, Assistance Publique-Hopitaux de Marseille Timone, and Universite Aix-Marseille, Marseille], Olivier Faugeras, Jonathan Touboul, Fabrice Wendling [INSERM, U642, Rennes].

We propose a general framework for studying neural mass models defined by ordinary differential equations. By studying the bifurcations of the solutions to these equations and their sensitivity to noise we establish an important relation, similar to a dictionary, between their behaviors and normal and pathological, especially epileptic, cortical patterns of activity. We then apply this framework to the analysis of two models that feature most phenomena of interest, the Jansen and Rit model, and the slightly more complex model recently proposed by Wendling and Chauvel. This model-based approach allows to test various neurophysiological hypotheses on the origin of pathological cortical behaviors and to investigate the effect of medication. We also study the effects of the stochastic nature of the inputs which gives us clues about the origins of such important phenomena as interictal spikes, inter-ical bursts and fast onset activity, that are of particular relevance in epilepsy.

This work has appeared in Neural Computation [27].
NUMED Project-Team (section vide)
6. New Results

6.1. A supervised clustering approach for fMRI-based inference of brain states

We propose a method that combines signals from many brain regions observed in functional Magnetic Resonance Imaging (fMRI) to predict the subject’s behavior during a scanning session. Such predictions suffer from the huge number of brain regions sampled on the voxel grid of standard fMRI data sets: the curse of dimensionality. Dimensionality reduction is thus needed, but it is often performed using a univariate feature selection procedure, that handles neither the spatial structure of the images, nor the multivariate nature of the signal. By introducing a hierarchical clustering of the brain volume that incorporates connectivity constraints, we reduce the span of the possible spatial configurations to a single tree of nested regions tailored to the signal. We then prune the tree in a supervised setting, hence the name supervised clustering, in order to extract a parcellation (division of the volume) such that parcel-based signal averages best predict the target information. Dimensionality reduction is thus achieved by feature agglomeration, and the constructed features now provide a multi-scale representation of the signal. Comparisons with reference methods on both simulated and real data show that our approach yields higher prediction accuracy than standard voxel-based approaches. Moreover, the method infers an explicit weighting of the regions involved in the regression or classification task. See also [14] and Fig. 1.

![Figure 1. Results for prediction of object size. Maps of weights found by supervised cut in the prediction of the size of an object. The proposed algorithm creates very interpretable clusters, compared to the reference methods that do not consider the spatial structure of the image.](image)

6.2. Multiclass Sparse Bayesian Regression for fMRI-Based Prediction

Inverse inference has recently become a popular approach for analyzing neuroimaging data, by quantifying the amount of information contained in brain images on perceptual, cognitive, and behavioral parameters. As it outlines brain regions that convey information for an accurate prediction of the parameter of interest, it allows to understand how the corresponding information is encoded in the brain. However, it relies on a prediction function that is plagued by the curse of dimensionality, as there are far more features (voxels) than samples (images), and dimension reduction is thus a mandatory step. We introduce in this work a new model, called Multiclass Sparse Bayesian Regression (MCBR), that, unlike classical alternatives, automatically adapts the amount of regularization to the available data. MCBR consists in grouping features into several classes and then regularizing each class differently in order to apply an adaptive and efficient regularization. We detail these framework and validate our algorithm on simulated and real neuroimaging data sets, showing that it performs better than reference methods while yielding interpretable clusters of features. See also [13] and Fig. 2.
Figure 2. Mental representation of size - Inter-subject analysis. Histogram of the weights found by Gibbs-MCBR, and corresponding class membership values (each color of dots represents a different class), for the inter-subject analyzes on the mental representation of size. We can see that Gibbs-MCBR creates clusters of informative and non-informative voxels, and that the different classes are regularized differently, according to the relevance of the features within them.

6.3. Total variation regularization for fMRI-based prediction of behaviour

While medical imaging typically provides massive amounts of data, the extraction of relevant information for predictive diagnosis remains a difficult challenge. Functional MRI (fMRI) data, that provide an indirect measure of task related or spontaneous neuronal activity, are classically analyzed in a mass-univariate procedure yielding statistical parametric maps. This analysis framework disregards some important principles of brain organization: population coding, distributed and overlapping representations. Multivariate pattern analysis, i.e., the prediction of behavioural variables from brain activation patterns better captures this structure. To cope with the high dimensionality of the data, the learning method has to be regularized. However, the spatial structure of the image is not taken into account in standard regularization methods, so that the extracted features are often hard to interpret. More informative and interpretable results can be obtained with the $\ell_1$ norm of the image gradient, a.k.a. its Total Variation (TV), as regularization. We apply for the first time this method to fMRI data, and show that TV regularization is well suited to the purpose of brain mapping while being a powerful tool for brain decoding. Moreover, this article presents the first use of TV regularization for classification. See also [15] and Fig. 3.

6.4. Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom.

As it provides the only method for mapping white matter fibers in vivo, diffusion MRI tractography is gaining importance in clinical and neuroscience research. However, despite the increasing availability of different diffusion models and tractography algorithms, it remains unclear how to select the optimal fiber reconstruction method, given certain imaging parameters. Consequently, it is of utmost importance to have a quantitative comparison of these models and algorithms and a deeper understanding of the corresponding strengths and weaknesses. In this work, we use a common dataset with known ground truth and a reproducible methodology to quantitatively evaluate the performance of various diffusion models and tractography algorithms. To examine a wide range of methods, the dataset, but not the ground truth, was released to the public for evaluation in a contest, the "Fiber Cup". 10 fiber reconstruction methods were evaluated. The results provide evidence
Figure 3. Regression - Sizes prediction experiment - Inter-subject analysis. Maps of weights found by TV regression for various values of the regularization parameter $\lambda$. When $\lambda$ decreases, the TV regression algorithm creates different clusters of weights with constant values. These clusters are easily interpretable, compared to voxel-based map (see below). The TV regression algorithm is very stable for different values of $\lambda$. 
that: 1. For high SNR datasets, diffusion models such as (fiber) orientation distribution functions correctly model the underlying fiber distribution and can be used in conjunction with streamline tractography, and 2. For medium or low SNR datasets, a prior on the spatial smoothness of either the diffusion model or the fibers is recommended for correct modelling of the fiber distribution and proper tractography results. The phantom dataset, the ground truth fibers, the evaluation methodology and the results obtained so far will remain publicly available on http://www.lnao.fr See also [10].

6.5. Multi-subject dictionary learning (MSDL) to segment an atlas of brain spontaneous activity

Fluctuations in brain on-going activity can be used to reveal its intrinsic functional organization. To mine this information, we give a new hierarchical probabilistic model for brain activity patterns that does not require an experimental design to be specified. We estimate this model in the dictionary learning framework, learning simultaneously latent spatial maps and the corresponding brain activity time-series. Unlike previous dictionary learning frameworks, we introduce an explicit difference between subject-level spatial maps and their corresponding population-level maps, forming an atlas. We give a novel algorithm using convex optimization techniques to solve efficiently this problem with non-smooth penalties well-suited to image denoising. We show on simulated data that it can recover population-level maps as well as subject specificities. On resting-state fMRI data, we extract the first atlas of spontaneous brain activity and show how it defines a subject-specific functional parcellation of the brain in localized regions. See also [25] and Fif 4 .

Figure 4. Outlines at 33% of all dictionary elements estimated by MSDL for 2 different set of cutting planes. The motor system is divided in (1) dorsal, (2) lateral, and (3) ventral regions. Similarly, the visual system is divided in (4) a primary region centered on the Calcarine sulcus, overlapping with (5) a region centered on the striate cortex, and (6) extrastriate regions. (7), (8): fine details of the vascular system segmented in several maps.

6.6. Functional brain imaging with M/EEG using structured sparsity in time-frequency dictionaries
Magnetoencephalography (MEG) and electroencephalography (EEG) allow functional brain imaging with high temporal resolution. While time-frequency analysis is often used in the field, it is not commonly employed in the context of the ill-posed inverse problem that maps the MEG and EEG measurements to the source space in the brain. In this work, we detail how convex structured sparsity can be exploited to achieve a principled and more accurate functional imaging approach. Importantly, time-frequency dictionaries can capture the non-stationary nature of brain signals and state-of-the-art convex optimization procedures based on proximal operators allow the derivation of a fast estimation algorithm. We compare the accuracy of our new method to recently proposed inverse solvers with help of simulations and analysis of real MEG data. See also [22].

6.7. A probabilistic framework to infer brain functional connectivity from anatomical connections

We present a novel probabilistic framework to learn across several subjects a mapping from brain anatomical connectivity to functional connectivity, i.e. the covariance structure of brain activity. This prediction problem must be formulated as a structured-output learning task, as the predicted parameters are strongly correlated. We introduce a model selection framework based on cross-validation with a parametrization-independent loss function suitable to the manifold of covariance matrices. Our model is based on constraining the conditional independence structure of functional activity by the anatomical connectivity. Subsequently, we learn a linear predictor of a stationary multivariate autoregressive model. This natural parameterization of functional connectivity also enforces the positive-definiteness of the predicted covariance and thus matches the structure of the output space. Our results show that functional connectivity can be explained by anatomical connectivity on a rigorous statistical basis, and that a proper model of functional connectivity is essential to assess this link. See also [20] and Fig. 5.

6.8. M/EEG source reconstruction based on Gabor thresholding in the source space

Thanks to their high temporal resolution, source reconstruction based on Magnetoencephalography (MEG) and/or Electroencephalography (EEG) is an important tool for noninvasive functional brain imaging. Since the MEG/EEG inverse problem is ill-posed, inverse solvers employ priors on the sources. While priors are generally applied in the time domain, the time-frequency (TF) characteristics of brain signals are rarely employed as a spatio-temporal prior. In this work, we present an inverse solver which employs a structured sparse prior formed by the sum of $\ell_2$ and $\ell_1$ norms on the coefficients of the Gabor TF decomposition of the

![Figure 5. Identifying structural connections associated with the default mode network. With yellow is represented the lateral parietal cortex, green areas represent the posterior cingulate gyrus (PCC), blue and light blue represent the medial prefrontal and orbito-frontal areas, respectively. The right model performs much better in terms of cross-validated data likelihood.](image-url)
source activations. The resulting convex optimization problem is solved using a first-order scheme based on proximal operators. We provide empirical evidence based on EEG simulations that the proposed method is able to recover neural activations that are spatially sparse, temporally smooth and non-stationary. We compare our approach to alternative solvers based also on convex sparse priors, and demonstrate the benefit of promoting sparse Gabor decompositions via a mathematically principled iterative thresholding procedure. See also [24].

6.9. Multifractal Analysis of Resting State Networks in Functional MRI

It has been known for at least one decade that functional MRI time series display long-memory properties, such as power-law scaling in the frequency spectrum. Concomitantly, multivariate model-free analysis of spatial patterns, such as spatial Independent Component Analysis (sICA), has been successfully used to segment from spontaneous activity Resting-State Networks (RSN) that correspond to known brain function. As recent neuroscientific studies suggest a link between spectral properties of brain activity and cognitive processes, a burning question emerges: can temporal scaling properties offer new markers of brain states encoded in these large scale networks? In this work, we combine two recent methodologies: group-level canonical ICA for multi-subject segmentation of brain network, and wavelet leader-based multifractal formalism for the analysis of RSN scaling properties. We identify the brain networks that elicit self-similarity or multifractality and explore which spectral properties correspond specifically to known functionally relevant processes in spontaneous activity. See also [19].

6.10. Multi-scale Mining of fMRI Data with Hierarchical Structured Sparsity

Inverse inference, or "brain reading", is a recent paradigm for analyzing functional magnetic resonance imaging (fMRI) data, based on pattern recognition tools. By predicting some cognitive variables related to brain activation maps, this approach aims at decoding brain activity. Inverse inference takes into account the multivariate information between voxels and is currently the only way to assess how precisely some cognitive information is encoded by the activity of neural populations within the whole brain. However, it relies on a prediction function that is plagued by the curse of dimensionality, as we have far more features than samples, i.e., more voxels than fMRI volumes. To address this problem, different methods have been proposed. Among them are univariate feature selection, feature agglomeration and regularization techniques. In this work, we consider a hierarchical structured regularization. Specifically, the penalization we use is constructed from a tree that is obtained by spatially constrained agglomerative clustering. This approach encodes the spatial prior information in the regularization process, which makes the overall prediction procedure more robust to inter-subject variability. We test our algorithm on a real data acquired for studying the mental representation of objects, and we show that the proposed algorithm yields better prediction accuracy than reference methods. See also [29] and Fig. 6.

6.11. Detecting Outlying Subjects in High-Dimensional Neuroimaging Datasets with Regularized Minimum Covariance Determinant

Medical imaging datasets used in clinical studies or basic research often comprise highly variable multi-subject data. Statistically-controlled inclusion of a subject in a group study, i.e. deciding whether its images should be considered as samples from a given population or whether they should be rejected as outlier data, is a challenging issue. While the informal approaches often used do not provide any statistical assessment that a given dataset is indeed an outlier, traditional statistical procedures are not well-suited to the noisy, high-dimensional, settings encountered in medical imaging, e.g. with functional brain images. In this work, we modify the classical Minimum Covariance Determinant approach by adding a regularization term, that ensures that the estimation is well-posed in high-dimensional settings and in the presence of many outliers. We show on simulated and real data that outliers can be detected satisfactorily, even in situations where the number of dimensions of the data exceeds the number of observations. See also [21] and Fig. 7.
Figure 6. Principle of structured sparsity: Example of a tree $\mathcal{T}$ when $p = 5$, with three voxels and two parcels. The parcel 2 is defined as the averaged intensity of the voxels $\{1, 2\}$, while the parcel 1 is obtained by averaging the parcel 2 and voxel 3. In red dashed lines are represented the five groups of variables that compose $\mathcal{S}$. If the group containing the parcel 2 is set to zero, the voxels $\{1, 2\}$ are also (and necessarily) zeroed out.

Figure 7. Regularized-MCD-based Mahalanobis distances of a small sample. The higher the Mahalanobis distance, the higher the probability for an observation to be tagged as outlying. Points in red are outliers subjects according to the whole population.
6.12. Connectivity-informed fMRI Activation Detection

A growing interest has emerged in studying the correlation structure of spontaneous and task-induced brain activity to elucidate the functional architecture of the brain. In particular, functional networks estimated from resting state (RS) data were shown to exhibit high resemblance to those evoked by stimuli. Motivated by these findings, we propose a novel generative model that integrates RS-connectivity and stimulus-evoked responses under a unified analytical framework. Our model permits exact closed-form solutions for both the posterior activation effect estimates and the model evidence. To learn RS networks, graphical LASSO and the oracle approximating shrinkage technique are deployed. On a cohort of 65 subjects, we demonstrate increased sensitivity in fMRI activation detection using our connectivity-informed model over the standard univariate approach. Our results thus provide further evidence for the presence of an intrinsic relationship between brain activity during rest and task, the exploitation of which enables higher detection power in task-driven studies. See also [23] and Fig 8.

Figure 8. Real data results. (a) rate of parcels with significant activation differences averaged across contrasts vs. p-value thresholds. (b) Parcels detected by contrasting computation against sentence processing task, and (c) auditory against visual task. Red = detected by only OAS-CM. Purple = detected by both OAS-CM and GL-CM. Blue = detected by all methods.

6.13. Beyond brain reading: identify and predict with clustering and randomized sparsity

The prediction of behavioral covariates from functional MRI (fMRI) is known as brain reading. From a statistical standpoint, this challenge is a supervised learning task. The ability to predict cognitive states from new data gives a model selection criterion: prediction accuracy. While a good prediction score implies that some of the voxels used by the classifier are relevant, one cannot state that these voxels form the brain regions involved in the cognitive task. The best predictive model may have selected by chance non-informative regions, and neglected relevant regions providing duplicate information. In this contribution, we address the support identification problem. The proposed approach relies on randomization techniques which have been proved to be consistent for support recovery. To account for the spatial correlations between voxels, our approach makes use of a spatially constrained hierarchical clustering algorithm. Results are provided on simulations and a visual experiment. See Fig. 9.

6.14. Joint T1 and Brain Fiber Diffeomorphic Registration Using the Demons

Non-linear image registration is one of the most challenging tasks in medical image analysis. In this work, we propose an extension of the well-established diffeomorphic Demons registration algorithm to take into
account geometric constraints. Combining the deformation field induced by the image and the geometry, we
define a mathematically sound framework to jointly register images and geometric descriptors such as fibers
or sulcal lines. We demonstrate this framework by registering simultaneously T1 images and 50 fiber bundles
consistently extracted in 12 subjects. Results show the improvement of fibers alignment while maintaining,
and sometimes improving image registration. Further comparisons with non-linear T1 and tensor registration
demonstrate the superiority of the Geometric Demons over their purely iconic counterparts. See also [28]
and Fig. 10.

Figure 9. Results on fMRI object recognition task (face vs. house). The selected voxels are in the Fusiform Face

Figure 10. Influence of the fiber weighting term on the registration accuracy. Fibers of 11 subjects were
overlapped after registration with the Geometric Demons for three values of the fiber weighting parameter.
Corresponding fibers in different subjects share colors.
6. New Results

6.1. Mathematical and numerical analysis of fluid-structure interaction problems

Participants: Cristóbal Bertoglio Beltran, Muriel Boulakia, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Jimmy Mullaert.

- Over the last decade, the numerical simulation of incompressible fluid-structure interaction has been a very active research field and the subject of numerous works. In [19], we review some of the coupling schemes recently proposed in the literature. Some numerical results that show the effectiveness of the novel approaches are also presented.

- In [21], we propose a new class of time-marching schemes for the explicit coupling of an incompressible fluid and a general elastic solid (i.e., not necessarily thin [46] and possibly damped). We state a general energy-based stability result and illustrate the accuracy of the different variants in a numerical benchmark.

- [30]: This paper focuses on Eulerian-based algorithms for fluid-structure applications featuring large structural motions and/or deformations in the context of compressible flows. First, it presents a numerical method for treating simultaneously the fluid pressure and velocity conditions on static and dynamic embedded interfaces. This method is based on the exact solution of local, one-dimensional, fluid-structure Riemann problems. Next, it describes two consistent and conservative approaches for computing the flow-induced loads on rigid and flexible embedded structures.

- In [39], we present some issues encountered in fluid-structure interaction simulation (this text is targeted to a non-expert audience).

- In [42], we present a robust and efficient parameter estimation strategy for fluid-structure interaction problems. The method is based on a filtering algorithm restricted to the parameter space, known as the reduced order Unscented Kalman Filter. We illustrate our methodology with the estimation of the artery wall stiffness and the proximal Windkessel resistance.

- [44]: In this paper, we are interested in the three-dimensional coupling between an incompressible fluid and a rigid body. The fluid is modeled by the Navier-Stokes equations, while the solid satisfies the Newton’s laws. In the main result of the paper we prove that, with the help of a distributed control, we can drive the fluid and structure velocities to zero and the solid to a reference position provided that the initial velocities are small enough and the initial position of the structure is close to the reference position.

- In [46], we introduce a class of incremental displacement-correction schemes for the explicit coupling of a thin-structure with an incompressible fluid. We provide a general stability and convergence analysis that covers both the incremental and the non-incremental variants, and also the fully implicit case. The incremental variant with first-order extrapolation is unconditionally stable and yields optimal first-order accuracy in time.

6.2. Numerical methods for fluid mechanics and application to blood flows

Participants: Jean-Frédéric Gerbeau, Marc Thiriet, Irène Vignon-Clementel.
[13]: In this work, a virtual planning of three different surgical Fontan repairs was performed to test the predictive capability of a closed-loop multi-scale model (3D-0D), constructed based on preoperative patient-specific data. Results from this multi-scale approach showed that the pre-operative caval flows should not be used as boundary conditions in post-operative simulations. The Y-TCPC repair seemed to perform better than all other TCPC models both at rest and under exercise conditions. Further work is needed to correlate results from these simulations with clinical outcomes.

[18]: Flow reversal at an outlet, although perfectly physical, can lead to rapid numerical divergence in computational fluid dynamics. Several remedies have been proposed in the literature and are discussed in the present finite element study. The most robust one was found to be a boundary advective stabilization term. The comparison was done on simple examples, as well as realistic three-dimensional multi-branched models of blood flow.

[22] & [23]: Treatments for coarctation of the aorta (CoA) can alleviate blood pressure gradients, but long-term morbidity still exists that can be explained by altered indices of hemodynamics and biomechanics. These articles present a combination of CFD methods (physiologically realistic boundary conditions and FSI with viscoelastic tissue support) to explore these indices in untreated and treated CoA, comparing them to normal subjects under rest and exercise conditions. These studies showed in particular that CoA disturbs normal patterns of wall shear stress and oscillatory shear index throughout the thoracic aorta (potentially linked to the development of atherosclerosis) and that restoring favorable anatomy may not restore normal hemodynamics.

[24]: The objective of this work is to address the formulation of an adequate model of the external tissue environment when studying a portion of the arterial tree with fluid-structure interaction. The simulations are quantitatively assessed by detailed comparisons with dynamical medical image sequences, and the model results are shown to be in very good adequacy with the data.

[26]: The wide range of existing viscoelastic wall models may produce significantly different blood flow, pressure, and vessel deformation solutions in cardiovascular simulations. In this paper, we have successfully implemented and verified two viscoelastic wall models in a nonlinear 1D finite element blood flow solver and analyzed differences between these models in various idealized and physiological simulations, including exercise.

[28]: High-intensity focused ultrasound (HIFU) is used as a thermal ablation process to eliminate tumors in different body’s organs. Blood flow has a cooling effect. Conversely, ultrasounds are responsible for acoustic streaming. A three-dimensional acoustics-thermal-fluid coupling model is carried out to compute the temperature field a given hepatic cancerous region.

[29]: Imaged-based patient-specific models of the multi-branched pulmonary arteries and superior vena cava were built for five cavopulmonary connection (i.e. Glenn) patients prior to their third surgery to alleviate their congenital heart disease. Inflow and outflow boundary conditions for computational blood flow simulations (CFD) were constructed based on an iterative procedure to match available MRI and catheterization clinical data. Common trends and differences emerged from this three-dimensional CFD study; in particular low wall shear stress was found for all subjects, which is potentially deleterious. A sensitivity analysis was performed to investigate the impact of input data (clinical and modeling) to construct boundary conditions on several clinical and mechanical indicators. Among other findings, this study suggests that although 6-10% flow split imprecision seemed reasonable in terms of patient comparison, the common practice of imposing a right pulmonary artery/left pulmonary artery flow split of 55%/45% when performing patient specific simulations should be avoided.

[31]: A novel Y-shaped baffle has been proposed for the Fontan operation with promising initial results on idealized models or a single patient-specific model. The objective of this study is to comprehensively compare the hemodynamic performance and hepatic blood flow distribution of the Y-graft Fontan baffle with two current designs on multiple patient-specific models. Methods
include virtual geometrical design, computational fluid dynamics based on preoperative patient-specific data, particle tracking and sensitivity analysis, including rest and exercise conditions. The Y-graft Fontan design achieves overall superior hemodynamic performance compared with traditional designs. However, the results emphasize that designs should be customized for individual patients before clinical application.

6.3. Numerical methods for cardiac electrophysiology

Participants: Muriel Boulakia, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Vincent Martin, Elisa Schenone.

- [32]: We consider the problem of estimating some parameters of a model of electrocardiograms from the data of the Einthoven leads. The direct model is based on the bidomain equations in the heart and a Poisson equation in the torso. To keep the computational time reasonable, the evaluation of the direct problem is approximated with a reduced order model based on Proper Orthogonal Decomposition. The optimization problem is solved using an evolutionary algorithm. Numerical tests show that, with noisy synthetic data, the proposed procedure allows to recover ionic parameters and initial activation regions with a fair accuracy.

- [34] and [48]: In presence of a high magnetic field, the blood flow in the aorta induces an electrical potential which is responsible for an increase of the T-wave in the electrocardiogram (ECG). This phenomenon may perturb ECG-gated imaging. The aim of this numerical study is to reproduce this experimental observation through computer simulations. The proposed model consists of three components: magnetohydrodynamics (MHD) in the aorta, bidomain equations in the heart and electrical diffusion in the rest of the body. These models are strongly coupled together and solved with finite elements.

- [38]: We present an overview of our works about electrocardiogram numerical simulations.

- [45]: A reduced-order model based on Proper Orthogonal Decomposition is proposed for the bidomain equations of cardiac electrophysiology. Its accuracy is assessed through electrocardiograms in various configurations, including myocardium infarctions and long-time simulations. We show in particular that a restitution curve can efficiently be approximated by this approach. The reduced-order model is then used in an inverse problem solved by an evolutionary algorithm. Some attempts are presented to identify infarction locations from synthetic electrocardiograms.

6.4. Lung and respiration modeling

Participants: Laurent Boudin, Bérénice Grec, Muriel Boulakia, Anne-Claire Egloffé, Céline Grandmont, Ayman Moussa.

- [9]: This paper is concerned with a system that couples the incompressible Navier–Stokes equations to the Vlasov–Fokker–Planck equation. Such a system arises in the modeling of sprays, where a dense phase interacts with a disperse phase. The coupling arises from the Stokes drag force exerted by a phase on the other.

- [25]: We are concerned with the global well-posedness of a two-phase flow system arising in the modelling of fluid-particle interactions. This system consists of the Vlasov-Fokker-Planck equation for the dispersed phase (particles) coupled to the incompressible Euler equations for a dense phase (fluid) through the friction forcing.

- [49]: We obtain the Maxwell-Stefan diffusion model by studying the asymptotic behaviour of a multicomponent kinetic model when the Knudsen number goes to 0.

- [50]: We are concerned here with identifiability, stability properties and estimates for the inverse problem of identifying a Robin coefficient on some non accessible part of the boundary from available data on the other part of boundary corresponding to solutions of the Stokes equations. We first study the identifiability of Robin coefficient and then we establish a stability estimate of logarithm type using Carleman inequality.
6.5. Miscellaneous

Participant: Laurent Boudin.

- [43]: We deal with a kinetic model to describe the evolution of the opinion in a closed group where there are two opposite behaviours: conciliatory and contradictory agents. We provide an existence and uniqueness result for the model and numerically test it in some relevant cases.
6. New Results

6.1. Interactive Simulation of Liver Resection PASSPORT demonstration at SIGGRAPH

The 3-year EU PASSPORT project is being finalized. In this context, we created a GPU-based interactive simulation of laparoscopic liver resection that was selected for SIGGRAPH Real Time Live! [21]. While similar medical simulators have been developed in the past, this demonstration rely on advanced methods and the computational power of today GPUs to simulate multiple organs with high-resolution deformations and collisions in real-time. We use detailed meshes generated from segmented CT scans, facilitating the reproduction of patient-specific scenarios, as is necessary for the pre-operative rehearsal of complex or risky medical cases. In the presented application, the user can examine the mechanical and collision models, and the generated contacts while the simulated patient is breathing. He can manipulate a laparoscopic instrument to navigate through the abdominal cavity, push on organs and perform a thermal ablation.

Figure 5. Left: mechanical FEM models (4k tetrahedra), Middle: Collision surfaces (15k triangles), Right: Visual Models. Bottom: the user view during liver resection. The resulting simulation runs at 25 FPS.

While similar medical simulators have been developed in the past 10 years, this demonstration is based on recently proposed methods: high-resolution Finite Element Model (FEM) with implicit time-integration implemented on GPU [1], volume contact constraints [2], an efficient numerical solver based on asynchronous preconditioning [17], as well as improvements in visual and haptic rendering. These methods allow to simulate in real-time all organs in the abdominal cavity using a improved level of precision compared to previous works. The FEM formulation enables to reproduce specific material properties. Contacts are handled by precise constraints with frictions on detailed surface meshes. Both methods support topological changes efficiently, as demonstrated by performing a resection of a portion of a liver, an important step in
surgical procedures performed to remove cancerous tumors. This work was presented at SIGGRAPH 2011 [21].

6.2. Biomechanical simulation of electrode migration for deep brain stimulation

Deep Brain Stimulation is a modern surgical technique for treating patients who suffer from affective or motion disorders such as Parkinson’s disease. The efficiency of the procedure relies heavily on the accuracy of the placement of a micro-electrode which sends electrical pulses to a specific part of the brain that controls motion and affective symptoms. However, targeting this small anatomical structure is rendered difficult due to a series of brain shifts that take place during and after the procedure. We introduce a biomechanical simulation of the intra and postoperative stages of the procedure in order to determine lead deformation and electrode migration due to brain shift. To achieve this goal, we propose a global approach, which accounts for brain deformation but also for the numerous interactions that take place during the procedure (contacts between the brain and the inner part of the skull and falx cerebri, effect of the cerebrospinal fluid, and biomechanical interactions between the brain and the electrodes and cannula used during the procedure). Preliminary results show a good correlation between our simulations and various results reported in the literature. This work was presented at MICCAI 2011 [16].

![Figure 6. Screenshot showing the simulated deflection of the right electrode immediately after operation (left) and several weeks after the operation (right).](image)

6.3. Preconditioner-Based Contact Response and Application to Cataract Surgery

We introduced a new method to compute, in real-time, the physical behavior of several colliding soft-tissues in a surgical simulation. The numerical approach is based on finite element modeling and allows for a fast update of a large number of tetrahedral elements. The speed-up is obtained by the use of a specific preconditioner that is updated at low frequency. The preconditioning enables an optimized computation of both large deformations and precise contact response. Moreover, homogeneous and inhomogeneous tissues are simulated with the same accuracy. This method was used in a simulation of one step in a cataract surgery procedure, which require to handle contacts with non homogeneous objects precisely. This work was presented at MICCAI 2011 [17].
6.4. Interactive Blood-Coil Simulation in Real-time during Aneurysm Embolization

We introduced a complete process for patient-specific simulations of coil embolization, from mesh generation with medical datasets to computation of coil-flow bilateral influence. We propose a new method for simulating the complex blood flow patterns that take place within the aneurysm, and for simulating the interaction of coils with this flow. Porous media was introduced to model the impact of the coil onto the flow (as a change of flow pattern and a decrease of velocity) from a statistical point of view, while the reverse effect on the coil (as a shift in the blood flow) was described by the local drag force. By solving the Navier-Stokes Equations with extra porous terms using the DEC method, the velocity field of blood flow was obtained, and then used to compute the drag force applied on the coil during aneurysm embolization. This work was published in the Computers & Graphics journal [13].

6.5. Constraint-based Haptic Rendering of Multirate Compliant Mechanisms
The research, that is published in IEEE transaction of Haptics [12], is dedicated to haptic rendering of complex physics-based environment in the context of surgical simulation. A new unified formalism for modeling the mechanical interactions between medical devices and anatomical structures and for computing accurately the haptic force feedback is presented. The approach deals with the mechanical interactions using appropriate force and/or motion transmission models named compliant mechanisms. These mechanisms are formulated as a constraint-based problem that is solved in two separate threads running at different frequencies. The first thread processes the whole simulation including the soft-tissue deformations, whereas the second one only deals with computer haptics. This method builds a bridge between the so called virtual mechanisms (that were proposed for haptic rendering of rigid bodies) and intermediate representations (used for rendering of complex simulations). With this approach, it is possible to describe the specific behavior of various medical devices while relying on a unified method for solving the mechanical interactions between deformable objects and haptic rendering. The technique is demonstrated in interactive simulation of flexible needle insertion through soft anatomical structures with force feedback.

![Figure 9. Constraint-based Haptic Rendering of Multirate Compliant Mechanisms](image)

### Figure 9. Constraint-based Haptic Rendering of Multirate Compliant Mechanisms

6.6. Asynchronous Haptic Simulation of Contacting Deformable Objects with Variable Stiffness
This research, published in IROS proceedings [18], presents a new asynchronous approach for haptic rendering of deformable objects. When stiff non-linear deformations take place, they introduce important and rapid variations of the force sent to the user. This problem is similar to the stiff virtual wall for which a high refresh rate is required to obtain a stable haptic feedback. However, when dealing with several interacting deformable objects, it is usually impossible to simulate all objects at high rates. To address this problem we propose a quasi-static framework that allows for stable interactions of asynchronously computed deformable objects. In the proposed approach, a deformable object can be computed at high refresh rates, while the remaining deformable virtual objects remain computed at low refresh rates. Moreover, contacts and other constraints between the different objects of the virtual environment are accurately solved using a shared Linear Complementarity Problem (LCP). Finally, we demonstrate our method on two test cases: a snap-in example involving non-linear deformations and a virtual thread interacting with a deformable object.

![Figure 10. Schematic visualization of the computational model](image-url)
6. New Results

6.1. Parallelism and convergence in Krylov methods

**Participants:** Édouard Canot, Jocelyne Erhel, Désiré Nuentsa Wakam, Bernard Philippe.

This work is done in the context of the Cinemas2 and the Libraero contracts, 7.2 and 8.1.3. It is also done in collaboration with the joint INRIA/ NCSA laboratory on petascale computing.

A Ph-D thesis was defended this year [12].

6.1.1. Some properties of Krylov methods

**Participant:** Jocelyne Erhel.

A survey was presented at a conference and published in a book chapter [37] [24].

Solving a linear system is at the heart of many scientific and engineering applications. Generally, this operation is the most time and memory consuming part of the simulation. This paper focuses on some properties of Krylov iterative methods. Iterative methods of Krylov type require less memory than direct methods, but the number of iterations increases rapidly with the size of the system. The convergence rate and the accuracy of the results depend on the condition number which can blow up at large scale. Therefore, it is essential to combine these methods with a preconditioner; the idea is to solve another system, close to the original one, but which is easier to solve; also, on parallel computers, it must be scalable. In Krylov iterative methods, the matrix is not transformed but the kernel operation is the matrix-vector product; thus it is possible to use matrix-free versions without storing the matrix. However, preconditioning will sometimes require the matrix. Krylov methods are described in many books. In this survey, we choose the framework of polynomial and projection methods. We first give general properties. Then, we study specific methods for the three different types of matrices: the case of SPD matrices is analyzed first, followed by the case of symmetric indefinite matrices. The general case of nonsymmetric matrices is studied with the description of several Krylov methods. Finally, some practical issues, preconditioning and parallelism are discussed.

6.1.2. Generation of Krylov subspace bases

**Participant:** Bernard Philippe.

This work was done in collaboration with L. Reichel, from University of Kent, USA.

It has been published in a journal [19].

Many problems in scientific computing involving a large sparse square matrix A are solved by Krylov subspace methods. This includes methods for the solution of large linear systems of equations with A, for the computation of a few eigenvalues and associated eigenvectors of A, and for the approximation of nonlinear matrix functions of A. When the matrix A is non-Hermitian, the Arnoldi process commonly is used to compute an orthonormal basis for a Krylov subspace associated with A. The Arnoldi process often is implemented with the aid of the modified Gram–Schmidt method. It is well known that the latter constitutes a bottleneck in parallel computing environments, and to some extent also on sequential computers. Several approaches to circumvent orthogonalization by the modified Gram–Schmidt method have been described in the literature, including the generation of Krylov subspace bases with the aid of suitably chosen Chebyshev or Newton polynomials. We review these schemes and describe new ones. Numerical examples are presented.

6.1.3. Parallel preconditioned GMRES with Multiplicative Schwarz

**Participants:** Édouard Canot, Jocelyne Erhel, Désiré Nuentsa Wakam, Bernard Philippe.

This work was published in a journal [18].
This paper presents a robust hybrid solver for linear systems that combines a Krylov subspace method as accelerator with a Schwarz-based preconditioner. This preconditioner uses an explicit formulation associated to one iteration of the multiplicative Schwarz method. The Newton-basis GMRES, which aim at expressing a good data parallelism between subdomains is used as accelerator. In the first part of this paper, we present the pipeline parallelism that is obtained when the multiplicative Schwarz preconditioner is used to build the Krylov basis for the GMRES method. This is referred as the first level of parallelism. In the second part, we introduce a second level of parallelism inside the subdomains. For Schwarz-based preconditioners, the number of subdomains is kept small to provide a robust solver. Therefore, the linear systems associated to subdomains are solved efficiently with this approach. Numerical experiments are performed on several problems to demonstrate the benefits of using these two levels of parallelism in the solver, mainly in terms of numerical robustness and global efficiency.

6.1.4. Adaptive deflation in preconditioned GMRES algorithm using a combined preconditioning

Participants: Jocelyne Erhel, Désiré Nuentsa Wakam, Bernard Philippe.

This work has been presented at a conference and a workshop [35], [27] and submitted to the proceedings of DD20 [45]. The software module DGMRES is integrated in the Petsc distribution.

Many scientific libraries are currently based on the GMRES method as a Krylov subspace iterative method for solving large linear systems. The restarted formulation known as GMRES(m) has been extensively studied and several approaches have been proposed to reduce the negative effects due to the restarting procedure. A common effect in GMRES(m) is a slow convergence rate or a stagnation in the iterative process. In this situation, it is less attractive as a general solver in industrial applications. In this work, we propose an adaptive deflation strategy which retains useful information at the time of restart to avoid stagnation in GMRES(m) and improve its convergence rate. We give a parallel implementation in the PETSc package. The provided numerical results show that this approach can be effectively used in the hybrid direct/iterative methods to solve large-scale systems.

6.1.5. Adaptive deflation in preconditioned GMRES algorithm using an augmented subspace

Participants: Jocelyne Erhel, Désiré Nuentsa Wakam, Bernard Philippe.

This work has been presented at a conference [31] and submitted to the journal ETNA [46].

The GMRES iterative method is widely used as Krylov subspace accelerator for solving sparse linear systems when the coefficient matrix is nonsymmetric and indefinite. The Newton basis implementation has been proposed on distributed memory computers as an alternative to the classical approach with the Arnoldi process. The aim of our work here is to introduce a modification based on deflation techniques. This approach builds an augmented subspace in an adaptive way to accelerate the convergence of the restarted formulation. In our numerical experiments, we show the benefits of using this implementation with hybrid direct/iterative methods to solve large linear systems.

6.1.6. Using deflated preconditioned GMRES for industrial CFD problems

Participant: Désiré Nuentsa Wakam.

This work has been submitted to the journal Computers and Fluids [47].

This paper deals with the solution of large and sparse linear systems arising from design optimization in Computational Fluid Dynamics. From the algebraic decomposition of the input matrix, a hybrid robust direct/iterative solver is often defined with a Krylov subspace method as accelerator, a domain decomposition method as preconditioner and a direct method as subdomain solver. The goal of this paper is to reduce the memory requirements and indirectly the computational cost at different steps of this scheme. To this end, we use a grid-point induced block approach for the data storage and the partitioning part, a Krylov subspace method based on the restarted GMRES accelerated by deflation, a preconditioner formulated with the restricted additive Schwarz method and an aerodynamic/turbulent fields split at the subdomain level. Numerical results are presented with industrial test cases to show the benefits of these choices.
6.2. Parallel numerical algorithms

6.2.1. High Performance Scientific Computing

Participant: Bernard Philippe.

This work was done in collaboration with several authors, from US, Greece, etc.

A book will appear on this subject in 2012 [39] and a chapter of this book is devoted to a historical perspective [38].

This comprehensive text/reference, inspired by the visionary work of Prof. Ahmed H. Sameh, represents the state of the art in parallel numerical algorithms, applications, architectures, and system software. Articles in this collection address solutions to various challenges arising from concurrency, scale, energy efficiency, and programmability. These solutions are discussed in the context of diverse applications, ranging from scientific simulations to large-scale data analysis and mining.

As exascale computing is looming on the horizon while multicore and GPU’s are routinely used, we survey the achievements of Ahmed H. Sameh, a pioneer in parallel matrix algorithms [38]. Studying his contributions since the days of Illiac IV as well as the work that he directed and inspired in the building of the Cedar multiprocessor and his recent research, unfolds a useful historical perspective in the field of parallel scientific computing.

6.2.2. Updating the Diagonalization of a Symmetric Matrix

Participant: Bernard Philippe.

This work is done in the context of the DIAMS project.

Two methods are compared: Jacobi method and first order correction of the spectral projectors [25],[26].

6.2.3. Counting eigenvalues in domains of the complex field

Participant: Bernard Philippe.

This work is done in collaboration with E. Kamgnia, from the University of Yaounde 1, Cameroon, in the context of the MOMAPLI project at LIRIMA.

It has been submitted to a journal [43].

A procedure for counting the number of eigenvalues of a matrix in a region surrounded by a closed curve is presented. It is based on the application of the residual theorem. The quadrature is performed by evaluating the principal argument of the logarithm of a function. A strategy is proposed for selecting a path length that insures that the same branch of the logarithm is followed during the integration. Numerical tests are reported for matrices obtained from conventional matrix test sets.

6.2.4. Rescaling for time integration

Participant: Jocelyne Erhel.

This work is done in collaboration with N. Makhoul and N. Nassif, from the American University of Beirut, Lebanon.

It is published in a journal [17].
This paper considers the mathematical framework of a sliced-time computation method for explosive solutions to systems of ordinary differential equations: $Y(t) \in \mathbb{R}^k : \frac{dY}{dt} = F(Y), \quad 0 < t, \quad Y(0) = Y_0,$ that have finite or infinite explosion time. The method used generates automatically a sequence of non uniform slices $\{[T_{n-1}, T_n] | n \geq 1\}$ determined by an end-of-slice condition that controls the growth of the solution within each slice. It also uses rescaling of the variables, whereas $t = T_{n-1} + \beta_n s$ and $Y(t) = Y(T_{n-1}) + D_n Z_n(s)$, $D_n \in \mathbb{R}^{k \times k}$ and $\beta_n$ being respectively an invertible diagonal matrix and a rescaling time factor. Thus, the original system is transformed into a sequence of slices-dependent initial-value shooting problems: $\frac{dZ_n}{ds} = G_n(Z_n), \quad 0 < s \leq s_n, \quad Z_n(0) = 0$. A suitable selection of $\beta_n$ and $D_n$ leads the rescaled systems to verify a concept of uniform similarity, allowing to disable the extreme stiffness of the original ODE problem.

Then, on each time slice, the uniformly rescaled systems are locally solved using a 4th order explicit Runge-Kutta scheme, within a computational tolerance of $\epsilon_{\text{loc}}$. After sequentially implementing the local solver on a total of $N$ slices, a global tolerance $\epsilon_{\text{glob}}$ would result in approximating the solution $Y(t)$ of the original system.

The proper definition of Uniform Similarity leads to deriving, under a stability assumption, a relationship between $\epsilon_{\text{loc}}, \epsilon_{\text{glob}}$ and $N$. Numerical experiments are conducted for infinite and finite times explosive discrete reaction diffusion problems. These experiments validate the theoretical results and attest for the efficiency of the method in terms of stability and high accuracy.

6.3. Numerical models and simulations applied to physics

6.3.1. Heat and mass transfer in soil and prehistoric fires

Participant: Édouard Canot.

This work is done in the context of the Arphymat project, in collaboration with Archeosciences, IPR and Lebanese International University (LIU), Lebanon.

This work is published in a journal [16].

This paper is devoted to the simulation of water forced evaporation in a porous saturated medium in a 3D-axisymmetric domain by resolution of partial differential algebraic equations (PDAE) that are encountered in different engineering applications. The goal of this paper is an attempt to present effective realizations, in order to determine the minimal duration of burning for prehistoric occupations. This multidisciplinary work includes scientists in Mathematics, Physics and Archaeology. The model proposed here couples the heat conduction in a water saturated soil with the water steam flow in the medium. We propose an efficient and robust global numerical method, based on a method of lines and differential algebraic equations (DAE) solvers, combined with a Newton method using a powerful sparse linear solver. After a brief overview of classes for numerical techniques applied for moving boundary problems, the Apparent Heat Capacity method (AHC) is used, and in order to validate our codes, a comparison with experiments is done.

Recent work concerns the optimal choice of the temperature interval across which the phase change occurs in the apparent capacity method, because we have to make a compromise between the smoothness of the solution and its accuracy.

6.3.2. Rheology of granular systems flowing out of silo

Participant: Édouard Canot.

This work is done in the framework of a project funded by the Region Bretagne. A PhD thesis (Merline Djouwe-Tankeo), coadvised with Patrick Richard, who is from the Physics Institute at the University of Rennes (IPR), started in February 2009 and will be defended in January 2012.

It has been presented at a conference [36] and a paper is submitted.
We first studied the granular flows by the "discrete elements" method in silo geometries. By changing the micro-mechanical properties of the grains (restitution and friction), we showed that they had a significant influence on the flow discharge. Although models such as "discrete elements" provide access to all the individual properties of the grains, they have one major drawback: the computation time is very important that prohibits the modeling of geophysical and industrial situations. To overcome this problem, we used the "continuous medium" approach, which consider that the granular medium studied follows a rheology recently proposed in the literature. After discussing the numerical implementation, we have studied this rheology for steady and fully developed flows with a semi-analytical method in two configurations: a shear cell and a channel. This allowed us to highlight the differences between a granular medium and a Newtonian fluid.

### 6.4. Models and simulations for transport in porous media

#### 6.4.1. Transport in highly heterogeneous porous medium

**Participants:** Jocelyne Erhel, Géraldine Pichot, Nadir Soualem.

This work is done in collaboration with A. Beaudoin, from University of Poitiers (Pprime) and J.-R. de Dreuzy, from Geosciences Rennes. It is done in the context of the Micas project (8.1.2).

It has been presented at a conference and a paper is in preparation [28].

We study the transport of an inert species in a 2D heterogeneous porous medium via a Random Walk Particle Tracking (RWPT) method. The main objective is to derive the macroscopic properties of the transport by the means of Monte-Carlo simulations in large domains. Conditions to reach asymptotic macro-dispersion coefficients are given. We also present our on-going research about the RWPT method in presence of discontinuities within the domain.

#### 6.4.2. Transport in discontinuous porous medium

**Participants:** Jocelyne Erhel, Géraldine Pichot.

This work is done in collaboration with A. Lejay, from Inria Nancy. It is done in the context of the Micas project (8.1.2).

It is published in the proceedings of a conference and submitted in a journal [30], [44].

We study a diffusion process in a 1D discontinuous medium using a random walk approach. Our main contribution is to encompass two existing numerical methods in the unified framework of the Skew Brownian Motion. This theoretical approach allows to detail and justify the derived algorithms. Numerical simulations are performed on two test cases to show that the algorithms can deal with the discontinuity in the diffusion coefficient.

#### 6.4.3. Reactive transport

**Participants:** Édouard Canot, Jocelyne Erhel, Souhila Sabit, Nadir Soualem.

This work is done in the context of the MOMAS GNR (8.1.1) and the contract with Andra (7.1).

It has been presented at a workshop and a paper is in preparation [33]. The software GRT3D (see section 5.6) is described in a report [48].

We have developed a method coupling transport and chemistry, based on a method of lines such that spatial discretization leads to a semi-discrete system of algebraic differential equations (DAE system). The main advantage is to use a complex DAE solver, which controls simultaneously the timestep and the convergence of Newton nonlinear iterations [53]. Analysis done with several numerical experiments showed that most of CPU time is spent in solving the linear systems of Newton iterations. We have reduced this computational time by reducing the size of the system; numerical experiments with large 2D domains show the efficiency.
6.5. Models and simulations for flow in fractured media

This work is done in collaboration with J.-R. de Dreuzy, from the department of Geosciences at the University of Rennes 1 (who is on leave until 2013 at UPC, Barcelona, Spain). It is done in the context of the Micas project (8.1.2).

A Ph-D thesis was defended this year [13].

6.5.1. Domain decomposition method for flow in 3D networks of fractures

Participants: Jocelyne Erhel, Baptiste Poirriez.

This work was presented at a conference and published in the proceedings of another conference [29], [32]. A paper is in preparation.

This paper aims at solving efficiently the linear system arising from flow computations in Discrete Fracture Networks (DFN). We define a partition of fractures into connected sets and apply a Schur domain decomposition method. Conjugate Gradient is preconditioned by Neumann-Neumann and deflation. Preliminary results with one network show the ability of our method to reduce both the number of iterations and the computational time.

6.5.2. Mortar method for flow in 3D networks of fractures

Participants: Jocelyne Erhel, Géraldine Pichot.

This work is published in a journal [20].

The simulation of flow in fractured media requires handling both a large number of fractures and a complex interconnecting network of these fractures. Networks considered in this paper are 3D domains made up of 2D fractures intersecting each other and randomly generated. Due to the stochastic generation of fractures, intersections can be highly intricate. The numerical method must generate a mesh and define a discrete problem for any Discrete Fracture Network (DFN). A first approach [51] is to generate a conforming mesh and to apply a mixed hybrid finite element method. However the resulting linear system becomes very large when the network contains many fractures. Hence a second approach [52] is to generate a non conforming mesh, using an independent mesh generation for each fracture. Then a Mortar technique applied to the mixed hybrid finite element method deals with the non-matching grids. When intersections do not cross nor overlap, pairwise Mortar relations for each intersection are efficient [52]. But for most of random networks, discretized intersections involve more than two fractures. In this paper, we design a new method generalizing the previous one and applicable for stochastic networks. The main idea is to combine pairwise Mortar relations with additional relations for the overlapping part. This method still ensures the continuity of fluxes and heads and still yields a symmetric positive definite linear system. Numerical experiments show the efficiency of the method applied to complex stochastic fracture networks. We also study numerical convergence when reducing the mesh step. This method makes it easy to perform mesh optimization and appears as a very promising tool to simulate flow in multiscale fracture networks.

6.6. Uncertainty quantification in hydrogeology

This work is done in collaboration with A. Debussche, from ENS-Cachan-Rennes and Ipso INRIA team. It is done in the context of the Micas project (8.1.2).

A PhD thesis was defended this year [11].

6.6.1. Strong and weak error estimates for elliptic partial differential equations with random coefficients

Participant: Julia Charrier.

This work has been presented at a workshop and is published in a journal [23], [14].
We consider the problem of numerically approximating the solution of an elliptic partial differential equation with random coefficients and homogenous Dirichlet boundary conditions. We focus on the case of a lognormal coefficient, we have then to deal with the lack of uniform coercivity and uniform boundedness with respect to the randomness. This model is frequently used in hydrogeology. We approximate this coefficient by a finite dimensional noise using a truncated Karhunen-Loève expansion. We give then estimates of the corresponding error on the solution, both a strong error estimate and a weak error estimate, that is to say an estimate of the error committed on the law of the solution. We obtain a weak rate of convergence which is twice the strong one. Besides this, we give a complete error estimate for the stochastic collocation method in this case, where neither coercivity nor boundedness are stochastically uniform. To conclude, we apply these results of strong and weak convergence to two classical cases of covariance kernel choices: the case of an exponential covariance kernel on a box and the case of an analytic covariance kernel, yielding explicit weak and strong convergence rates.

6.6.2. Numerical analysis of a multilevel Monte Carlo method for elliptic PDEs with random coefficients

Participant: Julia Charrier.

This work has been presented at a conference and is submitted in a journal [42], [22].

We consider a finite element approximation of elliptic partial differential equations with random coefficients. Such equations arise, for example, in uncertainty quantification in subsurface flow modelling. Models for random coefficients frequently used in these applications, such as log-normal random fields with exponential covariance, have only very limited spatial regularity, and lead to variational problems that lack uniform coercivity and boundedness with respect to the random parameter. In our analysis we overcome these challenges by a careful treatment of the model problem almost surely in the random parameter, which then enables us to prove uniform bounds on the finite element error in standard Bochner spaces. These new bounds can then be used to perform a rigorous analysis of the multilevel Monte Carlo method for these elliptic problems that lack full regularity and uniform coercivity and boundedness. To conclude, we give some numerical results that confirm the new bounds.

6.6.3. Numerical analysis of the advection-diffusion of a solute in random media

Participant: Julia Charrier.

This work is submitted in a journal [41].

We consider the problem of numerically approximating the solution of the coupling of the flow equation in a random porous medium, with the advection-diffusion equation. More precisely, we present and analyse a numerical method to compute the mean value of the spread of a solute introduced at the initial time, and the mean value of the macro-dispersion, defined at the temporal derivative of the spread. We propose a Monte-Carlo method to deal with the uncertainty, i.e. with the randomness of the permeability field. The flow equation is solved using finite element. The advection-diffusion equation is seen as a Fokker-Planck equation, and its solution is approximated thanks to a probabilistic particular method. The spread is indeed the expected value of a function of the solution of the corresponding stochastic differential equation, and is computed using an Euler scheme for the stochastic differential equation and a Monte-Carlo method. Error estimates on the mean spread and on the mean dispersion are established, under various assumptions, in particular on the permeability random field.

6.6.4. Model reduction for a 1D stochastic elliptic PDE

Participants: Jocelyne Erhel, Mestapha Oumouni.

This work is done in collaboration with Z. Mghazli, from the university of Kenitra, Morocco, in the context of the Co-Advise and Hydromed projects (8.2.1, 8.3.4).

This work has been presented at a conference and published in a journal [15] [34].
In this paper, we present an efficient method to approximate the expectation of the response of a one-dimensional elliptic problem with stochastic inputs. In conventional methods, the computational effort and cost of the approximation of the response can be dramatic. Our method presented here is based on the Karhunen–Loève (K-L) expansion of the inverse of the diffusion parameter, allowing us to build a base of random variables in reduced numbers, from which we construct a projected solution. We show that the expectation of this projected solution is a good approximation, and give an a priori error estimate. A numerical example is presented to show the efficiency of this approach.

6.6.5. Inverse problems in hydrogeology

Participant: Sinda Khalfallah.

This work is done in collaboration with A. ben Abda, from LAMSIN, Tunisia, in the context of the Hydromed and Co-Advise projects (8.2.1, 8.3.4). It is also done in collaboration with B. T. Johansson, from University of Birmingham, GB.

This work has been submitted to a journal [40].

This work is an initial study of a numerical method for identifying multiple leak zones in saturated unsteady flow. Using the conventional saturated groundwater flow equation, the leak identification problem is modelled as a Cauchy problem for the heat equation and the aim is to find the regions on the boundary of the solution domain where the solution vanishes, since leak zones correspond to null pressure values. To reconstruct the solution to the Cauchy problem in a stable way, we modify and employ an iterative regularizing method proposed recently. In this method, one solves mixed well-posed problems (obtained by changing the boundary conditions) for the heat operator as well as for its adjoint, to get a sequence of approximations to the original Cauchy problem. The mixed problems are solved using a Finite element method (FEM), and the numerical results obtained show that the leak zones can be accurately identified also when there is noise in the data.
6. New Results

6.1. Aggregation methods for optical flow computation

Participants: Charles Kervrann, Denis Fortun.

We address the problem of optical flow estimation, that is recovering the dense apparent motion of the pixels in a sequence of images. It is a fundamental computer vision task at the basis of a large variety of applications: object tracking, video compression, motion segmentation, movement detection, 3D reconstruction... Most of state-of-the-art methods rely on a common global variational framework [22], [31]. Computing optical flow amounts to minimizing a global energy. Our experiments demonstrated that the restriction of the minimization to local regions yields significant improvements of the estimation. Motivated by this fact, we developed a novel method to take advantage of this local approach by deriving the global estimate of the flow field from an aggregation of several local estimates. Our work can thus be seen as a general semi-local framework which can be used to improve the performance of any global variational method (Figure 3). We evaluated the performance of our approach on real and synthetic sequences. We investigate adaptations of this methodology to time-lapse fluorescence microscopy and we have recently performed comparisons with usual correlation techniques.

6.2. Lifetime estimation of moving vesicles in FLIM microscopy

Participants: Charles Kervrann, Philippe Roudot.

Fluorescence lifetime imaging microscopy (FLIM) is a widely spread imaging technique for sensing fluorophore environment in a living biological sample (like pH, ions...). Fluorescence lifetime (i.e. the average time a fluorophore stays in excited state before relaxing to its ground state possibly emitting a photon) is particularly useful to detect the Förster resonance energy transfer (FRET) which quantifies spatial proximity between molecules. We have proposed a statistical framework that exploits the intensity model of the frequency-domain FLIM output to jointly estimate trajectories and lifetimes of tracked vesicles. The proposed tracker, inspired from template cross-correlation or gaussian fitting, combines lifetime estimation and robust M-estimation in a efficient and fast way. Estimation of movement and lifetime are decoupled and alternatively performed, while particle/spot detection is performed on the first frame (Figure 4). To improve the results on real image sequences depicting moving vesicles, the background (cytoplasmic auto-fluorescence) model parameters and the scale parameters involved in the M-estimation procedure are estimated in our approach.

Partner: F. Waharte and J. Boulanger (UMR 144 CNRS PICT IBiSA Institut Curie)

6.3. Repetitive and transient event detection in fluorescence video-microscopy

Participants: Charles Kervrann, Pierre Hellier.

Progresses in imaging dynamic behaviours of molecules including fast video microscopy and the application of evanescent wave microscopy have allowed to image intracellular vesicular movements, exocytosis and endocytosis of fluorescent-tagged proteins. For an unbiased quantification of repetitive and transient events, we have proposed an approach which is versatile enough, to be applicable to diverse although complementary modes of microscopy. The proposed detection method described in [21], [11], [16] can be decomposed into three main steps: i) a first pre-processing step is dedicated to the normalization of the image sequence; ii) the second step is the patch-based detection procedure to detect unusual patterns; iii) a third post-processing step allows us to cluster and count detected events in space and time. In a more recent case study, we have used this approach to analyse image sequences depicting M10 cells stably expressing Langerin-YFP and to get deeper insights in the recycling pathway and dynamics of this molecule.
Figure 3. Average Angular Error (AAE) of vector flows obtained with the global variational method and three aggregation methods applied to three image sequences of the Middlebury (http://vision.middlebury.edu/flow/) database: row #1: first frame of the sequence; row #2: global variational method; row #3: mean aggregation; row #4: confidence-based weighted average; row #5: graph cut aggregation.
6.4. Atlas creation of fluorescence microscopy images

**Participant:** Pierre Hellier.

In this work, we consider the analysis of fluorescence images over time to account for two artifacts: fluorescence decreasing over time and geometric misalignment. A single exponential function is most commonly used to represent measured fluorescence decay profiles due to photobleaching. Accordingly, homologous points need to be geometrically aligned over time. Unfortunately the living cell exhibits slow motion over time. We have considered the iterative estimation of both geometrical alignment and intensity correction by the creation of a 3D atlas.

6.5. Averaging of 3D volumes and denoising for the analysis of cryo-electron tomograms

**Participant:** Charles Kervrann.

Trichocysts are large vesicles secreted by the ciliated protozoa, Paramecium. They are characterized by the presence of large three-dimensional crystals of proteins. Under chemical or physical stimuli, or facing a predator, trichocysts undergo an exocytosis, right after fusion of their membrane with the unicellular organism plasma membrane. The crystalline mesh changes its conformation from a condensed to an extended shape in a few milliseconds. Nowadays, cryo-electron tomography (cryo-ET) allows one to visualize those crystals and so, to analyse their three-dimensional organization. However, two main impediments remain with this method. Samples are very sensitive to electron radiation involving the spreading out of the electron dose on the whole tilt series, causing the emergence of background noise in the images. Moreover, a lack of data occurs during image acquisition, called the “missing wedge”, due to uncovered angles at the moment of the acquisition of the tilt series. After tomogram reconstruction of four trichocysts, we have tested usual denoising methods (anisotropic diffusion, Fourier coefficient thresholding) and an unpublished patch-based denoising method
inspired from the nD-safir software (5.1). The denoising methods improved the alignment of different crystal sub-volumes. The sub-volume averaging allowed us to fill in partially the “missing wedge” and then, to obtain a more faithful three-dimensional crystal reconstruction [17], [15].

**Partners:** E. Pollet, A. Guesdon and D. Chrétien (UMR 6026 CNRS University of Rennes 1)

### 6.6. Analysis of spatio-temporal dynamics of cytoplasmic actin under geometrical confinement

**Participant:** Charles Kervrann.

The generation, cell-cycle regulation and maintenance of such cellular functions are often correlated with symmetry breaking and spatiotemporal reorganization of F-actin assembly. In this study, we analyzed the spatiotemporal evolution of actin filaments using Xenopus meiotic extracts artificially confined within a geometry mimicking the cell boundary. It turns out the confinement of the cytoplasm generates symmetry breaking in the spatial organization of actin filaments. Combination of quantitative image analysis and biochemical perturbations show that both spatial localization of F-actin nucleators and actin turnover play a decisive role in generating symmetry breaking. In this project, we proposed to combine an optical flow-based tracker (Kanade-Lucas-Tomasi tracker [40], [36]) to a photobleaching correction method in order to extract quantitative spatiotemporal characteristics of the actin dynamics.

**Partners:** Z. Gueroui (BioPhysics team, UMR 8640 Ecole Normale Superieure, Paris) and M. Pinot (UMR 144 CNRS Institut Curie, Paris)

### 6.7. Analysis of lateral organization of ordered domains at the plasma membrane surface

**Participant:** Charles Kervrann.

Modifications of plasma membrane physical properties are also known to be involved in the perception and response to environmental modifications such as temperature, mechanical and osmotic stress in various organisms. We have analysed a recently designed probe, di-4-ANEPPDHQ, that can change its fluorescent properties depending on whether it is residing in ordered or disordered phases of the tobacco cell plasma membranes. We performed a spatial analysis (covariograms, Markov Random Fields) of small (<200 nm) ordered domains observed in multispectral confocal microscopy. We focused on binary images, assumed to be realizations of a MRF-Ising model, depicting the spatial organization of ordered domains. Maximum pseudo-likelihood methods were investigated to estimate parameters able to describe the spatial properties of ordered domains. We analyzed the modifications of the whole plasma membrane fluidity, and the distribution of ordered domains occurring in the few minutes following addition of the elicitor of defense reaction cryptogein to tobacco cells.

**Partners:** P. Gerbeau-Pissot, F. Simon-Plas (UMR 1088 PME INRA, Dijon) and K. Kiêu (MIA Unit INRA, Jouy-en-Josas)
5. New Results

5.1. Modeling, observation and control: systems modeled by ordinary differential equations

5.1.1. Nonlinear system identification

Participants: Pierre-Alexandre Bliman, Michel Sorine, Qinghua Zhang.

Our current researches on nonlinear system identification are mainly in the framework of the joint Franco-Chinese ANR-NSFC EBONSI project (See Section 6.5), started in March 2011 for three years, in collaboration with the Laboratory of Industrial Process Monitoring and Optimization of Peking University and with Centre de Recherche en Automatique de Nancy (CRAN). Three topics have been studied this year: system identification with a continuous time autoregressive model, system identification with quantized data, and Hammerstein-Wiener system identification.

Though discrete time models are widely used in system identification, some advantages of continuous time models are also of practical importance, in particular, the ability of fully benefiting from fast sampling devices. Our studies on this topic have resulted in a continuous time black-box model structure for nonlinear system identification, together with an efficient model estimation method. This model structure belongs to the class of continuous time nonlinear ARX (AutoRegressive with eXogenous input) models, with the particularity of being integrable. By applying techniques of adaptive observer, models of the proposed structure can be efficiently estimated from input-output data. This result has been presented at the last Journées Identification et Modélisation Expérimentale [49].

System identification is usually based on sampled and quantized data, because of the important role of digital computers. When quantized data are coded with a sufficiently large number of bits, the effect of quantization is often ignored in the design of system identification methods. However, when data are quantized with few bits, sometimes to a single bit leading to binary data, then the effect of quantization must be explicitly taken into account. Data quantization can be modeled as a non differentiable hard nonlinearity, hence the well known gradient-based optimization methods cannot be used for the identification of such nonlinear systems. We have developed a quadratic programming-based method for system identification from quantized data, which, in contrast to most existing methods, can be applied to systems with general input excitations. This result has been presented at the last IFAC World Congress [46].

A Hammerstein-Wiener system is composed of a dynamic linear subsystem preceded and followed by two static nonlinearities. Typically, the nonlinearities of such a system is caused by actuator and sensor distortions. The identification of such systems with a continuous time model had been studied by colleagues of CRAN with the refined instrumental variable (RIV) method. Stable low-pass filters were used to overcome the difficulties related to the continuous time nature of the model. Our study of this year is about the application of the Kalman filter at the place of the previously used low-pass filters. The advantages of this new method include the stability of the numerical algorithm and the fact that the Kalman filter does not color white noises.

5.1.2. Model-based fault diagnosis

Participants: Abdouramane Moussa Ali, Qinghua Zhang.

The increasing requirements for higher performance, efficiency, reliability and safety of modern engineering systems call for continuous research investigations in the field of fault detection and isolation. In the framework of the MODIPRO project funded by Paris Region ASTech, we are currently studying model-based fault diagnosis for nonlinear systems. Motivated by an application in the MODIPRO project, the considered system is modeled by nonlinear algebro-differential equations, with the particularity that the differential part of the model is linear in state variables. Instead of using general numerical solvers of algebro-differential equations, we are developing a method based on ordinary differential equation solvers, by taking advantage of the particular algebro-differential structure of the considered system.
5.2. Observation, control and traveling waves in systems modeled by partial differential equations

5.2.1. Inverse scattering for soft fault diagnosis in electric transmission lines

**Participants:** Michel Sorine, Huaibin Tang, Qinghua Zhang.

The inverse scattering theory is helpful for efficient use of the reflectometry technology in the field of electric transmission line fault diagnosis. Our recent studies on this topic have been published in [35], [33], [32]. The main progress of this year in our study has been the experimental validation of the inverse scattering-based method for soft fault diagnosis. In collaboration with Florent Loete of LGEP (Laboratoire de Génie Electrique de Paris), we have tested the inverse scattering-based method on cables used in Trucks. By slightly separating the two wires of a twisted pair following a predefined spatial profile, a soft fault in the cable is physically simulated. The spatially smoothly varying characteristic impedance of the cable is computed from the physical and geometrical parameters of the cable, and also computed from the reflection coefficient measured with a network analyzer at one end of the cable. The two results are close enough to clearly detect and to locate the physically simulated soft fault from the measured reflection coefficient. A demonstration software has been developed and registered at Agence pour la Protection des Programmes (APP). See Section 4.2.

5.2.2. Modeling of electric transmission networks

**Participants:** Mohamed Oumri, Michel Sorine, Filippo Visco Comandini, Qinghua Zhang.

The increasing number and complexity of wired electric networks in modern engineering systems is amplifying the importance of the reliability of electric connections. In the framework of the ANR 0-DEFECT project, we have studied mathematical models of complex electric networks with the aim of designing an algorithm for fault diagnosis (see [34], [20] for some theoretical results). A generalization of the Baum-Liu-Tesche (BLT) equation to the case of inhomogeneous transmission lines has been developed this year. Efforts have been made in particular to elaborate a fully automatized method for numerical simulation of complex networks with inhomogeneous transmission lines. An efficient method has been designed for the computation of the propagation matrix of each inhomogeneous transmission line and also for the computation of the scattering matrix at each network node. The implemented numerical simulator is based on these propagation and scattering matrices associated to the BLT equation, and on a numerical solution of this equation.

5.2.3. Diagnosis of insulator degradation in long electric cables

**Participants:** Leila Djaziri, Michel Sorine, Huaibin Tang, Qinghua Zhang.

For the diagnosis of insulator degradation in long electric cables, the estimation of the shunt conductance of such cables have been studied, in the framework of the ANR INSCAN project. The shunt conductance of a healthy electric cable is usually very weak. Even when the insulator in the cable is significantly degraded, the shunt conductance can still remain at a quite low level. The main difficulty in this study is due to the fact that the measurements made at the ends of a cable are hardly sensitive to the variations of the shunt conductance. To overcome this difficulty, two methods have been studied. The first one is based on the processing of long time data records. It is designed for the estimation of distributed shunt conductance, in order to detect and to locate inhomogeneous degradation of the insulator. The main idea of this method is to compensate the weak sensitivity of the measurement by long time data records. Numerical simulations have confirmed its feasibility. The second method aims at assessing the average shunt conductance along a cable. It is based on the analysis of the sensitivity of the wave propagation velocity to the shunt conductance. This method is currently tested through experiments made on cables of SNCF (Société Nationale des Chemins de Fer français).

5.3. System theory approach of some quantum systems

5.3.1. Design of strict control-Lyapunov functions for quantum systems with QND measurements

**Participants:** Hadis Amini, Mazyar Mirrahimi, Pierre Rouchon.
We have proposed a feedback scheme for quantum systems undergoing discrete-in-time non-destructive measurements. Under some observability assumptions, the proposed feedback law ensures the stabilization of any desired equilibrium state of the measurement process. This theoretical contribution has given rise to a primary conference paper [40] and we are currently working on a more complete journal paper that proves the convergence in presence of various measurement noises and uncertainties. The proposed feedback scheme has been applied in a recent experiment realized by Serge Haroche and Jean-Michel Raimond’s group at Ecole Normale Supérieure and has given rise to a journal paper [29].

5.3.2. Approximate stabilization of an infinite dimensional quantum stochastic system

Participants: Ram Somaraju, Mazyar Mirrahimi, Pierre Rouchon.

Extending our previous feedback schemes dealing only with finite dimensional quantum systems, we have proven the approximate stabilization of any desired Fock state in the microwave cavity setup of Ecole Normale Supérieure. By appropriately choosing the Lyapunov function, we avoid the mass-loss phenomena through high energy levels and ensure the pre-compactness (in probability) of the trajectories. This work has given rise to a conference paper [48] and a submitted journal paper [56].

5.3.3. On stability of continuous-time quantum filters

Participants: Hadis Amini, Mazyar Mirrahimi, Pierre Rouchon.

We have studied the stability of quantum filters for continuous-in-time measurement processes. Indeed, we have proven that the filter between the quantum state governed by a continuous time stochastic master equation driven by a Wiener process and its associated quantum-filter state is a sub-martingale. This result has given rise to a conference paper [41].

5.4. Modeling, observation and control in biosciences - Reproductive system

5.4.1. Numerical simulation of the selection process of the ovarian follicles

Participants: Benjamin Aymard, Frédérique Clément.

Collaboration with Frédéric Coquel and Marie Postel.

We have designed and implemented a numerical method to simulate a multiscale model describing the selection process in ovarian follicles [9], [8]. The PDE model consists in a quasi-linear hyperbolic system of large size, namely $N_f \times N_f$, ruling the time evolution of the cell density functions of $N_f$ follicles (in practice $N_f$ is of the order of a few to twenty). These equations are weakly coupled through the sum of the first order moments of the density functions. The time-dependent equations make use of two structuring variables, age and maturity, which play the roles of space variables. The problem is naturally set over a compact domain of $\mathbb{R}^2$. The formulation of the time-dependent controlled transport coefficients accounts for available biological knowledge on follicular cell kinetics. We introduce a dedicated numerical scheme that is amenable to parallelization, by taking advantage of the weak coupling. Numerical illustrations assess the relevance of the proposed method both in term of accuracy and HPC achievements [50], [51].

5.4.2. Multiscale modeling of follicular ovulation as a mass and maturity dynamical system

Participants: Frédérique Clément, Philippe Michel, Danielle Monniaux.
We have analyzed the dynamics of the solutions using bifurcation tools on a reduced, ODE model \[ 73 \]. In a first stage, the 2D PDE model is reduced to a 1D PDE model, where the only remaining variable is the age. This reduction is based on a result of exponential convergence in maturity; we have proved that the granulosa cell density of each follicle converges to a “dirac mass in maturity”, which can be understood as: “the follicle becomes uniform in maturity”. The proof is based on the crucial decay property of the maturity speed rate with respect to the maturity variable, so that the support of the cell density of each follicle concentrates its mass around a curve given by a characteristic equation. In a second stage, the mitosis rate is averaged in age, reducing the 1D PDE to a simpler system of two coupled nonlinear ODE, where each follicle is characterized by its cell number (the follicle mass) and global maturity. These variables correspond respectively to the zero-order moment and first-order moment in maturity of the cell density in the original model. The dynamics of one given follicle can then be studied with respect to the pressure exerted collectively by all other growing follicles, in the framework of dynamical games. In some sense, the pressure can be considered as an exogenous parameter, so that we can detect dynamical bifurcations according to the pressure value. Each follicle plays against the others for its survival. In the course of its terminal development, a follicle first remains in the proliferative zone of the mass-maturity plane and then enters the differentiated zone. At the transition from proliferation to differentiation, the follicle is highly sensitive to the pressure. In the worst (doomed) case, the follicle becomes atretic, due to prolonged cell loss. In the best (saved) case, it manages to go through the vulnerability zone and becomes insensitive to the pressure of other follicles.

5.4.3. **Optimal control for a conservation law modeling the development of ovulation**

**Participants:** Frédérique Clément, Peipei Shang.

Collaboration with Jean-Michel Coron

We are now investigating control problems associated to the multiscale model of follicle selection. The conditions for the triggering of the ovulatory surge, coupled with the sorting of the ovulatory follicles, define a complex, nested reachability problem. We have considered a more tractable version of that problem, which is centered on defining the optimal local control corresponding to a single ovulatory trajectory. Under some simplifying assumptions (no loss term and constant aging velocity), we have obtained analytical and numerical results in the case when the density is idealized by one or several Dirac mass. We are extending our results to the PDE original model.

5.4.4. **Multiscale analysis of mixed-mode oscillations in a phantom bursting model**

**Participants:** Frédérique Clément, Mathieu Desroches, Maciej Krupa, Alexandre Vidal.

We have carried on the study of our fast-slow model of the GnRH (gonadotropin-releasing hormone) pulse and surge generator \[ 5 \], \[ 4 \]. If we relax a little the constraints imposed by the biological specifications on the parameters, very rich and complex behaviours can be further exhibited by the model. More precisely, both a delay to the surge and a post-surge pause (before pulsatility resumption) may occur. A detailed examination of the pause has revealed that it is shaped by mixed-mode oscillations (MMO). We are currently investigating how the precise sequence of MMO is determined by the global return map from the surge to the pulse regime.

5.4.5. **Transient synchronization of calcium oscillations in cultures of GnRH neurons.**

**Participants:** Frédérique Clément, Maciej Krupa, Alexandre Vidal.

We have started to study the individual dynamics of GnRH neurons and the conditions under which they may synchronize. We are more specifically tackling the issue of synchronization of calcium oscillations in cultures obtained from the olfactory placodes of rhesus monkey embryos \[ 81 \]. We have introduced a class of models explaining the synchronization events; their main idea was to introduce a global variable controlling the onset of synchronization that was subsequently reset by the subsequent high firing rate caused by the activation of an adaptation current.

5.5. **Clinical and physiological applications**

5.5.1. **DynPeak: An algorithm for pulse detection and frequency analysis in hormonal time series**

**Participants:** Frédérique Clément, Claire Médigue, Alexandre Vidal, Qinghua Zhang.
The endocrine control of the reproductive function is often studied from the analysis of luteinizing hormone (LH) pulsatile secretion by the pituitary gland. Whereas measurements in the cavernous sinus cumulate anatomical and technical difficulties, LH levels can be easily assessed from jugular blood. However, plasma levels result from a convolution process due to clearance effects when LH enters the general circulation. Simultaneous measurements comparing LH levels in the cavernous sinus and jugular blood have revealed clear differences in the pulse shape, the amplitude and the baseline. Besides, experimental sampling occurs at a relatively low frequency (typically every 10 min) with respect to LH highest frequency release (one pulse per hour) and the resulting LH measurements are noised by both experimental and assay errors. As a result, the pattern of plasma LH may be not so clearly pulsatile. Yet, reliable information on the InterPulse Intervals (IPI) is a prerequisite to study precisely the steroid feedback exerted on the pituitary level. Hence, there is a real need for robust IPI detection algorithms. We have designed an algorithm for the monitoring of LH pulse frequency, basing ourselves both on the available endocrinological knowledge on LH pulse (shape and duration with respect to the frequency regime) and synthetic LH data generated by a simple model [57]. We make use of synthetic data to make clear some basic notions underlying our algorithmic choices. We focus on explaining how the process of sampling affects drastically the original pattern of secretion, and especially the amplitude of the detectable pulses. We then describe the algorithm in details and perform it on different sets of both synthetic and experimental LH time series. We further comment on how to diagnose possible outliers from the series of IPIs which is the main output of the algorithm.
5. New Results

5.1. Definition of the research priorities, and first works

As STEEP addresses pluri-disciplinary research topics that are new both for INRIA and for the STEEP members, our very first goal of this year has been to define and specify the four research priorities described above: (a) development of numerical systemic models (economy / society / environment) at local scales; (b) calibration and convergence of integrated models; (c) consideration and management of uncertainties in integrated models; (d) environmental impacts of urban policies.

To gain better insight into these issues, we choose to start with the TRANUS model, as first experimental framework. We collaborate closely with IDDRI (Institut du Développement Durable et des Relations Internationales, Paris http://www.iddri.org/ ), that has implemented a TRANUS model applied to the city of Grenoble in the context of the AETIC project. We make use of this model as starting point to investigate the three first topics.

(a) We have started a project consisting in designing and implementing an integrated model by ourselves, in close collaboration with the EDDEN laboratory (laboratoire d’Économie du Développement Durable et de l’Energie, Grenoble http://webu2.upmf-grenoble.fr/LEPII/spip/spip.php?article22 ) and IDDRI. The goal in to enrich the transport / land use Grenoble model with energy sectors, using the data and analysis our partners are generating in the context of the AETIC project. The goal here is to find a way to encapsulate in or combine the energy models proposed by EDDEN with the TRANUS model implemented by IDDRI for Grenoble.

(b) We participated to the still on-going calibration procedure of the Grenoble models, which is leaded by IDDRI and Modelistica. This helped us to identify the keypoints of the calibration of such a model, and to propose first semi-automatic approaches based on simple parameters optimizations.

(c) We started an uncertainty analysis of this model.

Parallel to that, we have taken an interest in the LEAM model and URBANSIM and we are preparing an ANR project on TRANUS and URBANSIM which would allow us to deeply compare the very different representations on which they are based on.

5.2. Building of Partnerships

This year, we have spent a considerable amount of time and energy to build various partnerships.

We have started a collaboration with Modelistica and Tómas de la Barra, the author of TRANUS Model. Tómas de la Barra came at INRIA for a week in June 2011. We have submitted a project to ECOS-Nord program (TRACER). Among others, we are closely working with him on TRANUS calibration.

Our collaboration with EDDEN is now fully set up. We are working with them on the design of a transport-land use-energy integrated model for the city of Grenoble using the data and analysis they are generating in the context of the AETIC project. This work is based on the TRANUS model implemented by IDDRI for Grenoble. In the same way, our collaboration with IDDRI has started to be effective. With IDDRI, we work on TRANUS calibration issues and practically on the Grenoble model they are implementing. IDDRI is the coordinator of TRACER project submitted to ECOS-Nord program.

In other respects, with AURG (Agence d’Urbanisme de la Région Grenobloise http://www.aurg.org/ ) and EDDEN, we are one of the driving forces behind the MUTERA project. This project aims at gathering the main actors in urban planning and transport of the Rhône-Alpes region to work on the issues regarding land-use and transport models. This group includes technicians, politicians and researchers. The kick-off meeting should take place in the beginning of 2012.
We have also started a collaboration with the LECA laboratory (Laboratoire d’Ecologie Alpine, Grenoble http://www-leca.ujf-grenoble.fr/). In June 2011, we have submitted a project to FRB Flagship program with them (ESNET).

In parallel of that, we got in touch with a number a potential partners and colleagues ones of which are the LET laboratory (Laboratoire d’Economie des Transport, Lyon http://www.let.fr/), the department of urban and regional planning of the university of Illinois at Urbana-Champaign (visit to Brian Deal), Veolia Environnement, SOGREAT etc. A number of them have been contacted in the framework of the development of an ANR project which should be submitted to 2012 “Modèles numériques” program or in the framework of CIFRE or industrial partnerships.

### 5.3. First results on material flow, production and consumption analysis

This theme of research has been started this year through a 6 month internship with a student of Ecole Centrale de Lyon. A large database has been constructed, that allows to analyze material flows, production and consumption by type (and origin and destination in case of flows) at the level of the French region and departments. This database compiles data from a variety of public institutions and private organizations. Almost all the data are in the public domain. A large amount of work has gone into building the database, normalizing the different data sources at least in terms of classification, and into building tools to manipulate the data and produce usable information.

At present, this “physical accounting” has been performed on all French regions for a limited number of materials (in particular cereals, fossil fuels, and construction materials). Furthermore, it has been performed for the same materials at the departement level in the Rhone-Alpes region.

This work is now focusing on various issues:

- By correlating the relevant quantities with other variables (such as, e.g., population or land use, etc), it is possible to downscale the consumption and waste production data at the level of urban regions (where a large part of this information is unavailable). Such an information is critical for the development of decision-help tools at the level of urban areas.

- A process of quantification of the major source of information on material transports is the SitraM database (transportation database maintained by the Ministry of Ecology and Sustainable Development). These data are statistical in nature, but the statistical error is unknown. We are in the process of developing a way to estimate this error.

- The next step will then be to transform this material use in terms of environmental impact. Various types of impacts (global and local) have been identified. The quantification will rely on various known approaches (Life Cycle Analysis, regional Input-Analysis), notably drawing on part of the work performed in this area by the Global Footprint Network (even for impacts that cannot be quantified in terms of land, carbon or water footprint).
SYMBIOSE Project-Team

6. New Results

6.1. Advanced tools for data management

Participants: Olivier Collin [contact], Dominique Lavenier, François Moreews, Olivier Sallou, Anthony Bretaudeau, Jonathan Piat.

- **Annotation and databases**: The AnnotQTL server is a tool designed to gather the functional annotation of genes from several institutional databases for a specific chromosomal region. [14] [Online publication: http://dx.doi.org/10.1093/NAR/GKR361]. SigReannot-mart is a query environment populated with regularly updated annotations for different oligo sets. It stores the results of the SigReannot pipeline that has mainly been used on farm and aquaculture species [17] [Online publication: http://database.oxfordjournals.org/content/2011/bar025]. BioMart Central Portal is a first of its kind, community-driven effort to provide unified access to dozens of biological databases spanning genomics, proteomics, model organisms, cancer data, ontology information and more. [11] [Online publication: http://database.oxfordjournals.org/content/2011/bar041].

- **Bioinformatics Workflow for Intensive Computation** SLICEE proposes to abstract the calls to the cluster scheduler by handling command submission and takes care of exploiting the data parallelism. It enables an easy implementation maintaining and sharing for bioinformatics workflows using intensive computation resources [32]. OBIWEE is a virtual cluster deployment tool associated with SLICEE. It can be deployed either on private cloud or a public cloud architecture. It helps at facing the increasing demand for bioinformatics intensive treatments, in a context of large dissemination of sequencing technologies usages. [29] http://vapor.gforge.inria.fr/. We also developed a library of bioinformatics softwares implemented on manycore structures such as GPU [27].

6.2. Sequences assembly, alignment and comparison

Participants: Dominique Lavenier [contact], Claire Lemaitre, Pierre Peterlongo, Fabrice Legeai, Guillaume Chapuis, Rayan Chikhi, Nicolas Maillet, Delphine Naquin, Raluca Uricaru, Pavlos Antoniou, Thomas Derrien.

- **Hardware accelerator**: Designing FPGA-based accelerators is a difficult and time-consuming task that can be eased by High Level Synthesis Tools. A C-to-hardware methodology has been used to develop an efficient systolic array for the genomic sequence alignment problem. [42], [25] [Online publication 1: http://www.eetimes.com/design/programmable-logic/4217568/How-to-accelerate-genomic-sequence-alignment-4X-using-half-an-FPGA?Ecosystem=programmable-logic] [Online publication 2: http://www.springerlink.com/content/37j00567qm18h146/]

- **De novo assembly of NGS data**: A novel framework has been introduced for de novo assembly of next-generation sequencing data. The new paired string graphs and localized assembly models are implemented in the Monument assembler [24]. [Online publication: http://www.springerlink.com/content/f5g305j5k73x3k14/]

- **International competition of de novo genome assembly**: The Symbiose team (IRISA/CNRS/ENS Cachan Brittany) participated to this competition. [9] [Online publication: http://genome.cshlp.org/content/early/2011/09/16/gr.126599.111.abstract]

- **Indexation of NGS data**: A novel data structure is described for indexing NGS data. The structure is coupled with filtering algorithms that enable memory-efficient and parallel indexing. [23]

- **Breakpoints in genomes**: We analysed the correlation between 3D chromatin interaction data and breakpoint regions resulting from evolutionary rearrangements in the human genome. We found that two loci distant in the human genome but adjacent in the mouse genome are significantly more often observed in close proximity in the human nucleus than expected. [21] [Online publication: http://www.biomedcentral.com/1471-2164/12/303]
• **Repeat detection**: A tool has been presented for detecting long similar fragments that occur two or more times in a set of biological sequences. This is achieved by using a filter as a preprocessing step, and by using the information that the filter has gathered also in the successive inference phase. [26]. [Online publication: http://www.stringology.org/event/2011/p08.html]

• **Targeted assembly of NGS data**: Mapsembler is an iterative targeted assembler which processes large datasets of reads on commodity hardware. Mapsembler checks for the presence of given regions of interest in the reads and reconstructs their neighborhood, either as a plain sequence (consensus) or as a graph (full sequence structure). [39]

• **Transcriptome assembly and annotation**: We established and analyzed two catalogues of transcripts by assembling EST sequences, and performed their functional annotations using the gene ontology for the following 2 species: *spodoptera littoralis* [15] and *cabomba* [20].

• **Substitution matrices**: A general and simple methodology has been proposed to build new matrices fitted to specific compositional bias of proteins. It was then applied to the large scale comparison of Mollicute AT-rich genomes [16]. [Online publication: http://www.biomedcentral.com/1471-2105/12/457/]

6.3. Genome Structure

**Participants**: Jacques Nicolas [contact], Dominique Lavenier, François Coste, Catherine Belleannée, Olivier Sallou, Fabrice Legeai, Guillaume Rizk, Guillaume Chapuis, Matthias Gallé, Anthony Bretaudeau.

• **GPU accelerated RNA folding algorithm**: The main kernel of the widely used RNA folding package Unafold has been accelerated using GPU boards by reordering computations to enable tiled computations and good data reuse [37], [2].

• **GPU accelerated QTL algorithm**: Our GPU/multicore implementation performs up to 20 times faster than the previous multicore implementation and allows extensive QTL analysis to be conducted in a reasonable time, while maintaining the same level of precision [35].

• **Hierarchical structure of genomes**: In [7], we proposed to split the classical smallest grammar problem into two tasks: (1) choosing the constituents of the grammar and (2) finding the smallest grammar parsing given these constituents. This defines properly the search space for this problem and, as we have shown how to solve in polynomial time the second task, this opens doors for new algorithms finding smaller grammars as shown on a generic compression benchmark (up to 10%). In [6], we have worked on the scalability to propose a new algorithm able to handle whole genomes: on this kind of sequences, the size reduction is still about 10% for a comparable execution time with respect to state-of-the-art algorithms.

• **Data compression**: By using grammars with rigid patterns as words, we were able to achieve a compression rate up to 25% better compared to the previous best DNA grammar-based coder, and just below state-of-the-art dedicated DNA compressors [1].

• **CRISPR Modeling and identification**: CRISPR (Clustered regularly interspaced short palindromic repeats) are small repeats present in a number of bacterial and archaeal species. We proposed the most complete database on these elements (http://crispi.genouest.org), elaborating for the first time a complete study of the palindromic nature of these repeats. The analysis has made an extensive use of our Logol Parser to decipher stem-loop structures [40].

• **Aphid genetics**: We participated in a genetic study aiming at comparing the rates of evolution of genes enclosed in aphid sexual chromosome (X) to autosomal genes. In order to do so, we provided particular microsatellites for the selection of genomic sequences, as well as tools for studying their genomic environment. [13] [Online publication: http://mbe.oxfordjournals.org/content/early/2011/10/12/molbev.msr252]
6.4. Protein Sequences and Structures

Participants: Rumen Andonov [contact], Antonio Mucherino, François Coste, Jacques Nicolas, Andres Burgos, Gaëlle Garet, Pavel Senin, Mathilde Le Boudic-Jamin.

- **Branch & Prune Algorithm**: We proposed an extension of the Branch & Prune (BP) algorithm for the Discretizable Molecular Distance Geometry Problem (DMDGP) which is able to exploit all symmetries of the research domain of the corresponding combinatorial optimization problem [30].

- **Modeling protein sequences with long distance correlations**: To initiate this new line of research, we have set up a framework to learn context-free grammars on protein sequences based on the identification of conservation blocks and substitutability of non-terminals. A first implementation of the learning algorithms showed the interest of this approach [38].

- **Maximum Contact Map Overlap (CMO)** is a popular measure for quantifying the similarity between protein structures. A new integer programming model was presented for CMO and an exact branch-and-bound algorithm was designed with bounds obtained by a novel Lagrangian relaxation. The efficiency of the approach was demonstrated on known benchmarks on which sets our approach significantly outperforms the best existing exact algorithms [3].

- **Alignment of protein structures**: First successes were obtained on provably optimal pairwise alignment of protein inter-residue distance matrices, using the popular Dali scoring function. We proposed the first mathematical model for computing optimal structural alignments based on dense inter-residue distance matrices and present algorithm engineering techniques to handle the huge integer linear programs [22]. In a second paper, a strategy was proposed for sparsifying distance matrices in which only the distances needed for uniquely reconstructing the conformations of the proteins are kept. [31].

- **Protein Family Identification**: Identification of protein families is a computational biology challenge that needs efficient and reliable methods. First, we used the comparison tool A_purva, which is based on Contact Map Overlap (CMO), to classify protein structure coming from the CATH database. The obtained results showed that A_purva was able to correctly classify 92% of the structures, and that introducing the notion of dominance drastically reduced the computational time needed for classifying the protein structures [33]. Then, we introduced this concept of dominance in a novel combined approach based on Distance Alignment Search Tool (DAST), which contains an exact algorithm with bounds. Our experiments showed that this method successfully finds the most similar proteins in a set without solving all instances [28].

- **Local Protein Threading, sequence-structure alignment**: A novel approach to PTP has been investigated. It aligns a part of a protein structure onto a protein sequence in order to detect local similarities [8]. [Online publication: http://www.sciencedirect.com/science/article/B6TYW-50G78H4-1/2/947312da7a7bbf175cab7b3288ba4f03]

6.5. Systems Biology

Participants: Anne Siegel [contact], Jérémie Bourdon, Michel Le Borgne, Nathalie Theret, Geoffroy Andrieux, Oumarou Abdou-Arbi, Sylvain Prigent, Pierre Blavy, Andres Aravena, Santiago Videla, Valentin Wucher, Brivael Trelhu.

- **Average-case analysis for quantitative data integration**: We proposed a probabilistic modeling framework that integrates heterogeneous data. Average case analysis methods were used in combination with Markov chains to link qualitative information about transcriptional regulations to quantitative information about protein concentrations. The approach was illustrated by modeling the carbon starvation response in Escherichia coli. It accurately predicted the quantitative time-series evolution of several protein concentrations using only knowledge of discrete gene interactions and a small number of quantitative observations on a single protein concentration [5]. [Online publication: http://dx.plos.org/10.1371/journal.pcbi.1002157]
- **Combining genetic and metabolic regulations**: We mixed Gale-Nikaido reduction steps and differential inequalities to understand how genetic regulation modifies the behavior of a very abstracted model of lipid metabolism [18]. [Online publication: http://www.springerlink.com/content/n437048670560782/]

- **Extract relevant information with respect to a cancer phenotype**: We designed dedicated logical rules to model the static response of biomolecular interactions implied in the cancer network. This allowed us to trace back genes implied in the cancer phenotype [12]. [Online publication: http://www.computer.org/portal/web/csdl/doi/10.1109/TCBB.2010.71]

- **Integrative biology for brown algal**: We proposed a protocol focusing on integrating heterogeneous knowledge gained on brown algal metabolism. The resulting abstraction of the system helps understanding how brown algae cope with changes in abiotic parameters within their unique habitat [19].

- **Search for key regulators**: A method was proposed to model the effects of all transcriptional and metabolic regulations contained in transpath in a single influence network. The network was analyzed to find a set of candidates that explain the variations of a set of targets [34].

- **Identification of co-regulation patterns**: We introduced a new approach based on the compilation of Simple Shared Motifs (SSM), groups of sequences defined by their length and similarity and present in conserved sequences of gene promoters. We proved that Simple Shared Motifs analysis provides a clearer definition of expression networks [10]. [Online publication: http://www.biomedcentral.com/1471-2105/12/365]

- **Probabilistic models for systems biology**: We reviewed, in a book chapter, some classical concepts concerning probabilistic models and their applications in systems biology. Probabilistic boolean networks were presented in deep with a focus on the effect of synchronization of genes and on stochastic simulation of such networks [36].
5. New Results

5.1. Analysis of structures resulting from meristem activity

5.1.1. Acquisition and design of plant geometry

Participants: Chakkrit Preuksakarn, Frédéric Boudon, Christophe Pradal, Christophe Godin.

This research theme is supported by RTRA and ARC projects named PlantScan3D.

Virtual 3D model of plants are required in many areas of plant modeling. They can be used for instance to simulate physical interaction of real plant structures with their environment (light, rain, wind, pests, ...), to set up initial conditions of growth models or to assess their output against real data. In the past decade, methods have been developed to digitize plant architectures in 3D [42], [33]. These methods are based on direct measurements of position and shape of every plant organ in space. Although they provide accurate results, they are particularly time consuming. More rapid and automated methods are now required in order to collect plant architecture data of various types and sizes in a systematic way. In this aim, we explore the use of pictures, video, laser scanner and direct sketching.

- **Reconstruction of plant architecture from 3D laser scanner data.** (Chakkrit Preuksakarn, Mathilde Balduzzi, Frédéric Boudon, Jean-Baptiste Durand, Christophe Godin, Xinghua Song [INRIA, Galaad], Bernard Mourrain [INRIA, Galaad], Dobrina Boltcheva [INRIA, Imagine], Franck Hetry [INRIA, Morpheus], Marie-Paule Cani [Inria, Imagine], Pascal Ferraro [Labri, Bordeaux])

  We investigate the possibility to use 3D laser scanners to automate plant digitizing. We are developing algorithms to reconstruct branching systems without leaves or foliage from scanner data or from scan simulated on plant mock-up obtained using different digitizing method. For this we collaborate with the EPI Galaad from Sophia-Antipolis, the EPI Imagine from Grenoble, different INRA teams, UMR PIAF in Clermont Ferrand, UMR LEPSE and AFEF team in Montpellier and Lusignan, the University of Helsinki, Finland and the CFCC in England.

  We developed a reconstruction pipeline composed of several procedures. A contraction procedure, first aggregates points at the center of the point cloud. The team proposed a simple adaptive scheme to contract points. In parallel, the Galaad team explored uses of detection of circular patterns to be contracted toward their center. Comparison of these approaches has been carried out. In a second step, a skeleton procedure uses a Space Colonization Algorithm [41] to build the skeleton of the shape from the contracted point set. This method is adaptive to the local density of the point set. Then a pipe-model based procedure makes it possible to estimate locally diameters of the branches. Finally, an evaluation procedure has been designed to assess the accuracy of the reconstruction. Results Publication of this work is in progress.

- **Sketching of plants.** (Frédéric Boudon, Christophe Godin, Steven Longuay [University of Calgary, Canada], Przemyslaw Prusinkiewicz [University of Calgary, Canada])

  Modeling natural elements such as trees in a plausible way, while offering simple and rapid user control, is a challenge. In a first collaboration with the EPI Evasion we developed a method based on the design of plants from silhouettes [43]. This sketching paradigm allows quick and intuitive specification of foliage at multiple scales. On this topic, we started a collaboration with S. Longuay and P. Prusinkiewicz who develop iPad tools to design plants based on SCA. Combination of multitouch interface, sketching paradigm and powerful adaptive procedural model that generate realistic trees offer intuitive and flexible design tools. This work is part of the INRIA associated team with the University of Calgary.

- **Reconstruction of vineyards from video.** (Frédéric Boudon, Jerome Guenard [IRIT, Toulouse], Geraldine Morin [IRIT, Toulouse], Pierre Gurdjos [IRIT, Toulouse], Vincent Charvillat [IRIT, Toulouse])
In this work, we investigate the reconstruction of constrained plant geometry of a vineyard from a set of pictures coming from video. Pictures are segmented to identify the different trees of a same row in the yard. From this segmentation, a number of parameters are estimated, which makes it possible to instantiate a virtual model of a vine tree. In particular, paths of main branches and leaf volumes and densities are estimated. A preliminary version of this work has been presented to the AFIG conference [21].

- **Reconstruction of virtual fruits from pictures.** (Mik Cieslak, Nadia Bertin [Inra, Avignon], Frédéric Boudon, Christophe Godin, Christophe Pradal, Michel Genard [Inra, Avignon], Christophe Goz-Bac [Université Montpellier 2])

  This research theme is supported by the Agropolis project Fruit3D.

Understanding the controlling factors of fruit quality development is challenging, because fruit quality is the result of the interplay between physical and physiological processes that are under the control of genes and the environment. Although process-based models have been used to make significant progress in understanding these factors, they to a large extent ignore the shape and internal structure of the fruit. To help characterizing effects of fruit shape and internal structure on quality, the creation of a 3D virtual fruit model that integrates fruit structure and function with growth governed by environmental inputs has been investigated. For this purpose, a modelling pipeline has been created that includes the following steps: creation of a 3D volumetric mesh of the internal and external fruit structure, calculation of the fruit’s physical properties from the resulting mesh, and integration of aspects of fruit physiology into the 3D structure. This pipeline has been applied to study tomato fruit (Solanum lycopersicum) by constructing 3D volumetric meshes from two images of perpendicular fruit slices and from MRI data, and integrating water and carbon transport processes into one of these meshes. To illustrate the tomato model, a simulation of one season’s of the fruit’s growth has been performed and its results compared with an already published process-based tomato fruit model. The results of the two models were in general agreement, but our model provided additional information on the internal properties of the fruit, such as a gradient in sugar concentration. Once the model is calibrated and evaluated, our approach will be suitable for studying the effects of internal fruit heterogeneity and overall shape on fruit quality development [18].

5.1.2. **Modeling the plant ontogenic programme**

**Participants:** Christophe Godin, Yann Guédon, Evelyne Costes, Jean-Baptiste Durand, Pascal Ferraro, Anaëlle ambreville, Christophe Pradal, Catherine Trottier, Jean Peyhardi, Yassin Refahi, Etienne Farcot.
This research theme is supported by two PhD programmes.

The remarkable organization of plants at macroscopic scales may be used to infer particular aspects of meristem functioning. The fact that plants are made up of the repetition of many similar components at different scales, and the presence of morphological gradients, e.g. [23], [35], [36], [32], provides macroscopic evidence for the existence of regularities and identities in processes that drive meristem activity at microscopic scales. Different concepts have been proposed to explain these specific organizations such as “morphogenetic programme” [38], “age state” [31] or “physiological age” [24]. All these concepts state that meristem fate changes according to position within the plant structure and during its development. Even though these changes in meristem fate are specific to each species and lead to the differentiation of axes, general rules can be highlighted [31], [24]. Here we develop computational methods to decipher these rules.

- **Branching and axillary flowering structures of fruit tree shoots.** (Yann Guédon, Evelyne Costes, David Da Silva [UC Davis], Anna Davidson [UC Davis], Ted DeJong [UC Davis], Claudia Negron [UC Davis]).

  In the context of a collaboration with Claudia Negron, Anna Davidson, David Da Silva and Ted DeJong, stochastic models (hidden semi-Markov chains) for the branching and axillary flowering structures of different categories of peach and almond shoots corresponding to different genetic backgrounds, environment conditions and horticultural practices were built. These stochastic models have been integrated in simulation systems which combine stochastic models with different mechanistic models of biological function, in particular carbon partitioning models. This collaboration extends the work initiated on apple trees [40], [3]; see 5.1.3.

- **Genetic determinisms of the alternation of flowering in apple tree progenies.** (Jean-Baptiste Durand, Jean Peyhardi, Baptiste Guitton [DAP, APEF team], Catherine Trottier, Evelyne Costes, Yann Guédon)

  Previous approaches for a statistical quantification of the effect of factors on tree architecture were mainly oriented toward the structure of main axes, and environmental explanatory variables (see [34], [27]). To characterize genetic determinisms of the alternation of flowering in apple tree progenies at annual shoot (AS) scale, a model of the transitions between ASs was built. The ASs were of two types: flowering or vegetative. Two replications of each genotype were available. Our model operated on tree-structured data and relied on a second-order Markov tree. Generalized Linear Mixed Models (GLMMs) were used to model the effect of year, replications and genotypes (with their interactions with year or memories of the Markov model) on the transition probabilities. This work was the continuation of the Master 2 internship of Jean Peyhardi (Bordeaux 2 University) and was carried out in the context of the PhD thesis of Baptiste Guitton.

  This PhD thesis also comprised the study of alternation in flowering at individual scale, with annual time step. To relate alternation of flowering at AS and individual scales, indices were proposed to characterize alternation at individual scale. The difficulty is related to early detection of alternating genotypes, in a context where alternation is often concealed by a substantial increase of the number of flowers over consecutive years. To separate correctly base effect from alternation in flowering, our model relied on a parametric hypothesis on the base effect (random slopes specific to genotype and replications), which translated into mixed effect modeling. Different indices of alternation were then computed on the residuals. Clusters of individuals with contrasted patterns of bearing habits were identified. Our models highlighted significant correlations between indices of alternation at AS and individual scales. The roles of local alternation and asynchronism in regularity of flowering were assessed using an entropy-based criterion, which characterized asynchronism.

  As a perspective of this work, patterns in the production of children ASs (numbers of flowering and vegetative children) depending on the type of the parent AS must be analyzed using branching processes and different types of Markov trees, in the context of Pierre Fernique’s PhD Thesis.

- **Modeling branching patterns in fruit tree shoots through the characterization of their demographic properties** (Pierre Fernique, Jean-Baptiste Durand, Yann Guédon).
To test the effect of some properties of a given parent shoot on the properties of its children shoots, statistical models based on multitype branching processes were developed. This kind of dependence between parent and children shoots is frequently at stake in fruit trees, for which the number of flowering or vegetative children of a parent shoot depends on its nature, with potential interactions with other factors. Thus, controlling demographic patterns of the shoots (through varietal selection or crop management strategies) is expected to bring substantial improvements in the quantity and quality of yields.

Formally, the shoot properties are summed up using the notion of shoot state. The number of children shoots in each state is modeled through discrete multivariate distributions. Model selection procedures are necessary to specify parsimonious distributions. We developed an approach based on probabilistic graphical models to identify and exploit properties of conditional independence between numbers of children in different states, so as to simplify the specification of their joint distribution. The graph building stage was based on a Poissonian Generalized Linear Model for the contingency tables of the counts of joint children state configurations. Then, parametric families of distributions were implemented and compared statistically to provide probabilistic models compatible with the estimated independence graph.

This work was carried out in the context of Pierre Fernique’s Master 2 internship (Montpellier 2 University and AgroParisTech). It was applied to model dependencies between short or long, vegetative or flowering shoots in apple trees. The results highlighted contrasted patterns related to the parent shoot state, with interpretation in terms of alternation of flowering. This work will be continued during Pierre Fernique’s PhD thesis, with extensions to other fruit tree species (mango trees) and other strategies to build probabilistic graphical models and parametric discrete multivariate distributions including covariates and mixed effects.

Analyzing fruit tree phenology (Anaëlle Dambreville, Jean-Baptiste Durand, Yann Guédon, Christophe Pradal, Pierre-Eric Lauri [UMR AGAP], Frédéric Normand [UPR HortSys], Catherine Trottier)

Mango is a tropical tree characterized by strong asynchronisms within and between trees. Causation networks explaining the vegetative and reproductive growths within and between growing cycles were studied on the basis of generalized linear models. We highlighted in this way marked interplays between structural and temporal components of tree structure development at three scales. At growth unit scale, a growth unit appeared early in the growing cycle had higher rate of burst compared to late appeared growth units. At growing cycle scale, a growth unit which flowered delayed its future vegetative growth compared to a vegetative growth unit. At tree scale, a fruiting tree delayed further vegetative growth and flowering compared to a non-fruiting tree. These results evidenced that tree phenology is strongly affected by structural components and not only by the environment.

Self-nested structure of plants. (Christophe Godin, Pascal Ferraro)

To study the redundancy of structures embedded at various levels in tree architectures, we investigated the problem of approximating trees by trees with particular self-nested structures. Self-nested trees are such that all their subtrees of a given height are isomorphic. We show that these trees present remarkable compression properties, with high compression rates. In order to measure how far a tree is from being a self-nested tree, we introduced a quantitative measure of the degree of self-nestedness for any tree. For this, we want to find a self-nested tree that minimizes the distance of the original tree to the set of self-nested trees that embed the initial tree:

$$NEST(T) = \arg \min_{S \in S^+(T)} D(T, S),$$

where $T$ is a tree, $D(\cdot, \cdot)$ is a distance on the set of trees (chosen so as to preserve certain structural properties between the compared trees) and $S^+(T)$ is the set of self-nested trees that contain $T$, i.e. that can be obtained from $T$ by inserting nodes only. In a previous work we showed that this problem can be solved in polynomial time and gave the corresponding algorithm [6].
Now, we continue this work along different directions:

- **Approximate compression including geometry** (Anne-Laure Gaillard, Pascal Ferraro, Frédéric Boudon, Christophe Godin)

  We now investigate how to include branch geometry in the compression schemes. This problem constitutes a part of the PhD thesis of Anne-Laure Gaillard co-supervised with P. Ferraro. First results show that any tree architecture can be compressed at different degrees with a varying loss in the geometric information. A publication on this topic is currently in progress.

- **Search for the NST** (Farah Ben-Naoum, Christophe Godin, Pascal Ferraro)

  The NEST algorithm constructs a closest self-nested tree that embeds a given tree $T$. This means in particular that $T \in NEST(T)$. For some trees this definition might be too restrictive and one may want to get rid of the latter constraint. In this case, the closest self-nested tree (NST) to a tree $T$ would be defined by:

  $$NST(T) = \arg\min_{S \in S(T)} D(T, S),$$

  where $S(T)$ is now the set of all self-nested trees. Although it has not yet been demonstrated, the computation of $NST(T)$ is likely to be a NP-complete problem. Therefore we decided to develop a heuristic algorithm to carry out this optimization. This algorithm is based on evolutionary optimization algorithms. First results obtained at the University of Sidi Bel Abbes by Farah Ben Naoum are very encouraging. They should contribute to define a general method to compress tree-like structures based on self-similarity.

- **Stochastic NESTs** (Jean-Baptiste Durand, Christophe Godin, Pascal Ferraro, Yann Guédon)

  One of the limits of the original NEST algorithm in defining a compressed structure is that it does not take into account possible small variations in the identification of similar tree patterns in the global tree structure. In real trees, it is however seldom the case that similar tree parts are exactly isomorphic. They may differ by small details while keeping globally similar. A general compression technique should account for such variations from a theoretical point of view. This is what we intend to do here by developing a stochastic NEST approach. The idea is to extend the definition of tree compression to include the notion of tree distribution in compressed representations of trees. A first formalism has been designed for stochastic tree compression.

  - Statistical characterization of the branching properties of trees at individual scale – application to the quantification of approximate self-nestedness. (Jean-Baptiste Durand, Yann Guédon, Christophe Godin)

    To test different hypotheses related to the role of apical control on local branching properties within tree architecture, statistical models of the fates of the apical meristem and its production at a given scale were specified. They extend the hidden Markov tree models [29] to variable number of children of a given entity, and provide a model for their dependencies. Moreover, they allow for the comparison of different hypotheses regarding the relevance of the children ordering (absence of ordering, partial or total ordering). In the next years, these models will be used on mango and apple tree cultivars. Particularly, they will found new approaches to compare the effect of various strategies of culture.

    These models are based on a notion of state that extends the concept of equivalence classes for tree isomorphism used in the above paragraph, to the notion of classes of approximate isomorphism (i.e. isomorphism as a stochastic process). As a consequence, the hidden Markov tree models offer new insight for lossy compression of trees, which will be investigated in future work.

    - Analyzing perturbations in Arabidopsis thaliana phyllotaxis. (Yann Guédon, Yassin Refahi, Etienne Farcot, Christophe Godin, Fabrice Besnard [RDP, Lyon], Teva Vernoux [RDP, Lyon])
The cytokinin hormones are known to play a significant role in the regulation of phyllotaxis. To investigate this, Fabrice Besnard and Teva Vernoux are studying Arabidopsis thaliana ahp6 mutants, AHP6 being a protein known for its inhibitory effect in the cytokinin signaling pathway. At the macroscopic scale, this mutation induces perturbations of the phyllotaxis, barely sensible on single plants. In order to characterize these perturbations, we designed a pipeline of models and methods which decompose into three steps: (i) identification of perturbation patterns, (ii) characterization of perturbation patterns using hidden variable-order Markov chains and combinatorial mixture models both with von Mises observation distributions (Gaussian-like periodic distribution for circular variables), (iii) classification of plant phyllotaxis among wild-type and ahp6 mutant phyllotaxies. Using this pipeline of methods, we have shown that the perturbation patterns in both wild-type and mutant plants can be explained by permutations in the order of insertion along the stem of 2 or 3 consecutive organs. The number of successive synchronized organs between two permutations reveals unexpected patterns that depend on the nature of the preceding permutation (2- or 3-permutation). We identified significant individual deviations of the level of baseline segments with reference to 137.5°, which confirms theoretical model predictions. Finally, we highlighted a marked relationship between permutation of organs and defects in the elongation of the internodes in between these organs. All these results can be explained by the absence of a strict coupling between the timing of organ development and their angular and longitudinal position on the stem. A paper about these results is in revision for the journal Science. Another more methodological paper is in progress.

5.1.3. Analyzing the influence of the environment on the plant ontogenic programme

Participants: Jean-Baptiste Durand, Damien Fumey, Frédéric Boudon, Christophe Godin, Yann Guédon, Jean Peyhardi, Pierre Fernique, Christian Cilas, Evelyne Costes, Pascal Ferraro, Catherine Trottier.

This research theme is supported by a CIFRE contract and two PhD programmes.

The ontogenetic programme of a plant is actually sensitive to environmental changes. If, in particular cases, we can make the assumption that the environment is a fixed control variable (see section 5.1.2), in general the structure produced by meristem results from a tight interaction between the plant and its environment, throughout its lifetime. Based on observations, we thus aim to trace back to the different components of the growth (ontogenetic development and its modulation by the environment). This is made using two types of approaches. On the one hand, we develop a statistical approach in which stochastic models are augmented with additional time-varying explanatory variables that represent the environment variations. The design of estimation procedures for these models make it possible to separate the plant ontogenetic programme from its modulation by the environment. On the other hand, we build reactive models that make it possible to simulate in a mechanistic way the interaction between the plant development and its environment.

- Analyzing growth components in trees. (Yann Guédon, Jean-Baptiste Durand, Jean Peyhardi, Yves Caraglio [AMAP], Emilie Lebarbier [AgroParisTech], Catherine Trottier, Olivier Taugourdeau [AMAP])

Observed growth, as given for instance by the length of successive annual shoots along a forest tree trunk, is assumed to be mainly the result of three components: (i) an endogenous component assumed to be structured as a succession of roughly stationary phases separated by marked change points that are asynchronous between individuals [34], (ii) a time-varying environmental component assumed to take the form of fluctuations that are synchronous between individuals, (iii) an individual component corresponding to the local environment of each tree. This environmental component is thus assumed to be a “population” component as opposed to the individual component. In order to identify and characterize these three components, we proposed to use Markov and semi-Markov switching linear mixed models [27][2]. The underlying Markov or semi-Markov chain represents the succession of growth phases (endogenous component) while the linear mixed model attached to each state of the underlying Markov or semi-Markov chain represents -in the corresponding growth phase- both the influence of time-varying climatic explanatory variables (environmental component) as fixed effects, and inter-individual heterogeneity (individual component) as random effects. We
investigated the estimation of Markov and semi-Markov switching linear mixed models in a general framework using MCEM-like algorithms. These integrative statistical models were in particular applied in a forest ecology context to characterize the opportunistic development of understory samplings in relation to light environment [15]. Concerning the application to forest trees, the proposed statistical modeling approach relies on the availability of climatic data. In the case where climatic data are not available, we are studying Markov and semi-Markov switching linear mixed models with year random effects common to all the trees to model the synchronous part of the growth fluctuations. With Markov and semi-Markov switching linear mixed models, the response variable is constrained to be approximately normally distributed. We are now studying the statistical methodology for Markov switching linear mixed models to take into account non-normally distributed response variables (e.g., number of growth units, apex death/life, non-flowering/flowering character). It should be noted that the estimation algorithms proposed for Markov switching linear mixed models can be directly transposed to other families of hidden Markov models such as, for instance, hidden Markov tree models; see Section 5.1.2.

- **Coupling stochastic models with mechanistic models for plant development simulation.** (Damien Fumey, Yann Guédon, Christophe Godin, Thomas Cokelaer, Evelyne Costes, Pierre-Eric Lauri [UMR AGAP])

Arboricultural practices such as pruning, artificial bending or fruit thinning are crucial interventions in orchard management and are used for controlling tree size, penetration of light into the canopy and the equilibrium between vegetative and reproductive growth. In the PhD of Damien Fumey we explored the possibility of integrating such practices in a model of apple tree development. To this end, a field experiment was designed to study the effects of pruning (thinning or heading cuts) on two apple cultivars with contrasted architecture, ‘Fuji’ and ‘Braeburn’. Results of this experiment [30] showed that thinning cuts of laterals tended to be compensated by an increase in lateral branching. Based on these field experiments, a model is currently being developed to account for pruning practices on fruit trees. This model relies on a formalization of the competition of meristems by combining a carbon allocation strategy and a competition mechanism in a stochastic manner. The resulting model is reactive to human interventions and should enable us to capture plant reactions to pruning practices in a robust way [28]. A paper describing this model is currently in progress.

### 5.2. Meristem functioning and development

In axis 2 work focuses on the creation of a virtual meristem, at cell resolution, able to integrate the recent results in developmental biology and to simulate the feedback loops between physiology and growth. The approach is subdivided into several sub-areas of research.

#### 5.2.1. Data acquisition and design of meristem models

**Participants**: Frédéric Boudon, Christophe Godin, Vincent Mirabet, Jan Traas, Grégoire Malandain, Jean-Luc Verdeil.

*This research theme is supported by the ATP CIRAD Meristem and the ANR GeneShape and FlowerModel projects.*

Studies on plant development require the detailed observation of the tissue structure with cellular resolution. In this context it is important to develop methods that enable us to observe the inner parts of the organs, in order to analyze and simulate their behavior. Here we focus on the apical meristems, that have been extensively studied using live imaging techniques and confocal microscopy. An important limitation of the confocal microscope lies in the data anisotropy. To overcome this limitation, we designed new protocols to achieve an accurate segmentation of the cells. Using these segmentations, a geometrical and topological representation of the meristem is built. Such representations may be used to analyze the meristem structure at cell level, to support the description of gene expression patterns and to initiate and assess virtual meristem simulations.

- **Microscopy image reconstruction and automatic lineage tracking of the growing meristem cells**
Participants: Romain Fernandez, Christophe Godin, Grégoire Malandain, Jan Traas, Pradeep Das, Vincent Mirabet.

In previous work [5], we studied the tracking of meristem cells using time-lapse confocal microscopy acquisition on early stages flowers of Arabidopsis shoot apical meristems. We designed a reconstruction method (MARS, Figure 4) and a tracking algorithm (ALT) in order to map the segmentations of the same meristem at different times, based on a network flow representation in order to solve the cell assignment problem. We validated the MARS-ALT pipeline on a four-steps time course of an early stage floral bud. In 2011, we worked to improve the robustness of the MARS-ALT pipeline. The software pipeline was completely re-engineered so as to be easily available in OpenAlea with documentation. We also designed new tests to improve the method and get high-quality results on different types of organisms.

Figure 4. Surface view of a flower meristem automatically segmented using MARS at cell resolution and a transversal cut showing the inner segmented tissues

- Design of 3D virtual maps for specifying gene expression patterns (Jérôme Chopard, Christophe Godin, Jan Traas, Françoise Monéger [ENS Lyon])

This research theme is supported the ANR GeneShape and FlowerModel projects.

To organize the various genetic, physiological, physical, temporal and positional informations, we build a spatialized and dynamic database. This database makes it possible to store all the collected information on a virtual 3D structure representing a typical organ. Each piece of information has to be located spatially and temporally in the database. Tools to visually retrieve and manipulate the information, quantitatively through space and time are being developed. For this, the 3D structure of a typical organ has been created at the different stages of development of the flower bud. This virtual structure contains spatial and temporal information on mean cell numbers, cell size, cell lineages, possible cell polarization (transporters, microtubules), and gene expression patterns. Such 3D virtual map is mainly descriptive. However, like for classical databases, specific tools make it possible to explore the virtual map according to main index keys, in particular spatial and temporal keys. Both a dedicated language and a 3D user interface are being designed to investigate and query the 3D virtual map.
A prototype version of this 3D virtual map is currently being built and is integrated in V-Plants, see figure 5 where a cell-based volumic tissue that can contain different types of information (cell lineage, cell size, cell identity, etc...).

![Cell-based map diagram](image)

Figure 5. A 3D database represented as a cell-based map that can contain different types of information (cell lineage, cell size, cell identity, ...), from [14].

Using this 3D virtual map prototype, along with piecewise-linear models of gene networks, and optimization techniques, our colleagues from ENS/Lyon have been able to develop a model of the gene network regulating sepal polarization in *A. thaliana*. This model, which is consistent with a very comprehensive literature data set, also include new predictions for certain gene interactions, and will appear next year in Plant Cell.[14]

5.2.2. Shape analysis of meristems

(Jean-Baptiste Durand, Frederic Boudon, Yann Guedon, Francois Mankessi [BURST, DAP], Olivier Monteuius [BURST, DAP], Jean Luc Verdeil [PHIV, DAP])

Plants that grow several forms or type of leaves along a shoot, depending on age or shoot length, are called heteroblastic. The influence of heteroblasty on morphological and histocytological characteristics of Acacia mangium shoot apical meristems (SAMs) was assessed comparing materials with mature and juvenile leaf morphology in natural and in vitro conditions. For this we introduced a workflow for characterizing dome shape with few parameters (SAM dome height (H), basal diameter (D) and shape factor (S)) and their joint
statistical analysis to assess influence of conditions on SAM shape. In particular, a new statistical test is introduced here for multivariate analysis. This is a generalization of univariate ANOVA that takes into account statistical dependencies between the shape parameters. As a result, we found that SAM dome height (H) and basal diameter (D) were highly correlated. The joint analysis revealed that H, D, and shape (S) varied significantly according to the four plant origins investigated, with the higher scores for the outdoor mature source “Mat”. Overall, heteroblasty induced more conspicuous differences of SAM characteristics for the outdoor than for the in vitro materials.

5.2.3. Transport models

Participant: Michael Walker.

This research theme is supported by the ANR GeneShape and ERASysBio+ iSAM projects.

Active transport of the plant hormone auxin has been shown to play a key role in the initiation of organs at the shoot apex, and vein formation in both leaves and the shoot apical meristem. Polar localized membrane proteins of the PIN1 and AUX/LAX family facilitate this transport and observations and models suggest that the coherent organization of these proteins in the L1 layer is responsible for the creation of auxin maxima (surrounded by a depletion zone), which in turn triggers organ initiation close to the meristem center [39] [1]. Furthermore, canalized PIN allocations are thought to play a crucial role in vein formation in the leaf and in the L2.

Previous studies have typically modeled the L1 and L2 with different models to explain their different PIN allocations. For example [25] used the so-called ‘up-the-gradient’ model in the L1 to recreate the depletion zone around a maximum, and a ‘flux-mediated’ model to produce the midvein growing through the L2. A more unified approach was that of Stoma et.al. [10] who used the flux-mediated model in both tissues, but with different powers of the auxin flux, i.e. with a linear function in the L1 and a quadratic one in the L2.

We seek a completely unified model with no difference in the PIN allocation model. Our approach is based on inherent differences between the L1 and L2, specifically their dimensionality and the distribution of sources and sinks. We also use a flux-feedback function whose power is intermediate between one and two. For powers close to but slightly higher than one, the diffuse PIN distribution is retained in the L1 but a canalized one is formed in the L2.

We are also comparing two and three dimensional simulations to test the common approximation of three-dimensional tissues being modeled as two-dimensional. We find that source-sink models are well-approximated in two dimensions while sink- and source-driven systems show qualitatively different behavior.

A paper on this unified model is being written. This work is done in the context of the Geneshape and iSAM projects.

5.2.4. Mechanical model

Participants: Jérôme Chopard, Christophe Godin, Frédéric Boudon, Jan Traas, Olivier Hamant [ENS-Lyon], Arezki boudaoud [ENS-Lyon].

This research theme is supported by the ANR VirtualFlower and Geneshape projects.

The rigid cell walls that surround plant cells is responsible for their shape. These structures are under constraint due to turgor pressure inside the cell. To study the overall shape of a plant tissue and morphogenesis, its evolution through time, we therefore need a mechanical model of cells. We developed such a model, in which walls are characterized by their mechanical properties like the Young modulus which describes the elasticity of the material. Wall deformation results from forces due to turgor pressure. Growth results from an increase in cell wall synthesis when this deformation is too high. The final shape of the tissue integrates mechanically all the local deformations of each cell.
To model this process, we used a tensorial approach to describe both tissue deformation and stresses. Deformations were decomposed into elementary transformations that can be related to underlying biological processes. However, we showed that the observed deformations does not map directly local growth instructions given by genes and physiology in each cell. Instead, the growth is a two-stage process where genes are specifying by their activity a targeted shape for each cell (or small homogeneous region) and the final cell shape results from the confrontation between this specified shape and the physical constraints imposed by the cell neighbors. Hence the final shape of the tissue results from the integration of all these local rules and constraints at organ level. This work is being described in a paper which will be submitted for publication at the beginning of 2011.

5.2.5. Gene regulatory networks

Modeling gene activities within cells is of primary importance since cell identities correspond to stable combination of gene level activity.

- **The auxin signaling pathway** (Etienne Farcot, Yann Guédon, Christophe Godin, Yassin Refahi, Jonathan Legrand, Jan Traas, Teva Vernoux)

  The auxin signalling network involves about 50 potentially interacting factors. We applied a graph clustering method \cite{16} that relies on 0/1 interactions between factors deduced from yeast two-hybrid (Y2H) data. The Y2H analysis involves two independent tests (X-gal and HIS3 tests). Each possible interaction was tested in the two possible configurations, where each protein was alternatively the bait and the prey protein. A binary interaction is thus a summary of the four outputs of the X-gal and HIS3 tests. In order to limit the loss of information, we designed a standardization procedure to summarize the outputs of the X-gal and HIS3 tests as a distance defined on a continuous scale. This opens the possibility to studies the influence of phylogenetic distances between factors on their interactions using an extension of the mixture model for random graphs that incorporate explanatory variables. This work is the object of a collaboration with Jean-Benoist Leger and Stéphane Robin (MIA, AgroParisTech/INRA).

  As an output to this interaction extraction, it was possible to apply a clustering procedure (work of Y. Guédon and J. Legrand), based on a mixture model of random graphs. This allowed us to determine a well-founded simplified network, which was then used to develop a differential equation model. This model showed that the meristem could present well differentiated buffering abilities in its role of auxin perception, a prediction that was corroborated by experiments using a newly developed auxin sensor. The combination of computational techniques and novel experimental tools and data that were developed in this project have led to a new view of the meristem behavior, where auxin signaling, and not only its transport, plays a crucial role. These results have appeared in a systems biology journal with wide readership \cite{16}.

  As a follow up, extensions of the differential equation model are under study, with the aim of better understanding this system in more general contexts than the shoot apical meristem development.

- **Complex dynamics and spatial interactions in gene networks** (Yassin Refahi, Etienne Farcot, Christophe Godin)

  Complex computational and mathematical questions arise in the study of gene networks at two levels: (i) the single cell level, due to complex, nonlinear interactions, (ii) the tissue level, where multiple cells interact through molecular signals and growth, so that even simple local rules can challenge our intuition at higher scales.

  At the single cell level, new results were obtained in the framework of piecewise-linear models. Since their introduction in the late 1960’s, these models have been believed to present chaotic behavior in some parameter regimes. However, this was mostly observed numerically, based on intensive generation of random networks. In a collaboration between E. Farcot and R. Edwards (Univ. Victoria, Canada), with recent input from one of his students, E. Foxall, we have introduced a method to explicitly build piecewise affine models having a return map which is conjugate to a topological horseshoe. A paper presenting these results is currently in revision.
At a higher scale, we have also continued the study of gene regulation in meristematic tissues. Numerical tools for the simulation of gene network models have been further developed, in the context of Y. Refahi’s thesis. Using realistic 3D and 4D structures, reconstructed from microscopy data using the methods in [5], we have defined putative gene expression patterns in collaboration with biologist colleagues in Lyon (P. Das). We then have constructed an adimensionalized model with only 3 parameters, based on the existing literature [37]. Using standard optimization procedures to fit these parameters we were able to reproduced the in silico patterns with great accuracy. Moreover, alternative patterns were found, based on hypotheses of continuous deformation of gene patterns under tissue growth. These predictions remain to be tested, and this will constitute a large part of Y. Refahi’s post-doctoral project (INRA/INRIA CJS fund), in collaboration with ENS Lyon and the team of Henrik Jönsson in Lund, Sweden.

A related study has been initiated during a visit of E. Farcot at the Memorial University of Newfoundland, funded by the INRIA "Explorateur" program (see international relations section). This visit has allowed to initiate a mathematical study of gene regulated patterning in plants, with the aid of tools from equivariant dynamical systems theory, and more generally by relying on symmetry properties. Some first results have been obtained concerning homogeneous patterns in regular tissues, and their bifurcations.

5.2.6. Model integration

Participants: Mikaël Lucas, Michael Walker, Jérôme Chopard, Frédéric Boudon, Christophe Godin, Laurent Laplaze, Jan Traas, François Parcy.

This research theme is supported by the ANR/BBSRC project iSam.

Our approach consists in building a programmable tissue which is able to accept different modeling components. This includes a central data structure representing the tissue in either 2-D or 3-D, which is able to grow in time, models of gene activity and regulation, models of signal exchange (physical and chemical) between cells and models of cell cycle (which includes cell division). For each modeling component, one or several approaches are investigated in depth, possibly at different temporal and spatial scales, using the data available from the partners (imaging, gene networks, and expression patterns). Approaches are compared and assessed on the same data. The objective of each submodel component will be to provide plugin components, corresponding to simplified versions of their models if necessary, that can be injected in the programmable tissue platform.

- **Development of a computer platform for the ’programmable tissue’.** (Jérôme Chopard, Michael Walker, Frédéric Boudon, Etienne Farcot, Christophe Godin)

  One key aspect of our approach is the development of a computer platform dedicated to programming virtual tissue development. This platform will be used to carry out integration of the different models developed in this research axis. The platform is based on OpenAlea. Partner models can be integrated in the platform in a non-intrusive way (the code of their model need not be rewritten). In this context, model integration will i) consist of designing adequate data-structures at different levels that will be exchanged and reused among the different plug-in models and ii) defining control flows at adequate levels to avoid the burden of excessive interaction between components.

- **Design of a genetic model of inflorescence development.** (Etienne Farcot, Christophe Godin, François Parcy)

  At a scale involving organs and their geometric arrangement, we have developed a first model of the control of floral initiation by genes, and in particular the situation of cauliflower mutants, in which the meristem fails in making a complete transition to the flower. This work couples models at different scales, since gene regulation is described by a minimal gene network, which is used as a decision module in an L-system model of the inflorescence architecture. This coupled model has led us to hypothesize some interactions between genes and a particular plant hormone, and experiments are currently being made to verify this prediction.
6. New Results

6.1. Image Segmentation, Registration and Analysis

6.1.1. Quantitative Analysis of Open Curves in Brain Imaging: Applications to White Matter Fibers and Sulci

Participants: Meena Mani, Christian Barillot.

Shape, scale, orientation and position, the four physical features associated with open curves, have different properties so the usual approach has been to design different metrics and spaces to treat them individually. We took an alternative approach using a comprehensive Riemannian framework where joint feature spaces allow for analysis of combinations of features. We can compare curves by using geodesic distances, which quantify their differences. We validated the metrics we used, demonstrated practical uses and applied the tools to important clinical problems. To begin, specific tract configurations in the corpus callosum are used to showcase clustering results that depend on the Riemannian distance metric used. This nicely argues for the judicious selection of metrics in various applications, a central premise in our work. The framework also provides tools for computing statistical summaries of curves. We represented fiber bundles with a mean and variance, which describes their essential characteristics. This is both a convenient way to work with a large volume of fibers and is a first step towards statistical analysis. Next, we designed and implemented methods to detect morphological changes, which can potentially track progressive white matter disease. With sulci, we addressed the specific problem of labeling. An evaluation of physical features and methods such as clustering leads us to a pattern matching solution in which the sulcal configuration itself is the best feature.

6.1.2. Trimmed-likelihood estimation for focal lesions and tissue segmentation in multisequence MRI for multiple sclerosis

Participants: Sylvain Prima, Christian Barillot.

Following Daniel Garcia-Lorenzo’s PhD, we proposed a new automatic method for segmentation of multiple sclerosis (MS) lesions in magnetic resonance images. The method performs tissue classification using a model of intensities of the normal appearing brain tissues. In order to estimate the model, a trimmed likelihood estimator is initialized with a hierarchical random approach in order to be robust to MS lesions and other outliers present in real images. The algorithm was first evaluated with simulated images to assess the importance of the robust estimator in presence of outliers. The method was then validated using clinical data in which MS lesions were delineated manually by several experts. Our method obtains an average Dice similarity coefficient (DSC) of 0.65, which is close to the average DSC obtained by raters (0.66) [15].

6.1.3. Segmentation of Multimodal Brain Images using Spectral Gradient and Graph Cut

Participants: Camille Maumet, Jean-Christophe Ferré, Christian Barillot.

Following Jeremy Lecoeur’s PhD, we have introduced a new and original scale-space approach for segmenting normal and pathological tissue from multidimensional images. This method can perform a joint segmentation of three complementary imaging volumes at the same time by embedding a scale-space color invariant edge detector - i.e. the spectral gradient - as the boundary term in a graph cut optimization framework. Finally, we have proposed to extend this new scheme to more than three channels. We focussed the contribution onto the segmentation of tissues or structures of interest from multi-dimensional / multi-sequences brain MRI. This new multidimensional segmentation framework has been validated on simulated data and on clinical data (both pathological and healthy brains). We have exhibited the performances of this new method on various combinations of MRI sequences for the segmentation of normal and pathological tissues and showed how it is able to out perform competitive works. This work is under submission to an international journal.
6.1.4. Adaptive pixon represented segmentation for 3D MR brain images based on mean shift and Markov random fields  
**Participant:** Christian Barillot.

Following Lei Lin and Daniel Garcia Lorenzo’s PhDs, we proposed an adaptive pixon represented segmentation (APRS) algorithm for 3D magnetic resonance (MR) brain images. Different from traditional methods, an adaptive mean shift algorithm was adopted to adaptively smooth the query image and create a pixon-based image representation. Then K-means algorithm was employed to provide an initial segmentation by classifying the pixons into a predefined number of tissue classes. By using this segmentation as initialization, expectation-maximization (EM) iterations composed of bias correction, a priori digital brain atlas information, and Markov random field (MRF) segmentation were processed. Pixons were assigned with final labels when the algorithm converges. The adoption of bias correction and brain atlas made the current method more suitable for brain image segmentation than the previous pixon based segmentation algorithm. The proposed method was validated on both simulated normal brain images from BrainWeb and real brain images from the IBSR public dataset. Compared with some other popular MRI segmentation methods, the proposed method exhibited a higher degree of accuracy in segmenting both simulated and real 3D MRI brain data. The experimental results were numerically assessed using Dice and Tanimoto coefficient [18].

6.1.5. EM-ICP strategies for joint mean shape and correspondences estimation: applications to statistical analysis of shape and of asymmetry  
**Participant:** Sylvain Prima.

In collaboration with B. Combès, we proposed a new approach to compute the mean shape of unstructured, unlabelled point sets with an arbitrary number of points. This approach can be seen as an extension of the EM-ICP algorithm, where the fuzzy correspondences between each point set and the mean shape, the optimal non-linear transformations superposing them, and the mean shape itself, are iteratively estimated. Once the mean shape is computed, one can study the variability around this mean shape (e.g. using PCA) or perform statistical analysis of local anatomical characteristics (e.g. cortical thickness, asymmetry, curvature). To illustrate our method, we performed statistical shape analysis on human osseous labyrinths and statistical analysis of global cortical asymmetry on control subjects and subjects with situs inversus [29]. This work was led within the ARC 3D-MORPHINE (http://3dmorphine.inria.fr).

6.1.6. Surface-based method to evaluate global brain shape asymmetries in human and chimpanzee brains  
**Participant:** Sylvain Prima.

Following PhD and PostDoc works from Benoit Combès and Marc Fournier, in this work we used humans and chimpanzees brain MRI databases to develop methods for evaluating global brain asymmetries. We performed brain segmentation and hemispheric surface extraction on both populations. The human brain segmentation pipeline was adapted to chimpanzees in order to obtain results of good quality. To alleviate the problems due to cortical variability we proposed a mesh processing algorithm to compute the brain global shape. Surface-based global brain asymmetries were computed on chimpanzee and human subjects using individual mid-sagittal plane evaluation and population-level mean shape estimation. Asymmetry results were presented in terms of axis-wise components in order to perform more specific evaluation and comparison between the two populations [35]. This work was led within the ARC 3D-MORPHINE (http://3dmorphine.inria.fr).

6.1.7. Computational techniques for the analysis of endocranial cast and endocranial structures  
**Participant:** Sylvain Prima.
Following PhD and post-doc works from Benoit Combès and Marc Fournier, a series of studies were led within the ARC 3D-MORPHINE (http://3dmorphine.inria.fr) and were presented at the 1836th Journées de la Société d’Anthropologie de Paris (January 26-28) and at the 80th annual meeting of the American Association of Physical Anthropologists (April 12-16). These include: a method to assess 3D endocranial asymmetries in extant and fossil species: new insights into paleoneurology [48]; a method to map the distance between the brain and the inner surface of the skull [51], [34]; a method to compare bony labyrinths in humans, chimpanzees and baboons [28]; a method for the analysis of the endocranial shape and its relationship with ectocranial structures [41]; a new reconstruction of the frontal lobe and temporal pole of the Taung (Australopithecus africanus) endocast [32].

6.1.8. Evaluation of Registration Methods on Thoracic CT: The EMPIRE10 Challenge

Participant: Olivier Commowick.

We participated, as part of a collaboration with the Asclepios team, to the EMPIRE10 challenge on registration. EMPIRE10 (Evaluation of Methods for Pulmonary Image REgistration 2010) is a public platform for fair and meaningful comparison of registration algorithms which are applied to a database of intra-patient thoracic CT image pairs. Evaluation of non-rigid registration techniques is a non-trivial task. This is compounded by the fact that researchers typically test only on their own data, which varies widely. For this reason, reliable assessment and comparison of different registration algorithms has been virtually impossible in the past. In this work we present the results of the launch phase of EMPIRE10, which comprised the comprehensive evaluation and comparison of 20 individual algorithms from leading academic and industrial research groups. All algorithms are applied to the same set of 30 thoracic CT pairs. Algorithm settings and parameters are chosen by researchers expert in the configuration of their own method and the evaluation is independent, using the same criteria for all participants. All results are published on the EMPIRE10 website (http://empire10.isi.uu.nl). The challenge remains ongoing and open to new participants. Full results from 24 algorithms have been published at the time of writing. This article details the organisation of the challenge, the data and evaluation methods and the outcome of the initial launch with 20 algorithms. More details are available in [20].

6.2. Image processing on Diffusion Weighted Magnetic Resonance Imaging

6.2.1. Diffusion Directions Imaging (DDI)

Participants: Aymeric Stamm, Christian Barillot.

Diffusion magnetic resonance imaging (dMRI) is the reference in vivo modality to study the connectivity of the brain white matter. Images obtained through dMRI are indeed related to the probability density function (pdf) of displacement of water molecules subject to restricted diffusion in the brain white matter. The knowledge of this diffusion pdf is therefore of primary importance. Several methods have been devised to provide an estimate of it from noisy dMRI signal intensities. They include popular diffusion tensor imaging (DTI) as well as higher-order methods. These approaches suffer from important drawbacks. Standard DTI cannot directly cope with multiple fiber orientations. Higher-order approaches can alleviate these limitations but at the cost of increased acquisition time. We have proposed, in the same vein as DTI, a new parametric model of the diffusion pdf with a reasonably low number of parameters, the estimation of which does not require acquisitions longer than those used in clinics for DTI. This model also accounts for multiple fiber orientations. It is based on the assumption that, in a voxel, diffusing water molecules are divided into compartments. Each compartment is representative of a specific fiber orientation (which defines two opposite directions). In a given compartment, we further assume that water molecules that diffuse along each direction are in equal proportions. We then focus on modeling the pdf of the displacements of water molecules that diffuse only along one of the two directions. Under this model, we derive an analytical relation between the dMRI signal intensities and the parameters of the diffusion pdf. We exploit it to estimate these parameters from noisy signal intensities. We carry out a cone-of-uncertainty analysis to evaluate the accuracy of the estimation of the fiber orientations and we evaluate the angular resolution of our method. Finally, we show promising results on real data and propose a visualization of the diffusion parameters which is very informative to the neurologist. This work was conducted in collaboration with Patrick Perez from Technicolor [56].
6.2.2. Anatomy of the corticospinal tracts: evaluation of a deterministic tractography method

Participants: Romuald Seizeur, Nicolas Wiest-Daesslé, Sylvain Prima, Camille Maumet, Jean-Christophe Ferré, Xavier Morandi.

In this work, anatomical, diffusion-weighted and functional 3T MRI were acquired on 15 right-handed healthy subjects to analyse the portions of the corticospinal tract (CST) dedicated to hand motor and sensory functions. The three MR images were then registered and regions of interest were delineated i) in the mid-brain using 3D T1-weighted MRI, and ii) in the cortex using fMRI using hand motor and sensory tasks. Deterministic tractography was then performed using these two ROIs from diffusion-weighted MRI after the diffusion tensors were computed. The ventrolateral tract fibers of the CST were generally not properly identified, due to fiber crossing in the corona radiata [55].

6.2.3. Tracking of the Hand Motor Fibers within the Corticospinal Tract Using Functional, Anatomical and Diffusion MRI

Participants: Romuald Seizeur, Nicolas Wiest-Daesslé, Olivier Commowick, Sylvain Prima, Aymeric Stamm, Christian Barillot.

In this work, we proposed to compare three diffusion models to track the portion of the corticospinal tract dedicated to the hand motor function (called hand motor fibers hereafter), using diffusion, functional and anatomical MRI. The clinical diffusion data have few gradient directions and low b-values. In this context, we show that a newly introduced model, called diffusion directions imaging (DDI) outperforms both the DTI and the ODF models. This new model allows to capture several diffusion directions within a voxel, with only a low number of parameters. Two important results are that i) the DDI model is the only one allowing consistent tracking from the mesencephalon to the most lateral part of the cortical motor hand area, and that ii) the DDI model is the only model able to show that the number of hand motor fibers in the left hemisphere is larger than in the contralateral hemisphere for right-handed subjects; the DDI model, as the other two models, fails to find such a difference for left-handed subjects. To the best of our knowledge, this is the first time such results are reported, at least on clinical data. [44].

6.2.4. Multifiber Deterministic Streamline Tractography Based on a New Diffusion Model

Participants: Olivier Commowick, Romuald Seizeur, Nicolas Wiest-Daesslé, Sylvain Prima, Aymeric Stamm, Christian Barillot.

In this work, we have built upon a new model, describing the random motion of water molecules in fibrous tissues, to develop a multifiber deterministic tractography algorithm. We apply this algorithm to track the corticospinal tract of the human brain, in both controls and patients with tumors. [31].

6.2.5. Automated detection of white matter fiber bundles

Participant: Olivier Commowick.

This work is part of a collaboration with the Computational Radiology Laboratory headed by Simon Warfield in Boston, USA. For this topic, we have studied how white matter fiber bundles can be extracted in a reproducible way from diffusion tensor MRI. Usually, white matter (WM) fiber bundles of the brain can be delineated by diffusion tractography utilizing anatomical regions-of-interest (ROI). These ROIs can specify seed regions in which tract generation algorithms are initiated. Interactive identification of such anatomical ROIs enables the detection of the major WM fiber tracts, but suffers from inter-rater and intra-rater variability, and is time consuming. We developed and compared three techniques for automated delineation of ROIs for the detection of two major WM fiber tracts in 12 healthy subjects. Tracts identified automatically were compared quantitatively to reference standard tracts derived from carefully hand-drawn ROIs. Based on comparative performance of the experimental techniques, a multi-template label fusion algorithm was found to generate tracts most consistent with the reference standard. More details on this work are available in [43].
6.3. Management of Information and Semantic Processing

6.3.1. NeuroLOG project: Sharing of data and sharing of processing tools in neuroimaging

Participants: Bernard Gibaud, Bacem Wali.

The NeuroLOG project (ANR ANR-06-TLOG-024) came to its end in December 2010. However, we managed to maintain the NeuroLOG platform in operation, which is important with regards to publication. Several papers are in preparation. A lot of efforts were devoted in 2011 to submit a new proposal to ANR, building on NeuroLOG’s achievements. A NeuroLOG2 project was submitted in March to the ANR TECSAN program (health technology). This new project aimed both at going on developing the technology for sharing data and processing tools, while being more involved in neuroimaging applications. Two application fields were proposed, concerning research on Alzheimer Disease, on the one hand, and epilepsy, on the other hand. The consortium was enlarged accordingly, with the integration of new partners such as the EDELWEISS project (INRIA, Sophia), the U642 LTSI (INSERM, Rennes) and U1028 CNRL (INSERM, Lyon). A new submission is envisaged in 2012, taking into account the recommendations of ANR.

6.3.2. Semantic annotation of anatomic images in neuroimaging

Participants: Bernard Gibaud, Tristan Moreau, Xavier Morandi.

This project aims at exploring the feasibility of relying on symbolic knowledge provided by ontologies to assist the annotation of anatomical images. The basic assumption underlying this work is that ontologies not only can provide a reference vocabulary to annotate images, but they can also provide useful prior knowledge that may help the annotation process itself, an assumption supported by the interesting results obtained by Ammar Mechouche in his PhD work. The current study, initiated in 2010 in the context of the Master student work of Elsa Magro (analysis of intra-precentral connections and of the U-fibers connecting the precentral and postcentral gyri) was pursued in 2011 (PhD work of Tristan Moreau). Our most recent works try to establish a parcellation of the grey-matter white matter surface based on the connectivity profiles of individual points of this surface, valid for a population of subjects. This is a prerequisite before identifying the more salient fiber bundles to be modelled in our ontology.

6.3.3. Semantic annotation of models and simulated medical images

Participants: Bernard Gibaud, Germain Forestier.

This project is carried out in the context of the Virtual Imaging Platform (VIP) project, an ANR project aiming at setting up a platform for facilitating the use of image simulation software in medical imaging, and coordinated by Creatis (Lyon). The platform will integrate simulation software to generate image of different modalities (i.e. MR, CT, PET, US). In this project, VISAGES is in charge of coordinating the development of an application ontology to support the annotation of the data shared in this platform (simulated images, anatomical models and physiological models used in simulations), as well as the annotation of simulation software components, in order to facilitate their interoperability within the platform. The work completed in 2011 is a continuation of what was started in 2010. Our major result in 2011 is an ontology allowing to annotate the models used for medical image simulations. Actually models are composed of files containing images (3D voxel maps) or surfaces (meshes). Our ontology includes entities called model layers associated with those files and depicting the model contents in terms of : anatomical structures, pathological structures, foreign bodies, contrast agents etc. Each individual object present in the model is referred to by an object layer part to which physical parameter distributions can be associated, that are used by simulation software. The ontology was modelled as OntoSpec documents (a methodology defined by Gilles Kassel in Amiens), then implemented in OWL. A preliminary version of this model was presented at a workshop organized by EBI in Cambridge in March 2011 (in the context of the VPH/RICORDO project). A more complete version was presented at the CBMS’2011 Conference in Bristol. VIP is a collaborative project, supported by ANR (Agence National de la Recherche), through grant ANR-AA-PPPP-000. The partners with whom we have the tightest relations are: Creatis (Lyon), I3S (Sophia), CEA-LETI (Grenoble).
6.4. Image Guided Intervention

6.4.1. Classification of Surgical Process using Dynamic Time Warping

**Participants:** Pierre Jannin, Germain Forestier, Florent Lalys, Brivael Trelhu.

Toward the creation of new computer-assisted intervention systems, Surgical Process Models (SPMs) are more and more used as a tool for analyzing and assessing surgical interventions. SPMs represent Surgical Process (SPs) which are defined as symbolic structured descriptions of surgical interventions, using a pre-defined level of granularity and a dedicated terminology. In this context, an important challenge is the creation of new metrics for the comparison and the evaluation of SPs. Thus, correlations between these metrics and pre-operative data allow to classify surgeries and highlight specific information on the surgery itself and on the surgeon, such as its level of expertise. In this study, we explored the automatic classification of a set of SPs based on the Dynamic Time Warping (DTW) algorithm. DTW allows to compute a distance between two SPs that focuses on the different types of activities performed during the surgery and their sequencing, by minimizing the time differences. Indeed, it turns out to be a complementary approach to classical methods focusing only on the time and the number of activities differences. Experiments were carried out on 24 lumbar disc herniation surgeries to discriminate the level of expertise of surgeons according to prior classification of SPs. Supervised and unsupervised classification experiments have shown that this approach was able to automatically identify groups of surgeons according to their level of expertise (senior and junior), and opens many perspectives for the creation of new metrics for surgeries comparison and evaluation. This work was performed in collaboration with Dr. Laurent Riffaud, and was published in the International Journal of Biomedical Informatics [14].

6.4.2. Surgical phases detection from microscope videos by machine learning

**Participants:** Pierre Jannin, Florent Lalys, Xavier Morandi.

Surgical process analysis and modeling is a recent and important topic aiming at introducing a new generation of computer-assisted surgical systems. Among all of the techniques already in use for extracting data from the Operating Room, the use of image videos allows automating the surgeons’ assistance without altering the surgical routine. In collaboration with Carl Zeiss Medical Systems (Oberkochen, Germany), we proposed an application-dependent framework able to automatically extract the phases of the surgery only by using microscope videos as input data and that can be adaptable to different surgical specialties. First, four distinct types of classifiers based on image processing were implemented to extract visual cues from video frames. Each of these classifiers was related to one kind of visual cue: visual cues recognizable through color were detected with a color histogram approach, for shape-oriented visual cues we trained a Haar classifier, for texture-oriented visual cues we used a bag-of-word approach with SIFT descriptors, and for all other visual cues we used a classical image classification approach including a feature extraction, selection, and a supervised classification. The extraction of this semantic vector for each video frame then permitted to classify time series using either Hidden Markov Model or Dynamic Time Warping algorithms. The framework was validated on cataract surgeries, obtaining accuracies of 95%. This work was performed in collaboration with Laurent Riffaud and was published at the ORASIS and MICCAI conferences.

6.4.3. Surgical tools recognition and pupil segmentation for cataract surgery modeling

**Participants:** Pierre Jannin, Florent Lalys.

In the above project work performed through the MS internship of David Bouget, we focus on developing an application-depandent framework able to extract surgical phases from microscope videos. The aim of this study was to enhance results of this framework by adding new visual cues extraction modules. We studied two modules: one to segment the pupil and one to extract and recognize surgical tools. Validation studies, performed with cataract surgery videos, show an increase of the framework accuracy to detect eight surgical phases. This work has been accepted at the MMVR 2012 international conference.

6.4.4. Automatic computation of electrode trajectories for Deep Brain Stimulation: a hybrid symbolic and numerical approach

**Participants:** Pierre Jannin, Florent Lalys, Camille Maumet, Claire Haegelen.
The optimal electrode trajectory is needed to assist surgeons in planning Deep Brain Stimulation (DBS). We developed and tested a method for image-based trajectory planning. Rules governing the DBS surgical procedure were defined with geometric constraints. A formal geometric solver using multimodal brain images and a template built from 15 brain MRI scans were used to identify a space of possible solutions and select the optimal one. For validation, a retrospective study of 30 DBS electrode implantations from 18 patients was performed. A trajectory was computed in each case and compared with the trajectories of the electrodes that were actually implanted. Computed trajectories had an average difference of 6.45 degrees compared with reference trajectories and achieved a better overall score based on satisfaction of geometric constraints. Trajectories were computed in 2min for each case. We demonstrated that a rule-based solver using pre-operative MR brain images can automatically compute relevant and accurate patient-specific DBS electrode trajectories. This work was published in the International Journal of Computer Assisted Radiology and Surgery.

6.4.5. Analysis of electrodes’ placement and deformation in deep brain stimulation from medical images

Participants: Pierre Jannin, Florent Lalys, Alexandre Abadie, Xavier Morandi, Claire Haegelen.

This work was performed during the internship of Maroua Mehri. Deep brain stimulation (DBS) is used to reduce the motor symptoms such as rigidity or bradykinesia, in patients with Parkinson’s disease (PD). The Subthalamic Nucleus (STN) has emerged as prime target of DBS in idiopathic PD. However, DBS surgery is a difficult procedure requiring the exact positioning of electrodes in the pre-operative selected targets. This positioning is usually planned using patients’ pre-operative images, along with digital atlases, assuming that electrode’s trajectory is linear. However, it has been demonstrated that anatomical brain deformations induce electrode’s deformations resulting in errors in the intra-operative targeting stage. In order to meet the need of a higher degree of placement accuracy and to help constructing a computer-aided-placement tool, we studied the electrodes’ deformation in regards to patients’ clinical data (i.e., sex, mean PD duration and brain atrophy index). Firstly, we presented an automatic algorithm for the segmentation of electrode’s axis from post-operative CT images, which aims to localize the electrodes’ stimulated contacts. To assess our method, we applied our algorithm on 25 patients who had undergone bilateral STNDBS. We found a placement error of 0.91 +/- 0.38 mm. Then, from the segmented axis, we quantitatively analyzed the electrodes’ curvature and correlated it with patients’ clinical data. We found a positive significant correlation between mean curvature index of the electrode and brain atrophy index for male patients and between mean curvature index of the electrode and mean PD duration for female patients. These results help understanding DBS electrode’ deformations and would help ensuring better anticipation of electrodes’ placement. This work has been accepted at the SPIE Medical Imaging 2012 conference.

6.5. Medical Image Computing in Brain Pathologies

6.5.1. Detection of cortical abnormalities in drug resistant epilepsy

Participants: Elise Bannier, Camille Maumet, Jean-Christophe Ferré, Jean-Yves Gauvrit, Christian Barillot.

Focal cortical dysplasia and heterotopias are a recognized cause of epilepsy. Indication for drug resistant epilepsy surgery relies on precise localization and delineation of the epileptogenic zone and lesion identification is an important issue. Visual detection and delineation of small or occult focal cortical dysplasia and heterotopias on MR images are sometimes difficult. The Double Inversion Recovery (DIR) imaging, by nulling white matter and cerebrospinal fluid signal, seems particularly appropriate to detect intracortical lesions in MS and Epilepsy. In this work we evaluated at 3T and using voxel-based morphometry (VBM) the ability of a 9-minute 3D DIR sequence to detect cortical and juxtacortical lesions in drug resistant epileptic patients. Results on 21 patients and 20 healthy volunteers show the potential of 3D DIR VBM to detect cortical abnormalities. Further work will investigate the use of alternate registration frameworks (e.g. DARTEL), improved intensity normalization of 3D DIR images and joint 3D T1-w/DIR analysis to improve detection sensitivity and specificity.
6.5.2. Multi-modal NMR cartography of USPIO positive and negative tissues in MS human models

**Participants:** Olivier Luong, Olivier Commowick, Christian Barillot.

The main objective of this work was to build an input object for an MRI simulator. Each voxel of the object is defined by its three physical entities which are $T_1$, $T_2$ and $\rho$ MR relaxation parameters. In our case, this object comes from Multiple Sclerosis brains. We initially defined a simplified model with respect to pathological regions, based on a combination of the Brainweb template and the lesion manually delineated from pathological images. From this, we allocated relaxation parameters for each voxels of these ROI based on fixed values of $T_1$, $T_2$ and $\rho$ (initialized from in vivo relaxometry acquisitions). This model model does not allow to obtain a fine description of the pathological regions, as potentially defined by differential contrats between USPIO and Gd enhanced images. In order to obtain this finer description, we used an MRI simulator based on the Bloch’s equations, in order to estimated the $T_1$, $T_2$ and $\rho$ parameters on each voxel from initial conditions coming from in-vivo images acquired in Rennes by using the USPIO-6 protocol.

This work is part of the VIP collaborative project, supported by ANR (Agence National de la Recherche), through grant ANR-AA-PPPP-000. The partners with whom we have the tightest relations are: Creatis (Lyon), I3S (Sophia), CEA-LETI (Grenoble).

6.6. Vascular Imaging and Arterial Spin Labelling

6.6.1. Arterial spin labeling for motor activation mapping at 3T

**Participants:** Jan Petr, Aymeric Stamm, Elise Bannier, Jean-Christophe Ferré, Jean-Yves Gauvrit, Christian Barillot.

Functional arterial spin labeling (fASL) is an innovative biomarker of neuronal activation that allows direct and absolute quantification of activation-related CBF and is less sensitive to venous contamination than BOLD fMRI. This study evaluated fASL for motor activation mapping in comparison with BOLD fMRI in terms of involved anatomical area localization, intra-individual reproducibility of location, quantification of neuronal activation, and spatial accuracy. Imaging was performed at 3T with a 32-channel coil and dedicated post-processing tools were used. Twelve healthy right-handed subjects underwent fASL and BOLD fMRI while performing a right hand motor activation task. Three sessions were performed 7 days apart in similar physiological conditions. Our results showed an activation in the left primary hand motor area for all 36 sessions in both fASL and BOLD fMRI. The individual functional maps for fASL demonstrated activation in ipsilateral secondary motor areas more often than the BOLD fMRI maps. This finding was corroborated by the group maps. In terms of activation location, fASL reproducibility was comparable to BOLD fMRI, with a distance between activated volumes of 2.1 mm and an overlap ratio for activated volumes of 0.76, over the 3 sessions. In terms of activation quantification, fASL reproducibility was higher, although not significantly, with a CVintra of 11.6% and an ICC value of 0.75. Functional ASL detected smaller activation volumes than BOLD fMRI but the areas had a high degree of co-localization. In terms of spatial accuracy in detecting activation in the hand motor area, fASL had a higher specificity (43.5%) and a higher positive predictive value (69.8%) than BOLD fMRI while maintaining high sensitivity (90.7%). The high intra-individual reproducibility and spatial accuracy of fASL revealed in the present study will subsequently be applied to pathological subjects [25].

6.6.2. Construction and evaluation of a quantitative ASL brain perfusion template at 3T

**Participants:** Jan Petr, Elise Bannier, Jean-Christophe Ferré, Jean-Yves Gauvrit, Christian Barillot.
Arterial spin labeling (ASL) allows non-invasive imaging and quantification of brain perfusion by magnetically labeling blood in the brain-feeding arteries. ASL has been used to study cerebrovascular diseases, brain tumors and neurodegenerative disorders as well as for functional imaging. The use of a perfusion template could be of great interest to study inter-subject regional variation of perfusion and to perform automatic detection of individual perfusion abnormalities. However, low spatial resolution and partial volume effects (PVE) issues inherent to ASL acquisitions remain to be solved. The purpose of this study is to enhance the template quality by using DARTEL non-rigid registration and by correcting for PVE. PICORE-Q2TIPS ASL datasets were acquired on 25 healthy volunteers at 3T. Four methods of creating the template were evaluated using leave-one-out cross correlation. Subsequently, these methods were applied to hyper-perfusion detection on functional ASL data of 8 healthy volunteers and compared with the standard generalized linear model (GLM) activation detection [40].

6.6.3. Evaluation of functional arterial spin labeling data using a perfusion template

**Participants:** Jan Petr, Elise Bannier, Jean-Christophe Ferré, Jean-Yves Gauvrit, Christian Barillot.

ASL allows non-invasive imaging and quantification of brain perfusion by magnetically labeling blood in the brain-feeding arteries. In this study, a template created from perfusion images of 25 resting healthy subjects was used to automatically detect hyper perfusion patterns of 8 other subjects. DARTEL registration was used to improve the precision of the template and partial volume correction to prevent interpolation artifacts. This study showed that a perfusion template can be used to assess task-related activation zones in functional ASL data while using only activated phase. Two assumptions can be made to explain why standard functional analysis yields slightly larger activation regions. First, the use of FWHM 6mm Gaussian kernel possibly enlarges the detected zones. Second, the data analyzed using SPM contains both resting and activated phases whereas only the activated phase was compared to the template. Future work will focus on detection of hyperperfusion in different neurodegenerative diseases taking into account registration issues of pathological T1 images. [24].

6.6.4. A contrario detection of focal brain perfusion abnormalities based on an ASL template

**Participants:** Camille Maumet, Elise Bannier, Jean-Christophe Ferré, Pierre Maurel, Christian Barillot.

Arterial Spin Labeling (ASL) is a recent MRI perfusion technique which enables quantification of cerebral blood flow (CBF). The presence of regions with atypical CBF can characterize a pathology. In brain tumors for instance, perfusion increase can be directly related to the grading of the malignant tissues. It is therefore of great interest to identify these regions in order to provide the patients with the most appropriate therapy. In this work, we proposed to detect abnormal brain perfusion by using an a contrario framework and an ASL template as a model of normal perfusion. Validation was undertaken by qualitative comparison with CBF extracted from dynamic susceptibility weighted contrast enhanced (DSC) sequence. We experimented this framework on four patients presenting brain tumors. Results show that high perfusion regions found in DSC CBF maps are correctly identified as hyperperfusions with a contrario detection. Automatic detection has clear advantages over manual delineation since it is less time-consuming, does not depend on medical expertise and allows quantification of perfusion abnormalities within the detected regions.

6.6.5. Peripheral angiography using non-contrast enhanced imaging

**Participants:** Elise Bannier, Isabelle Corouge, Nicolas Wiest-Daesslé.

Arteriography, CT and MR angiography are routinely performed in patients presenting peripheral arteriopathy. Yet, contrast agent injection is contraindicated in patients with renal insufficiency and the underlying risk of developing nephrogenic systemic fibrosis further encourages research on non-contrast enhanced MR angiography techniques (NCE MRA). In this context, we evaluated at 3T the ability of a 2 NCE MRA new sequences to reliably detect peripheral vascular abnormalities from the abdominal aorta to the calf in comparison with CE MRA.
A first study including 20 healthy volunteers and 4 patients evaluated the NCE ECG-gated T2 TSE NATIVE SPACE MRI sequence. It demonstrated its potential in noninvasively imaging peripheral vasculature, from the abdominal aorta to the calf, within a clinically acceptable acquisition duration. Although signal inhomogeneity and peristalsis artifacts were sometimes observed in the abdominal aortic station, very good image quality was obtained on all subjects on lower stations, with no venous contamination.

A second study evaluated the NCE ECG-gated Quiescent Interval Single Shot (QISS) MRA sequence. Preliminary results obtained on 11 patients show that several lesions were not detected with QISS MRA especially on the thigh station. Ongoing patient inclusions are required to confirm these results. Finally, a concomitant NCE and CE MRA reading will be performed to compare stenosis grading, stenosis-thrombosis mismatch and lesions not detected with NCE MRA.

6.7. Abnormal functional lateralization and activity of language brain areas in developmental dysphasia

Participants: Clément De Guibert, Camille Maumet, Jean-Christophe Ferré, Pierre Jannin, Christian Barillot.

Atypical functional lateralization and specialization for language have been proposed to account for developmental language disorders, yet results from functional neuroimaging studies are sparse and inconsistent. This functional magnetic resonance imaging study compared children with a specific subtype of specific language impairment affecting structural language, to a matched group of typically developing children using a panel of four language tasks neither requiring reading nor metalinguistic skills, including two auditory lexico-semantic tasks (category fluency and responsive naming) and two visual phonological tasks based on picture naming. Data processing involved normalizing the data with respect to a matched pairs paediatric template, groups and between-groups analysis, and laterality indices assessment within regions of interest using single and combined task analysis. Children with specific language impairment exhibited a significant lack of left lateralization in all core language regions (inferior frontal gyrus-opercularis, inferior frontal gyrus-triangularis, supramarginal gyrus and superior temporal gyrus), across single or combined task analysis, but no difference of lateralization for the rest of the brain. Between-group comparisons revealed a left hypoactivation of Wernicke’s area at the posterior superior temporal/supramarginal junction during the responsive naming task, and a right hyperactivation encompassing the anterior insula with adjacent inferior frontal gyrus and the head of the caudate nucleus during the first phonological task. This study thus provides evidence that this subtype of specific language impairment is associated with atypical lateralization and functioning of core language areas [12].
4. New Results

4.1. Spatial Computing approach and RFIDs

Participants: Michel Banâtre, Paul Couderc [contact], Yann Glouche, Arnab Sinha.

In the line of our previous research in pervasive computing, we are working on spatial computing approaches in the context of RFID. Spatial computing consists in data structures and computing processes directly supported by physical objects. RFID is an attractive technology for supporting spatial computing, enabling any object to interact in a smart environment. Traditional RFID solutions use a logical model, where the RFID tags are simple identifiers referring to data in a remote information system. In our approach, we use the memory of the tags to build self-contained data structures and self-describing objects. While featuring interesting properties, such as autonomous operation and high scalability, this approach also raises difficult challenges: the memory capacity of the tags is very limited, requiring compact and efficient data structures. Some results have been achieved for security applications, where we contributed to efficient integrity checking solution for coupled objects. Integrity checking based on objects group can also be used to provide reliable inventory protocol for RFID, which current readers are lacking. A paper has just been submitted on this aspect.

An applicative project (see 5.2.1 ) in the context of domestic waste management is broadly investigating the use of RFID at item level to provide early waste sorting, to avoid incompatible mix of waste and to prevent hazards. An ontology based system has been proposed to determine the possible interactions of tagged products based on their properties and the external conditions.

4.2. Integrity checking with coupled objects

Participants: Michel Banâtre [contact], Paul Couderc, Jean-Francois Verdonck.

Integrity checking is an important concern in many activities, such as logistic, telecommunication or even day to day tasks such as checking for someone missing in a group. While the computing and telecommunication worlds commonly use digital integrity checking, many activities from the real world do not benefit from automatic integrity control mechanisms. RFID technology offers promising perspectives for facing this problem, but also raises strong privacy concerns as most of the RFID-based systems rely on global identification and tracking. Previously we have already designed Ubi-Check to provide an approach aiming at coupling physical objects and enabling integrity control built on local interactions, without the support of a global information system. Ubi-Check led to the development of various novel applications running quite on the same technology. Most of the partners showed a high interest in defining coupled object-based security solutions, but they were lacking the possibility of defining hierarchical couplings. This is that we have studied and implemented this year. We have designed the Ubi-Tree software which strives to deal with those new requirements.

4.2.1. Hierarchical physical object coupling

Ubi-Tree is a new solution/software designed at INRIA aiming at setting-up and reading hierarchical couplings. It relies on a structure in which physical objects (also called fragments) are seen as external nodes of a tree that we call coupling tree. External nodes of a tree are called leaves. In the system, internal nodes are called coupling nodes. Each fragment embeds an RFID tag supporting coupling data. Coupling data stores the coupling tree. Each internal node can be checked, which means a lacking, illegally forged or corrupted node can be detected at any depth of a coupling. Ubi-Tree proposes a new API to create and check hierarchical couplings and an interactive editing GUI is under development.
4.2.2. Coupling tree structure

New algorithms and structure to store and read hierarchical couplings trees in its leaves (RFID tags) have been developed, making it possible to use multiple coupling levels. Let’s take an example: given three physical objects \( o_1, o_2, o_3 \). A user can couple \( o_1 \) and \( o_2 \) together. Let \( N_1 \) be the parent coupling of \( o_1 \) and \( o_2 \): \( N_1 = o_1, o_2 \). Then he couples this coupling with \( o_3 \) to create the \( N_2 = o_1, o_2, o_3 \) coupling node. Figure 2 gives an illustration of the described coupling.

The key idea of the coupling structure is that coupling data are spread in a way that only descendant leaves of a coupling node are required to read it and process its integrity control. This way, \( N_1 \) only needs \( o_1 \) and \( o_2 \) to be read and/or checked as \( o_1 \) and \( o_2 \) are descendants from \( N_1 \) whereas \( o_3 \) is not. This choice enables to process integrity controls at multiple coupling levels. It is very convenient, as an example, if \( o_3 \) is physically separated from \( o_1 \) and \( o_2 \). So if \( o_1, o_2 \) and \( o_3 \) are brought into the field of the RFID reader, \( N_1 \) and \( N_2 \) can be written, read and checked. If \( o_3 \) is not present in its field, \( N_2 \) will not be recovered from the read data but \( N_1 \) can still be read and checked. If checked, \( N_1 \) will notify it was not the root of the coupling tree when the coupling tree was written. This way, the user knows it did not read the whole structure \( o_1 \) and \( o_2 \) are part of. Indeed, coupling nodes can have the following status:

- Valid: the set of detected tags enabled to decode a structure in which the node has the children and the parent it is supposed to have.
- Partial: same as the valid status except the node should have a parent that could not be read from detected tags.
- Incomplete: the node is missing some of its children.

4.2.3. Ongoing work

Today the management algorithms of graphs are located on the memory available RFID reading and writing. However, in the various applications envisaged, only a subset of RFID memory are read/write, others are only accessible in read-only. Currently we are working on the development of our algorithms that take into account this kind of configuration. The other problem we are working on today is the interface provided to users to be sure that the association between RFID tag and physical object is the one that is perceived by our coupling software. The idea is to be able to identify in the right way the RFID tag associated to a physical object when we place one physical object \( O \) onto the support of the antenna linked to the RFID reader. The position of \( O \), and the tag associated to \( O \), in the physical space is determined using a camera coupled with an image recognition algorithm. The result is displayed onto a touch screen. In that way, when we want to couple a set of physical objects \( o_1, o_2, \ldots \) we place sequentially all these objects onto the support of the antenna, and...
from the image of these objects displayed onto the touch screen we touch those we want to couple and activate the coupling operation. This solution is under development.

4.3. Pervasive support for Smart Homes

Participants: Minh Tuan Ho, Michele Dominici, Bastien Pietropaoli, Frédéric Weis [contact].

Pervasive computing involves tight links between real world activities and computing process. While the perception of the real world events can be handled entirely by the application, we think that ad hoc approaches have limitations, in particular the complexity and the difficulty to re-use the code between applications. Instead, we promote the use of system level abstraction that leverage on tangible structures and processes. Important properties of this approach is that applications are, by design, operating in an implicit way ("in the background" of physical processes). They also often exhibits simpler architectures, and "natural" scalability in the sense that being build upon existing real-world process, they are strongly distributed design that relies essentially on local interactions between physical entities. We are applying this approach to "Smart Homes". A Smart Home is a residence equipped with computing and information technology devices conceived to collaborate in order to anticipate and respond to the needs of the occupants, working to promote their comfort, convenience, security and entertainment while preserving their natural interaction with the environment.

4.3.1. Definition of a system architecture

In a classical "logical" approach, all the intelligence of the Smart Home is condensed in a single entity that takes care of every device in the house. The sensors distributed in the environment have to send back all the gathered data to the central entity, that takes all the burden of parsing the sensitive information and infer the policies to be implemented. Our architecture is instead focused on a physical approach, where every device carries a part of the global intelligence: every single entity can analyze the part of information sensitive for its goal, derive useful data, and communicate meaningful information to the other devices.

Our work is based on a four-layer model [8], as showed in Figure 3. The first layer of our system should be simply composed of sensors, but some constraints have to be fitted. In order to reduce the global system cost and to protect the inhabitants’ privacy, the number of sensors dispatched in the environment has to be reduced as much as possible. However, a huge number of different sensors are required to sense context pieces and redundancy can significantly increase the reliability of the sources. With this idea in mind, the sensors are grouped in nodes. These nodes are able to preprocess the data with simple computation such as minimum, maximum and average. They also enable the sensors to communicate, using, for instance, 6LowPAN (IPv6 over LoW Power wireless Area Networks).
In the second layer, the raw data are processed to obtain more abstract data about context and occurring situations. It could be, for instance, a presence in a room, the number of people in this room or the posture of someone. The aggregation of raw data is realized thanks to a data fusion algorithm. The one we adopted is called the belief functions theory or theory of evidence [6].

The bridge between the second and the third layer is realized integrating the results of sensor data fusion into a context model called Context Spaces. This model uses geometrical metaphors to describe context and situations, relying on the following concepts: the context attributes, the application space, the situation spaces and the context state. The context attributes are information types that are relevant and obtainable by the system; in our case, the context attribute values are provided by the perception layer, together with a degree of confidence on them, needed to cope with the intrinsic uncertainty of sensing systems in real world scenarios. In the situation and context identification layer, the context state provided by the perception layer is analyzed to infer the ongoing situation spaces (representing real-life situations) and also produce a measure of confidence in their occurrence. As the same context state can correspond to several different situation spaces (and vice versa), reasoning techniques are needed to discern the actual ongoing real-life situations in spite of uncertainty [5].

4.3.2. Experimentation

The computations required by the second and the third layers to obtain abstract data and to analyze context and situations are too heavy for our nodes to be processed on. To remedy to this problem, more powerful nodes acting like sinks are used. These nodes are small "plug and play" computers called plug computers. Their role is to gather data from sensor nodes and to perform data fusion, required to produce the context attributes, and context space reasoning, used to identify ongoing situations.

The figure 4 gives an overview of our system architecture. The latter has been demonstrated by ACES team at EDF R&D in November.

Figure 4. System architecture
6. New Results

6.1. Adaptive Middleware


In 2011, we pursued our goal to demonstrate that general and high level concepts and solutions can be proposed to design multi-scale middleware systems. The multi-scale aspect has particularly been put forward and we obtained several interesting results: we showed that the concepts of service, component, and software architecture can be successfully used, in the small for wireless sensor middleware platforms [19], [43] with applications to the Internet of Things (IoT) [27] and for embedded systems [14], in mid-size distributed environments such as digital home networks [17], and in the large in cloud computing platforms [30]. We focus below on two achievements which are illustrative in the sense that they address both ends of the targeted spectrum of sizes.

At the scale of small systems, we proposed the REMORA platform [19], [43], [27] which defines a lightweight event-based programming model for wireless sensor networks. A C-like language for component implementation and an extension of the state-of-the-art Service Component Architecture (SCA) standard for service-oriented systems are proposed. The platform has been successfully deployed on the Contiki operating system. We showed that despite the characteristics of such resource-constrained environments, we are still able to obtain reconfigurability and adaptability properties for the deployed systems.

At the scale of very large systems, we showed in [30] first results that illustrate the fact that the FraSCAti platform [18] can be used to achieve interoperability between applications deployed on heterogeneous cloud platforms. The experiment is currently deployed on 13 public IaaS and PaaS cloud infrastructures. The very same concepts of service and software architecture that are used at smaller scales are put into practice here. Furthermore, we benefit from the same adaptability properties to address the heterogeneity of concepts needed to fit these very large scale infrastructures.

6.2. Context-awareness and Ambient Intelligence Software

Participants: Laurence Duchien, Sébastien Mosser, Clément Quinton.

Context-aware applications are applications that can react to changes on their environment. To achieve such reacting behavior, several challenges have to be faced in terms of: context management, support for dynamic reconfiguration, automation of development, and a consistent development process. One possible way to face those challenges is to use the principles of Software Product Line (SPL) and specifically dynamic SPL (DSPL). DSPLs focus on variability management and aim at deriving different products from a same product family. Additionally, DSPLs allow for products to be derived both at design and at runtime. This enables applications to be adapted during execution and dynamically fit new requirements or resource changes. In [16] we have proposed an approach to unify adaptation at design and at runtime based on Aspect Oriented Modeling. Our approach proposes a unified aspect metamodel and a platform that realizes two different weaving processes to achieve design and runtime adaptations. This approach is used in a Dynamic Software Product Line which derives products that can be configured at design time and adapted at runtime in order to dynamically fit new requirements or resource changes. Such products are implemented using the Service Component Architecture and Java. Finally, we have illustrated the use of our approach based on an adaptive e-shopping scenario. This work corresponds to Carlos Parra’s PhD thesis [11] and is partially funded by the CAPPUCINO project. Finally, in [40], we have proposed to develop an application for mobile devices using Software Product Lines (SPL). Considering variation factors, SPL allows the conception and the development of a software products family minimizing realization cost and time. The result is the APPLIDE framework, which provides SPL for smartphones, and we show how it works with a short demonstration.
6.3. SCeSAME: Formal Definition of Software Architecture Adaptation

Participants: Rubby Casallas, Laurence Duchien, Nohra Villegas, Gabriel Tamura.

In order to define properties on adaptation process, we need to formally model the architecture reconfiguration of a component-based (CB) system as an action performed by itself. These actions are performed in response to the disruption of Quality of Service (QoS) contracts, in the spirit of the Effeil’s rescue clause in object-oriented programming. By doing this, we aim to develop on the vision of the CBSE as a sound base to produce software systems enabled to automatically and safely reconfigure themselves by reconfiguring their abstract (reflection) architectures at runtime. For such structural reconfigurations, a system architect may reuse design patterns from other disciplines with the purpose of restoring QoS contracts, thus preserving them.

Our approach, named SCeSAME for “A Safe Contract-based Self-Adaptive Framework to Preserve QoS Properties on Mobile Devices” is built on the theory of extended graph (e-graph) rewriting proposed in e-graph [60], as a formalism to represent QoS contracts on component. We have given a formal definition of component-based structure systems, QoS contracts, and architecture reconfiguration rules. Based on these definitions, we built a framework that enables a component-based system to preserve its QoS contracts through architecture self-reconfiguration as a responding action to QoS contract violations. Our approach extends a theory of graph rewriting and defines a process calculus as formalisms to model the structure and reconfiguration process of architecture reconfiguration. The reconfiguration process, once parameterized with reconfiguration rules, can be verified as safe, i.e., component structural-compliant, terminating and confluent. This result is a part of Gabriel Tamura’s PhD and the results have been published in [36], [44].
6. New Results

6.1. Structuring of Applications for Scalability

**Participants:** Sylvain Contassot-Vivier, Thomas Jost, Jens Gustedt, Soumeya Leila Hernane, Constantinos Makassikis, Stéphane Vialle.

6.1.1. Large Scale and Interactive Fine Grained Simulations

Our library *parXXL* allows the validation of a wide range of fine grained applications and problems. We were able to test the interactive simulation of PDEs in physics, see [5], on a large scale. Also, biologically inspired neural networks have been investigated using *parXXL* and the InterCell software suite. The InterCell suite and these applicative results have been presented in [29].

6.1.2. Large Scale Models and Algorithms for Random Structures

A realistic generation of graphs is crucial as an input for testing large scale algorithms, theoretical graph algorithms as well as network algorithms, e.g platform generators.

Commonly used techniques for the random generation of graphs have two disadvantages, namely their lack of bias with respect to history of the evolution of the graph, and their incapability to produce families of graphs with non-vanishing prescribed clustering coefficient. In this work we propose a model for the genesis of graphs that tackles these two issues. When translated into random generation procedures it generalizes well-known procedures such as those of Erdős & Rényi and Barabási & Albert. When just seen as composition schemes for graphs they generalize the perfect elimination schemes of chordal graphs. The model iteratively adds so-called *contexts* that introduce an explicit dependency to the previous evolution of the graph. Thereby they reflect a historical bias during this evolution that goes beyond the simple degree constraint of preference edge attachment. Fixing certain simple statical quantities during the genesis leads to families of random graphs with a clustering coefficient that can be bounded away from zero.

A journal article describing intensive simulations of these models that confirm the theoretical results and that show the ability of that approach to model the properties of graphs from application domains has been published as [13].

6.1.3. Development environment for co-processing units

In the framework of the PhD thesis of Wilfried Kirschenmann, co-supervised by Stéphane Vialle (SUPELEC & AlGorille team) and Laurent Plagne (EDF SINETICS team), we have designed and implemented a unified framework based on generic programming to achieve a development environment adapted both to multi-core CPUs, multi-core CPUS with SSE units, and GPUs, for linear algebra applied to neutronic computations.

Our framework is composed of two layers: (1) MTPS is a low-level layer hiding the real parallel architecture used, and (2) Legolas++ is a high-level layer allowing the application developer to rapidly implement linear algebra operations. The Legolas++ layer aims at decreasing the development time, while the MTPS layer aims at automatically generating very optimized code for the target architecture, thus leading to decreased execution times. Experimental performances of the MTPS layer appeared very good, the same source code achieved performances close to 100% of the theoretical ones, on any of the supported target architectures. Our strategy is to generate optimized data storage and data access code for each target architecture, not just different computing codes.

A new version of Legolas++ is under development, and a minimal version has been implemented in 2011. It is optimized to use the MTPS layer: source code is generic while an optimized code is automatically generated to efficiently use all SSE/AVX vector units of a multicore CPU. An article on that work is accepted in the post-proceedings of PARA 2010 and will be published at the end of 2011; the thesis of Wilfried Kirschenmann will be defended in early 2012.
6.1.4. Structuring algorithms for co-processing units

Since 2009, we have designed and experimented several algorithms and applications, in the fields of option pricing for financial computations, generic relaxation methods, and PDE solving applied to a 3D transport model simulating chemical species in shallow waters. We aim at designing a large range of algorithms for GPU cluster architectures, to develop a real knowledge about mixed coarse and fine grained parallel algorithms, and to accumulate practical experience about heterogeneous cluster programming.

Our PDE solver on GPU cluster has been designed in the context of a larger project on the study of asynchronism (see 3.1 and 6.1.5). The iterations of the asynchronous parallel algorithm runs faster, but it requires more iterations and a more complex detection of convergence, see Section 6.1.5 below. We measured both computing and energy performances of our PDE solver in order to track the best solution, in function of the problem size, the cluster size and the features of the cluster nodes. We are tracking the most efficient solution for each configuration. It can be based on a CPU or a GPU computing kernel, and on a synchronous or asynchronous parallel algorithm. Moreover, the fastest solution is not always the less energy consuming. Our recent results are introduced in [26], and in an article accepted in the post-proceedings of PARA 2010. In 2011 we improved our asynchronous implementation. However, the most asynchronous version has led to significantly more complex code (with an increased probability of remaining bugs) but to similar performances. At the opposite, we designed and implemented different convergence detection mechanisms in our asynchronous version, and some versions seem to achieve really better performances. Execution time and energy consumption performances have now to be measured again for many configurations. We aim to get new complete performance evaluation at the beginning of 2012. Then we will design an automatic selection of the right kernel and the right algorithm, and we will implement an auto-setting application function of a global instruction of the user (to achieve a fast run, or a low consumption run, or a compromise...).

At last, we have continued to design option pricers on clusters of GPUs, with Lokman Abbas-Turki (PhD student at University of Marne-la-Valée) and some colleagues from financial computing. In the past we developed some European option pricers, distributing independent Monte-Carlo computations on the nodes of a GPU cluster. In 2010 we succeeded to develop an American Option pricer on our GPU clusters, distributing strongly coupled Monte-Carlo computations. The Monte-Carlo trajectories depend on each others, and lead to many data transfers between CPUs and GPUs, and to many communications between cluster nodes. First results were encouraging, we achieve speedup and size up. In 2011 we optimized a major step of our algorithm, consisting in a 4D to 2D reduction on GPU. Performances have increased, and are significantly easier to achieve. The configuration tuning of the application, function of the problem size and the number of computing nodes, has been simplified. Again, we investigate both computing and energy performances of our developments, in order to compare interests of CPU clusters and GPU clusters considering execution speed and the exploitation cost of our solution.

6.1.5. Asynchronism

In the previous paragraph is mentioned a project including the study of sparse linear solvers on GPU. That project deals with the study of asynchronism in hierarchical and hybrid clusters mentioned in 3.1.

In that context, we study the adaptation of asynchronous iterative algorithms on a cluster of GPUs for solving PDE problems. In our solver, the space is discretized by finite differences and all the derivatives are approximated by Euler equations. The inner computations of our PDE solver consist in solving linear equations (generally sparse). Thus, a linear solver is included in our solver. As this part is the most time consuming, to decrease the overall computation time it is essential to get a version that is as fast as possible. This is why we have decided to implement it on GPU, as discussed in the previous paragraph. Our parallel scheme uses the Multisplitting-Newton which is a more flexible kind of block decomposition. In particular, it allows for asynchronous iterations.

Our first experiments, conducted on an advection-diffusion problem, have shown very interesting results in terms of performances [8]. However, we investigate the possibility to insert periodic synchronous iterations inside the asynchronous scheme in order to improve the convergence detection delay. This is especially interesting on small/middle clusters with efficient networks.
Moreover, another aspect which is worth being studied is the full use of all the computational power present on each node, in particular the multiple cores, in conjunction with the GPU. This is still a work in progress.

6.1.6. New Control and Data Structures for Efficiently Overlapping Computations, Communications and I/O

With the thesis of Pierre-Nicolas Clauss we introduced the framework of ordered read-write locks, ORWL, see [3]. These are characterized by two main features: a strict FIFO policy for access and the attribution of access to lock-handles instead of processes or threads. These two properties allow applications to have a controlled pro-active access to resources and thereby to achieve a high degree of asynchronism between different tasks of the same application. For the case of iterative computations with many parallel tasks which access their resources in a cyclic pattern we provide a generic technique to implement them by means of ORWL. It was shown that the possible execution patterns for such a system correspond to a combinatorial lattice structure and that this lattice is finite iff the configuration contains a potential deadlock. In addition, we provide efficient algorithms: one that allows for a deadlock-free initialization of such a system and another one for the detection of deadlocks in an already initialized system.

We have developed a standalone distributed implementation of the API that is uniquely based on C and POSIX socket communications. Our goal is to simplify the usage of ORWL and to allow portability to a large variety of platforms. This implementation runs on different flavors of Linux and BSD, on different processor types Intel and ARM, and different compilers, gcc, clang, opencc and icc. An experimental evaluation of the performance is on its way. An engineering support from the local INRIA center has allowed to advance this implementation and to perform intensive benchmarks. The results have been presented in [28].

Data Handover, DHO, is a general purpose API that combines locking and mapping of data in a single interface. The access strategies are similar to ORWL, but locks and maps can also be hold only partially for a consecutive range of the data object. It is designed to ease the access to data for client code, by ensuring data consistency and efficiency at the same time.

In the thesis of Soumeya Hernane, we use the Grid Reality And Simulation (GRAS) environment of SimGrid, see 5.4, as a support for an implementation of DHO. GRAS has the advantage of allowing the execution in either the simulator or on a real platform. A first series of tests and benchmarks of that implementation demonstrates the ability of DHO to provide a robust and scalable framework, [18]. A step forward towards a distributed algorithm that allows distributed read-write locks with dynamic participation of process has been achieved in [30].

6.1.7. Energy performance measurement and optimization

Several experiments have been done on the GPU clusters of SUPÉLEC with different kinds of problems ranging from an embarrassingly parallel one to a strongly coupled one, via some intermediate levels. Our first results tend to confirm our first intuition that the GPUs are a good alternative to CPUs for problems which can be formulated in a SIMD or massively multi-threading way. However, when considering not embarrassingly parallel applications the supremacy of a GPU cluster tends to decrease when the number of nodes increases. This observation was the starting point of our participation to the COST-IC0804 about energy efficiency in large scale distributed systems, and an article accepted in the post-proceedings of PARA 2010 introduces our results achieved with our PDE solver distributed on our GPU clusters.

In 2011 we conducted new experiments and optimizations of our PDE solver and our American option pricer on the GPU clusters of SUPELEC. These experiments are still ongoing, and the optimization of this software should be achieved at the beginning of 2012. Simultaneously, we designed the foundations of a complete software architecture of self-configuring applications, choosing the right compute kernel and the right parallel algorithm to use, automatically. The global objectives to respect would either the overall speed, low energy consumption, or a speed-energy compromise. This global objective can be set by the user, or by an intelligent scheduler that aims to optimize a set of runs on a large cluster. This software architecture foundation has been introduced to a COST-IC0804 meeting in Budapest in June 2011.
In order to achieve this goal we need to establish some models of energy consumption of our applications on our CPU+GPU clusters, to be able to implement some heuristic of auto-setting. In 2011 we published a book chapter [26] introducing our first modeling strategies. Next step will be to implement a first auto-setting application, and to experiment its performances.

6.1.8. Load balancing

A load-balancing algorithm based on asynchronous diffusion with bounded delays has been designed to work on dynamical networks [34]. It is by nature iterative and we have provided a proof of its convergence in the context of load conservation. Also, we have given some constraints on the load migration ratios on the nodes in order to ensure the convergence. This work has been extended, especially with a detailed study of the imbalance of the system during the execution of a parallel algorithm simulated in the SimGrid platform.

The perspectives of that work are double. The first one concerns the internal functioning of our algorithm. There is an intrinsic parameter which tunes the load migration ratios and we would like to determine the optimal value of that ratio. The other aspect is on the application side in a real parallel environment. Indeed, we are currently applying this algorithm to a parallel version of the AdaBoost learning algorithm. This will allow us to study the best parameter to choose and to compare our load-balancing scheme to other existing ones.

Concerning the Neurad project, our parallel learning proceeds by decomposing the data-set to be learned. However, using a simple regular decomposition is not sufficient as the obtained sub-domains may have very different learning times. Thus, we have designed a first domain decomposition of the data set yielding subsets of similar learning times [40]. One of the main issues in this work has been the determination of the best estimator of the learning time of a sub-domain. As the learning time of a data set is directly linked to the complexity of the signal, several estimators taking into account that complexity have been tested, among which the entropy. However, the entropy is not the best estimator in that context, and we had to design a specific estimator. Also, we have optimized the decomposition process and added a selection phase that produces learning subsets of the same size [20]. Finally, we have also developed a parallel multi-threaded version of that decomposition/selection process.

6.1.9. Fault Tolerance

6.1.9.1. Application-level fault tolerance

Concerning fault tolerance, we have worked with Marc Sauget, from the University of Franche-Comté, on a parallel and robust algorithm for neural network learning in the context of the Neurad project [35]. A short description of that project is given in Section 4.1.5.

Our fault-tolerance strategy has shown to be rather efficient and robust in our different experiments performed with real data on a local cluster where faults were generated. Although those results are rather satisfying, we would like to investigate yet more reactive mechanisms as well as the insertion of robustness at the server level.

6.1.9.2. Programming model and frameworks for fault-tolerant applications

During the PhD thesis of Constantinos Makassikis [11], supervised by Stéphane Vialle, we have designed a new fault tolerance programming model (MoLoToF) to ease the development of fault-tolerant distributed applications. Main features of MoLoToF include so-called “fault-tolerant skeletons” to embed checkpoint-based fault tolerance within applications, and enable various collaborations, such as application-semantic knowledge supplied by users to the underlying system (e.g.: middleware), in order to fine tune fault tolerance.

Two development frameworks have been designed according to two different parallel programming paradigms: ToMaWork for Master-Workers applications [17] and FT-GReLoSSS (FTG) for some kind of SPMD applications including inter-node communications [10]. The programmer’s task is limited. He only needs to supply some computing routines (functions of the application), has to add some extra code to specify a fault-tolerant parallel programming skeletons and then to tune the checkpointing frequency.
Our experiments have exhibited limited runtime overheads when no failure occurs and acceptable runtime overheads in the worst case failures. Observed runtime overheads are less than the ones obtained with all other system-level fault tolerance solutions we have experimented, while maintaining very limited development time overhead. Moreover, detailed experiments up to 256 nodes of our cluster have shown that it is possible to finely tune the checkpointing policies of the frameworks in order to implement different fault tolerance strategies, for example, according to cluster reliability.

In 2011, we have used the FTG framework to make fault-tolerant an existing parallel financial application [31] from EDF R&D, where it is used for gas storage valuation. The resulting application kept its initial runtime performance despite some source code modifications which are required in order to use FTG. As it was the case in earlier experiments with other applications, these modifications accounted for a limited development time overhead and fault tolerance remained more efficient than system-level fault tolerance solutions.

6.2. Experimentation Methodology

Participants: Tomasz Buchert, Sébastien Badia, Pierre-Nicolas Clauss, El Mehdi Fekari, Jens Gustedt, Lucas Nussbaum, Martin Quinson, Cristian Rosa, Luc Sarzyniec, Sylvain Contassot-Vivier.

6.2.1. Overall Improvements of the SimGrid Framework

See 4.2.1 for the scientific context of this result.

This year was the third year of the USS-SimGrid project on the simulation of distributed applications. We are principal investigator (see 8.2.7) of this project, funded by the ANR. It was prolonged until October 2012, giving us the ability to finish properly what was started. Several improvements have therefore been added to the framework, with numerous contributions from the ANR participants. This served as a flagship for the whole SimGrid project and hosted several of our research efforts, detailed in the subsequent sections (up to 6.2.5). Also this year, the SONGS project got accepted by the ANR, paving the road for our research in this context for the next four years. Our team also coordinates this project, devoted to the “Simulation Of Next Generation Systems” (see 8.2.7).

In addition, the software quality efforts were pursued further through the second year of the INRIA ADT project (see 8.2.1) to maximize the impact of our research on our user community. First, we improved further our automated regression tests (by increasing the test coverage from below 60% to almost 80%) and by fixing the bugs found through the automated builds conducted on the INRIA pipol infrastructure. We also reduced the amount of possible configurations to reduce the test and maintenance burden. As usual, performance tuning deserved a lot of our attention this year. The bindings were solidified and improved. They are very well received by the user community. Finally, the port to the Mac architecture was improved while the experimental port to Windows was revived.

Finally, several operations were conducted to increase our user community. A publication summarizing all improvements made in the recent years were written and submitted [27]. The SimGrid team was represented at SuperComputing’11 (through our partners of Lyon) to meet potential users and distribute informative leaflets designed and printed to that extend.

6.2.2. Formal Verification of Distributed Applications

The context of this work is presented in 4.2.2.

In 2011, we started using the model-checker integrated last year into the SimGrid framework with the goal to evaluate its limitations. Due to its generic design, it is able to verify protocols written using several APIs of SimGrid. We tested it on both a MPI toy program written to that extend and on an implementation of the Chord P2P protocol. In this later case, the tested program was not written for the purpose of being model-checked but to assess the scalability limits of the simulator. The model-checker was used to track down a bug that was near to impossible to find with the simulator alone. This experiment and the formalism underlying our model-checker were described in the publication [19]. It is also described in further detail in Cristian Rosa’s PhD, defended this year [12].
A second axis of our work this year consisted in extending the semantic power of the verified properties. In the work presented above, only local assertions and invariants can be verified. We started to investigate how to improve this during the internship of Marion Guthmuller. The major difficulty is that the reduction techniques based on the transition independences that we used so far are not sufficient for vivacity properties and must be extended to deal with the visibility of atomic properties [37]. One of the specificity of our work is the use of actual implementations were most of the literature uses handmade abstract models. This work continued in a PhD program, but didn’t lead to any publication, yet.

6.2.3. Parallel Simulation within SimGrid

In addition to the software tuning and improvement described in 6.2.1, we tackled the issue of running SimGrid simulators in parallel. Our work differs from the state of the art, because we do not aim to parallelize the simulation kernel itself but the execution of the user code processes running on top of the simulated system. Interestingly enough, this benefits greatly from the work on formal verification introduced in the previous section, and particularly of the new network abstraction layer that was added. It greatly reduced the code locations where the global state is modified, making the parallel execution possible.

This allowed for example the simulation of up to 2 million Chord hosts on a single computer. This work was described in [33] and a publication in a major conference is under preparation. Since the available memory constitutes the main scalability limit now, we will work on distributing the simulation to leverage the memory of several computers at the same time.

6.2.4. Simulating MPI Applications

The final goal of SMPI is to simulate a C/C++/FORTRAN MPI program designed for a multi-processor system on a single computer without any source code modification. This addresses one of the main limitation of SimGrid, which requires the application to be written using one of the specific interfaces atop the simulator. New efforts have been put since July 2009 in this project, hereby continuing the work initiated by Henri Casanova and Mark Stilwell at University of Hawai’i at Manoa.

Previous work included a prototype implementation of various MPI primitives such as send, recv, isend, irecv and wait. Since the project’s revival, many of the collective operations (such as bcast, alltoall, reduce) have been implemented. The standard network model used in SimGrid has also been reworked to reach a higher precision in communication timings. Indeed, MPI programs are traditionally run on high performance computers such as clusters, and this requires to capture fine network details to correctly model the program behavior. Starting from the existing, validated network model of SimGrid, we have derived for SMPI a specific piece-wise linear model which closely fits real measurements. In particular, it enables to correctly model small messages and messages above the eager/rendezvous protocol limit. This work has been published at the IPDPS conference this year [15].

Ongoing work is now targeting a panel of MPI applications to have a better understanding of the applicability of our proposition. Pierre-Nicolas Clauss, who has been working full-time on the project between mid-2010 and mid-2011 has left, and we plan to put new workforce on SMPI with the support of the SONGS ANR project in 2012.

6.2.5. Simulating Real Applications

This work aims at providing a solution to simulate arbitrary applications on top of SimGrid. The approach consists in intercepting the application actions at system level while they are executed on a test platform, and then replay these actions on top of the simulator.

Concerning trace capture, we continued our work on the Simterpose software, which intercepts the actions of the application and save them to file for further use by the simulator. This work, presented in a national conference [22], will be continued during the PhD work of Marion Guthmuller.
Concerning trace replay, we proposed a replay mechanism specific to MPI applications in collaboration with F. Suter from the Computing Center at IN2P3 together with F. Desprez and G. Markomanolis from the Graal team at INRIA Rhônes-Alpes. The originality is to rely on time-independent execution traces. This allows to completely decouple the acquisition process from the actual replay of the traces in a simulation context. We are able to acquire traces for large application instances without being limited to an execution on a single cluster. Finally, our replay framework is built directly on top of the SIMGrid simulation kernel. This work was published in [16].

6.2.6. Emulation & Distem

During the internship of Luc Sarzyniec, we re-implemented an emulator from scratch with the goal of having a more reliable basis for further developments. This new development, Distem (see 5.2), already includes support for CPU performance emulation (internship of Tomasz Buchert in 2010) and network emulation. We are currently preparing a first release of Distem, and are working on its validation.

6.2.7. Grid’5000 and ADT Aladdin-G5K

Grid’5000 is an experimental platform for research on distributed systems. Two new sites were added to Grid’5000 in 2011: Reims and Luxembourg. This should reinforce the impact of Grid’5000 in the east of France. It is worth noting that the system administrator of the Luxembourg Grid’5000 site was formerly a student in Nancy, and did a student project using Grid’5000 managed by Lucas Nussbaum. Also, more collaboration on technical aspects is expected thanks to this geographical proximity.

On the local level, power consumption sensors are being added to the graphene cluster, which will allow an accurate monitoring of energy consumption during experiments.

On the national level, Lucas Nussbaum is now mandated by the Grid’5000 executive committee to follow the work of the technical team. He contributed to two publications [23], [24] at Journées Réseaux 2011 that describe the Grid’5000 software stack. He also gave invited talks during a Grid’5000 day at RenPar, and during the Support for experimental computer science workshop as SuperComputing’11.

Local scientific contributions include the automation of the deployment of the gLite middleware on Grid’5000. That work [21] was presented at Rencontres France Grilles and received the Best Poster award. We hope that this work will serve as a basis for further collaborations with the production grids community.

We also started the ADT Kadeploy project that will continue the development of the Kadeploy software, which already plays a key role on Grid’5000.

6.2.8. Experimental cluster of GPUs

The experimental platform of SUPÉLEC for "GPGPU", see Section 4.2.6, is composed of two GPU clusters, and its electrical line has been improved in 2011.

The first cluster is currently composed of 16 PCs, each one hosting a dual-core CPU and a GPU card: an nVIDIA GeForce GT285, with 1GB of RAM (on the GPU card). The 16 nodes are interconnected across a devoted Gigabit Ethernet switch. The second cluster has 16 more recent nodes, composed of an Intel Nehalem CPU with 4 hyper-threaded cores at 2.67GHz, and an nVIDIA GTX480 ("Fermi") GPU card with 1.5GB of memory. This cluster has a Gigabit Ethernet interconnection network too. These 2 clusters can be accessed and used like one 32-nodes heterogeneous cluster of hybrid nodes. This platform has allowed us to experiment different algorithms on an heterogeneous cluster of GPUs.

The energy consumption of each node of the cluster hosting the GTX285 GPUs is monitored by a Raritan DPXS20A-16 device that continuously measures the electric power consumption (in Watts). The nodes of the cluster hosting the GTX480 are monitored by two Raritan devices, because the energy consumed by this cluster exceeds the maximum energy supported by a Raritan DPXS20A-16 device.

A set of Perl and shell scripts, developed by our team, sample the electrical power (Watt) measured by the Raritan devices and compute the energy (Joule or Watt Hour) consumed by the computation on each node and on the complete cluster (including the interconnection switch).
In 2011 we have increased the amount of electrical energy supplied to these cluster, in order to support the experiments of our new distributed American option pricer. This application achieves high performances but consumes more energy on our GPU clusters than our previous codes, and exceeded the limit of our previous electrical line.

This platform has been intensively used to get experimental performance measures introduced in 2011 meetings of the COST IC0804 about *Energy efficiency in large scale distributed systems*, and published in a book chapter [26].
6. New Results

6.1. SouthBound results

6.1.1. L4 micro kernels

As part of our investigations about what software architectures were the best candidates to base our Ambient Middleware Stack upon, we studied different micro-kernel operating systems such as CodeZero [33], OKL4, and L4/Fiasco. The objective here is to try and quantify the development effort that would be needed before being able to execute a Java application on top of a micro-kernel. These studies included, in addition to a lot of bibliographic research, several technical experiments such as booting each of these various micro-kernel systems in QEmu, as well as on real hardware. We use a BeagleBoard as a representative example of the kind of hardware platforms typically encountered in Ambient Intelligence scenarios.

6.1.2. Virtual machines

In parallel to our study of micro-kernel architectures, we worked on virtual machines as well, in the perspective of bridging the gap between the two. The basic question here is: what does it takes to to cut down a Java virtual machine into pieces so as to run each of these pieces as a separate software component in the system. We ran two actions in order to investigate this question. First, we ported the JamVM virtual machine to run on top of the Genode operating system framework [37]. This provided us with better understanding of what are the real requirements of a Java virtual machine in terms of underlying operating systems support. Second, we focused on one particular service of the virtual machine, the garbage collector, and we precisely identified and studied the coupling between this component and all other parts of the virtual machine (bytecode interpreter, scheduler, etc). This work was done as part of a student summer internship (Yann Chevalier, INSA-Lyon 3IF). Removing a garbage collector at runtime, and “plugging in” another one dynamically proved to be vastly harder than expected. Still, this work provided us with great insights about the coupling relationships between different OS components.

6.1.3. HiKoB

Antoine Fraboulet (Amazone team), Guillaume Chelius (D-NET team) and Christophe Braillon (INRIA SED) started a new company called HiKob http://openlab.hikob.com/ in July 2011. HiKoB is a development project following several successful research projects completed these last 6 years at INSA Lyon and INRIA. HiKoB hardware and software products help in building complex, large-scale and distributed applications in the domains of: motion capture, biomechanical study, biologging study, building instrumentation and many more applications targeting wireless sensor network solutions for distributed and embedded measurement. HiKoB business model is built on two major directions: complete solutions for industrial applications and software and hardware tools for research and innovation in the fields of sensor networking and embedded wireless measure. HiKoB is supported by IT-Translation and INSArval.

6.1.4. Service-Oriented Tainted Object Propagation

Many Java technologies allow the execution of code provided by multiple parties. Service-oriented platforms based on components such as OSGi are good examples of such a scenario. Those extensible component-based platforms are service-oriented, as components may directly interact with each other via the services they provide. However, even robust languages such as Java were not designed to handle safely code interaction between trusted and untrusted parties.

In [38], we show how basic Java interactions can break encapsulation or execution safety and why the Java security layers’ threat coverage is incomplete. We also review flaws in the Java access control design that can allow untrusted code to bypass restrictions by exploiting vulnerabilities in trusted code.
As component-based platforms become more and more integrated to our daily life, we improved our Service-Oriented Tainted Object Propagation technique to find such vulnerabilities and used it on several open-source components to further demonstrate the real exposure that those vulnerabilities bring to the fore.

6.2. NorthBound results

Another key issue in the Amazones architecture was to bring together formal methods and service oriented programming such as OSGi/Java approach. We developed the Logos framework that observes and records communications that occurs between an OSGi client and a corresponding server. This architecture is developed in the LISE ANR project, and we made various improvements to the architecture.

6.2.1. Amazones Protocol

It aims at building automata from running and observing applications. The logos framework observes a running application at builds at run-time an automata that represents the application behavior. Julien Ponge wrote the corresponding Scala code.

6.2.2. Monitored oriented programming

Another Logos extension integrates Monitored Oriented architectures such as JavaMOP and Larva. We are currently using and working with the larva people. Each time we intercept a call, it is transfered to the Larva automata manager.

6.2.3. Real time SOA

Admission control for service oriented application in real time infrastructure. This work led by Lionel Morel tries to bring together real time architecture configuration and component based architectures. The deal is to find a better way of managing the dynamicity of applications in real time context.

6.2.4. B Model Slicing to Generate Tests

In a model-based testing approach as well as for the verification of properties, B models provide an interesting modelling solution. However, for industrial applications, the size of their state space often makes them hard to handle. To reduce the amount of states, an abstraction function can be used. The abstraction is often a domain abstraction of the state variables that requires many proof obligations to be discharged, which can be very time consuming for real applications.

we propose a contribution to this problem that complements an approach based on domain abstraction for test generation, by adding a preliminary syntactic abstraction phase, based on variable elimination. We define a syntactic transformation that suppresses some variables from a B event model, in addition to three methods that choose relevant variables according to a test purpose. This way, we propose a method that computes an abstraction of a source model $M$ according to a set of selected relevant variables. Depending on the method used, the abstraction can be computed as a simulation or as a bi-simulation of $M$. With this approach, the abstraction process produces a finite state system. We apply this abstraction computation to a Model Based Testing process. We evaluate experimentally the impact of the model simplification by variables elimination on the size of the models, on the number of proof obligations to discharge, on the precision of the abstraction and on the coverage achieved by the test generation.

This work is based on a B model approach. However, in the context of AMAZONES, one of our objectives is to extend it, in order to consider models automatically generated from the usage a the tested service on a particular context.

6.2.5. Distributed Data Centric Programs Verification

Netlog is a language designed to describe distributed programs. It has a precise semantics, provides a high-level of abstraction thanks to its datalog flavor and benefits from an efficient implementation. This makes it a very interesting target language for proofs of distributed programs. In [34], with the Coq proof assistant, we formalized the distributed computation model based on message passing with either synchronous or
asynchronous behaviors; built the translation of Netlog programs; modeled the embedded machine evaluating Netlog programs, and thus established a framework to formally verify properties of distributed programs in Netlog. To test the framework, we proved the correctness of a concrete distributed program for constructing spanning trees over connected graphs.

6.2.6. Managing dynamic service substitution at runtime

The service oriented approach is a paradigm allowing the introduction of dynamicity in developments. If there are many advantages with this approach, there are also some new problems associated to service disappearance. The particular case of service substitution is often studied and many propositions exist. However, proposed solutions are mainly server-side in the context of web-services.

In this work, we propose a client side API-based approach to allow service substitution without any restart of the client and without any assumption on external services. Our proposition is based on a transactional approach, defined to authorized substitutions of services dynamically, by preserving the current run and collected data.

We designed a framework organized by Julien Ponge [14].

6.3. Application domain results

An emerging trend into Amazones team is to apply our northBound/southBound approach to the InternetOfThing wave. We try to apply our architectures to the IoT application domain.

6.3.1. Data Centric Applications Distribution

Peer to peer systems have been widely used to alleviate the burden of servers by transferring to peers in a network tasks that do not require a centralization of the information. A wide range of applications are now emerging over peer-to-peer, such as social networking, multiplayer games, mobile messaging, etc. Most of these applications are essentially data-centric, they rely on exchange of data between peers, and can be expressed by queries over the database.

We propose a tool that allows for such applications, programmed as a collection of queries over a database, to be ported seamlessly without changing the initial queries from a client/server system to peer-to-peer system. The distribution is done with overlays network defined by declarative data centric programs specified in the Netlog language, thus resulting in a fully data centric modeling of the peer-to-peer application. The communication between peers relies on implicit addresses which can be evaluated on the fly to ensure the persistence of data.

We demonstrate the technique on a multiplayer online game, written in SQL, with players who connect to a mobile ad hoc network through their portable devices. The overlay is defined by a combination of an ad hoc routing protocol, DSDV, together with a DHT. The application runs on the QuestMonitor system, which allows to monitor the communication between peers, the evolution of the local data stores, as well as the execution of the declarative code.

6.3.2. Service Deployment in Disrupted Networks

OLD / REMOVE ? Ambient environments classically use wireless connections that suffers from frequent disconnections. The hard research point is to ensure service continuity. This disconnection problem has been widely tackled for application data with proxy and prefetching approaches. For services, disconnections are more difficult to anticipate, since service calls are only solved at run-time.

We are currently working on service deployment and invocations in disrupted networks with a network coding approach. This research is a joint work with the Swing team and with Aline Viana (INRIA Saclay @ TU Berlin). The main idea is to study how social-oriented applications, that need inter-dependent services and updates to be distributed to all or part of the mobile users community, could benefit from a network coding approach. The project aims at assessing for the first time the performance, in terms of latency, energy efficiency and capacity, of standard network coding techniques in presence of realistic user mobility and service demands.
Building on these results, we plan to propose original social-aware network coding techniques that take advantage of the heterogeneous nature of the opportunistic network to reduce delays and energy consumption, in presence of multiple concurrent service flows targeting either all users or specific groups of interests.

These performance issues tackle at the same time the overall network capacity optimizations, as well as the overall software stack optimizations of a device with local and autonomous network coding strategies.

An INRIA ARC project proposal, entitled SoCool, has been submitted jointly with INRIA AMAZONES, INRIA SWING, INRIA MAESTRO, INRIA Saclay, University of Nice, TU Berlin and Fordham University.
6. New Results

6.1. Introduction

The ARLES project-team investigates solutions in the forms of languages, methods, tools and supporting middleware to assist the development of distributed software systems, with a special emphasis on mobile distributed systems enabling the ambient intelligence/pervasive computing vision. Our research activities in 2011 have focused on the following areas:

- Dynamic interoperability among networked systems toward making them eternal, by way of on-the-fly generation of connectors based on adequate system models (§ 6.2);
- Pervasive service-oriented software engineering, focusing on supporting service composition in an increasingly heterogeneous and dynamic networking environment, while enforcing quality of service (§ 6.3);
- Service oriented middleware for the ultra large scale future Internet of Things (§ 6.4);
- Abstractions for enabling domain experts to easily compose applications on the Internet of Things (§ 6.5); and
- System-level support for application development in the context of mobile social ecosystems, while taking into account privacy, performance, and data interoperability (§ 6.6).

6.2. Emergent Middleware Supporting Interoperability in Extreme Distributed Systems

**Participants:** Emil Andriescu, Nelly Bencomo, Amel Bennaceur, Luca Cavallaro, Nikolaos Georgantas, Sneha-Sham Godbole, Valérie Issarny, Rachid Saadi, Daniel Sykes.

Interoperability is a fundamental challenge for today’s extreme distributed systems. Indeed, the high-level of heterogeneity in both the application layer and the underlying infrastructure, together with the conflicting assumptions that each system makes about its execution environment hinder the successful interoperation of independently developed systems. A wide range of approaches have been proposed to address the interoperability challenge [31]. Solutions that require performing changes to the systems are usually not feasible since the systems to be integrated may be legacy systems, COTS (Commercial Off-The-Shelf) components or built by third parties; neither are the approaches that prune the behavior leading to mismatches since they also restrict the systems’ functionality. Therefore, many solutions that aggregate the disparate systems in a non-intrusive way have been proposed. These solutions use intermediary software entities, called *mediators*, to interconnect systems despite disparities in their data and/or interaction models by performing the necessary coordination and translations while keeping them loosely-coupled. However, creating mediators requires a substantial development effort and a thorough knowledge of the application-domain, which is best understood by domain experts. Moreover, the increasing complexity of today’s distributed systems, sometimes referred to as Systems of Systems, makes it almost impossible to develop ‘correct’ mediators manually. Therefore, formal approaches are used to synthesize mediators automatically.
In light of the above, we have introduced the notion of emergent middleware for realizing mediators. Our research on enabling emergent mediators is done in collaboration with our partners of the CONNECT project (§ 7.1.1). Our work during the year has more specifically focused on:

- **Supporting architecture.** We have been working together with our partners in the CONNECT project on the refinement of an overall architecture supporting emergent middleware, from the discovery of networked systems to the learning of their respective behavior, and synthesis of emergent middleware enabling them to interoperate [30].

- **Affordance inference.** We have proposed an ontology-based formal model of networked systems based on their affordances, interfaces, behavior, and non-functional properties, each of which describes a different facet of the system [2]. However, legacy systems do not necessarily specify all of the aforementioned facets. Therefore, we are currently exploring techniques to infer the affordance by using textual descriptions of the interface of networked systems. More specifically, we rely on machine learning techniques to automate the inference of the affordance from the interface description by classifying the natural-language text according to a predefined ontology of affordances [17].

- **Mediator synthesis for emergent connectors.** We focus on systems that have compatible functionality, i.e., semantically matching affordances, but are unable to interact successfully due to mismatching interfaces or behaviors. We propose two approaches to enable communication between such systems:
  1. A mapping based approach, whose goal is to automatically synthesize a mediator model that ensures their safe interaction, i.e., deadlock-freedom and the absence of unspecified receptions. Our approach combines semantic reasoning and constraint programming to identify the semantic correspondence between networked systems’ interfaces, i.e., interface mapping. Unlike existing approaches that only tackle the one-to-one correspondence between actions, this approach handles the more general cases of one-to-many and many-to-many mappings.
  2. A goal based approach, which enables the communication of two networked systems, so that the communication satisfies a given user goal. It aligns their actions using ontology matching. The aligned processes as well as the user goal are encoded as a satisfiability problem. It relies on model checking to determine if a feasible communication trace exists that satisfies the user goal. The model checking process is reiterated so as to discover all the feasible satisfying traces, which are finally concatenated to build the mediator.

The feasibility of both of our approaches has been demonstrated through prototype tools and real-world scenarios involving heterogeneous systems.

- **Mediator synthesis for streaming connectors.** In the context of dynamic mediator synthesis, we have targeted the domain of mobile multimedia streaming, resulting in a first step that statically solves the hard problem of streaming interoperability across heterogeneous smartphone multimedia platforms. With the recent evolution of mobile phones, multimedia streaming is now commonly used in smartphones for purposes such as video broadcast, video conferencing and place shifting, which in turn highlights the importance of multimedia enabled applications. However, peer-to-peer solutions are difficult to implement because of increased node heterogeneity and their low processing power. Furthermore, existing mobile platforms such as Android, iOS, Blackberry and Windows Phone 7 support multimedia streaming (as resource consumers) either through platform specific APIs or system services. However, they use heterogeneous protocols and data formats, thus compromising interoperability.

Given the challenges above, we designed AmbiStream [11], a lightweight middleware for heterogeneous mobile devices, capable of “on the fly” adaptation. AmbiStream relies on the highly-optimized multimedia software stacks provided by smartphone platforms and adds the necessary
layers to solve interoperability. More specifically, the middleware targets: a) Streaming of prerecorded or live audio/video using an intermediary real-time protocol; b) Managing streaming protocol translation and multimedia container format adaptation to the ones supported natively by each device; and c) Extensibility in order to support new multimedia streaming protocols and multimedia container formats given its plug-in based architecture. We have used a model-driven approach to generate multi-platform plug-ins from higher level descriptions in the form of a Domain Specific Language (DSL). The defined DSL takes into account multimedia specific operations such as timing, fragmenting, multiplexing, congestion control and buffering.

- **Models@run.time.** We have recently integrated the notion of Models@run.time in our research towards emergent middleware. We use Models@run.time to extend the applicability of models and abstractions to the runtime environment. As is the case for software development models, a run-time model is often created to support reasoning. However, in contrast to development models, run-time models are used to reason about the operating environment and runtime behavior, and thus these models must capture abstractions of runtime phenomena. Different dimensions need to be balanced, including resource-efficiency (time, memory, energy), context-dependency (time, location, platform), as well as personalization (quality-of-service specifications, profiles). The hypothesis is that because Models@run.time provide meta-information for these dimensions during execution, run-time decisions can be facilitated and better automated. Thus, we anticipate that Models@run.time will play an integral role in the management of extremely distributed systems. Our work on the use of Models@run.time has two aspects:

  - We have used Models@run.time to tackle the crucial problem of uncertainty in extremely distributed systems that are aware of their own requirements. Requirements awareness helps optimize requirements satisfaction when factors that were uncertain at design time are resolved at runtime. Using our approach, we are able to maintain goal-based models in memory while the system is running. The executing system, therefore, is able to introspect and consult it goals during runtime. Crucially, at runtime we use the notion of claims to represent assumptions that cannot be verified with confidence at design time. Such claims are attached to the goal-based runtime models. By monitoring claims at runtime, their veracity can be tested. If falsified, the effect of claim negation can be propagated to the system’s goal model and an alternative means of goal realization can be selected automatically, allowing the dynamic adaptation of the system to the prevailing environmental context.

  - In a complementary way to the mediator synthesis approaches discussed above, we further promote the use of Models@run.time to support the runtime synthesis of software that will be part of the executing system. Specifically, we focus on the use of runtime models to support the realization of emergent middleware, i.e., the synthesis of mediators that define a sequences of actions to translate semantic actions of one system developed using a particular middleware protocol to the semantic actions of another system developed using an alternate middleware built with no prior knowledge on the former. Discovery and learning enablers capture the required knowledge of the context and environment during runtime. Supported by that knowledge, a runtime model of the mediator-to-be is reified. Reification means that the knowledge is explicitly formulated and made available for computational manipulation. The form of the runtime models is based on labeled transition systems (LTSs) which offer the behavioral semantics needed to model the interaction protocols. Ontologies complement the LTSs providing semantic reasoning about the mapping between protocols. Specifically the LTS of each protocol is annotated using ontologies to support the subsequent mapping between the protocols. From the LTS-based runtime models, mediators are synthesized.

---

11Models@run.time Dagstuhl Seminar, [http://www.dagstuhl.de/en/program/calendar/semhp/?semnr=11481](http://www.dagstuhl.de/en/program/calendar/semhp/?semnr=11481)
6.3. Revisiting the Abstractions of Service Oriented Computing for the Future Internet

Participants: Mohammad Ashiqur Rahaman, Dionysis Athanasopoulos, Sandrine Beauche, Nebil Ben Mabrouk, Nikolaos Georgantas, Valérie Issarny.

A software architecture style characterizes, via a set of abstractions, the types of: components (i.e., units of computation or data stores), connectors (i.e., interaction protocols) and possibly configurations (i.e., system structures) that serve to build a given class of systems. As such, the definition of a software architectural style is central toward eliciting appropriate design, development and runtime support for any family of systems. The service oriented architecture style may then be briefly defined as follows: (1) components map to services, which may be refined into consumer, producer or prosumer services; (2) connectors map to traditional client-service interaction protocols; and (3) configurations map to compositions of services through (service-oriented) connectors, e.g., choreography and orchestration structures. While the service-oriented architecture style is well suited to support the development of Internet-based distributed systems, it is largely challenged by the Future Internet that poses new demands in terms of sustaining qualities such as scalability, heterogeneity, mobility, awareness & adaptability that come in extreme degrees compared to the current Internet. Therefore, we have been working on eliciting software architectural abstractions for the Future Internet by building upon the service-oriented architecture style, as well as on applying them to system design, development and execution.

Complex distributed applications in the Future Internet will be to a large extent based on the open integration of extremely heterogeneous systems, such as lightweight embedded systems (e.g., sensors, actuators and networks of them), mobile systems (e.g., smartphone applications), and resource-rich IT systems (e.g., systems hosted on enterprise servers and Cloud infrastructures). These heterogeneous system domains differ significantly in terms of interaction paradigms, communication protocols, and data representation models, provided by supporting middleware platforms. Specifically considering interaction paradigms, the client/server (CS), publish/subscribe (PS), and tuple space (TS) paradigms are among the most widely employed ones today, with numerous related middleware platforms. In light of the above, we have aimed at eliciting abstractions that (i) leverage the diversity of interaction paradigms associated with today’s and future complex distributed systems, as well as (ii) enable cross-paradigm interaction to sustain interoperability in the highly heterogeneous Future Internet [19].

Existing cross-domain interoperability efforts are based on bridging communication protocols, wrapping systems behind standard technology interfaces, and/or providing common API abstractions. In particular, such techniques have been applied by the two widely established system integration paradigms, that is, service oriented architecture (SOA) and enterprise service bus (ESB). However, state of the art interoperability efforts do not or only poorly address interaction paradigm interoperability. Indeed, systems integrated via SOA and ESB solutions have their interaction semantics transformed to the CS paradigm. Then, potential loss of interaction semantics can result in suboptimal or even problematic system integration. To overcome the limitation of today’s ESB-based connectors for cross-domain interoperability in the Future Internet, we introduce a new connector type, called GA connector, which stands for “Generic Application connector”. The proposed connector type is based on the service bus paradigm in that it achieves bridging across heterogeneous connector types. However, the behavior of the GA connector type differs from that of classical ESB connectors by bridging protocols across heterogeneous paradigms, which is further realized by paying special attention to the preservation of the semantics of the composed protocols. Indeed, the GA connector type is based on the abstraction and semantic-preserving merging of the common high-level semantics of base interaction paradigms.

Eliciting Interaction Paradigm Abstractions: We introduce a systematic abstraction of interaction paradigms with the following features:

- First, we introduce base CS, PS and TS connector types, which formally characterize today’s core interaction paradigms. The proposed types comprehensively cover the essential semantics of the considered paradigms, based on a thorough survey of the related literature and representative
Then, we further abstract these connector types into a single higher-level one, the GA connector type. GA is a comprehensive connector type based on the abstract union of CS, PS, and TS, where precise identification of the commonalities or similarities between the latter has enabled the optimization of the former. Further, GA preserves by construction the semantics of CS, PS, and TS.

In more detail, connector types are formally specified in terms of: (i) their API (Application Programming Interface), and (ii) their roles, i.e., the semantics of interaction of the connected component(s) with the environment via the connector. Regarding the latter, the behavioral specification of roles from a middleware perspective relates to specifying the production and consumption of information in the network, while the semantics of the information are abstracted and dealt with at the application layer. The behaviors of the connector roles are then specified using Labeled Transition Systems (LTS). We precisely define the mapping of the roles implemented by the base connector types to/from the corresponding roles of the GA connector type.

For both the above abstraction transformations, we provide counterpart concretizations, which enable transforming GA connector primitives to CS, PS, or TS connector primitives and then to concrete middleware platforms primitives.

Furthermore, based on the GA abstraction, we introduce mapping transformations between any pair from the set \{CS, PS, TS\} via GA. The fine knowledge of CS, PS, and TS semantics, as embedded in GA, enables these mappings to be precise: differing semantics are mapped to each other in such a way that loss of semantics is limited to the minimum. These mappings relate to the definition of the glue process implemented by the GA connector, which defines how a pair of producer and consumer roles coordinates in the environment. The GA glue reconciles consumer and producer roles that may differ with respect to time and space coupling as well as scoping. Hence, GA connectors support interactions among highly heterogeneous services of the Future Internet, and especially across domains.

**eXtensible Service Bus:** We apply the above connector abstractions to introduce an enhanced bus paradigm, the eXtensible Service Bus (XSB). XSB features richer interaction semantics than common ESB implementations to deal effectively with the increased Future Internet heterogeneity. Moreover, from its very conception, XSB incorporates special consideration for the cross-integration of heterogeneous interaction paradigms. When mapping between such paradigms, special attention is paid to the preservation of interaction semantics. XSB has the following features:

- XSB is an abstract bus that prescribes only the high-level semantics of the common bus protocol. The XSB common bus protocol features GA semantics.
- Heterogeneous systems can be plugged into the XSB by employing binding components that adapt between the native middleware of the deployed system and the common bus protocol. This adaptation is based on the systematic abstractions and mappings discussed above.
- XSB, being an abstract bus, can have different implementations. This means that it needs to be complemented with a substrate which at least supports: (1) deployment (i.e., plugging) of various systems on the bus, and (2) a common bus protocol implementing GA semantics. With respect to the latter, we envision that a GA protocol realization may either be designed and built from scratch (still supposing at least an IP-based transport substrate) or be implemented by conveying GA semantics on top of an existing higher-level protocol used as transport carrier. The latter solution can be attractive, as it facilitates GA protocol realizations in different contexts and domains.

We have carried out an early realization of XSB on the PEtALS ESB. In particular, we addressed the workflow-based orchestration of heterogeneous systems, which is a preliminary step before dealing with peer-wise system integration. This work already provides a successful feasibility study of the XSB concept. This work comprises: (i) extending the BPEL workflow language with GA API primitives; (ii) introducing transformation between the GA-extended BPEL and the standard BPEL, which consists in encapsulating GA primitives into standard BPEL primitives and enables conveying GA semantics on top of BPEL primitives.
and subsequently on top of the common bus protocol primitives; (iii) providing templates for systematic and highly facilitated building of binding components; and (iv) introducing transformations between native system interface descriptions and GA-based interface descriptions.

6.4. Service Oriented Middleware facing the Challenges of the Internet of Things

Participants: Benjamin Billet, Nikolaos Georgantas, Sara Hachem, Valérie Issarny, Roberto Speicys Cardoso, Thiago Valladares Sabino Teixeira.

Over the years, the Internet has become the most important networking infrastructure, providing an integrated entity enabling sharing, contributing, creating, using, collaborating and integrating information and knowledge by all. As a result, the Internet is changing at fast pace and is expected to evolve into the Future Internet, i.e., service-aware and self-aware federated networks that provide built-in and integrated capabilities such as: contextualization, reliability, robustness, mobility, security, service support, and self-management of communication resources and services. In our vision, The Future Internet can be defined as the union and cooperation of the Internet of Content, Internet of Services, Internet of Things, and 3D interactive Internet, supported by an expanding network infrastructure foundation. In ARLES, we chose to pay special attention to the Internet of Things (IoT). IoT is characterized by the integration of large numbers of real-world objects (or “things”) onto the Internet, with the aim of turning high-level interactions with the physical world into a matter as simple as is interacting with the virtual world today. As such, two devices that will play a key role in the IoT are sensors and actuators. In fact, such devices are already seeing widespread adoption in the highly localized systems within our cars, mobile phones, laptops, home appliances, etc. In their current incarnation, however, sensors and actuators are used for little more than low-level inferences and basic services. This is partly due to their highly specialized domains (signal processing, estimation theory, robotics, etc.), which demand application programmers to also be domain experts, and partly due to a glaring lack of interconnectivity between the different devices. Our work within that domain was focused on three related directions:

- **Challenges related to IoT**: To prepare the ground for our research on middleware for the Internet of Things, we identified the set of challenges in the IoT, namely [10]: the large scale of the Internet of Things, heterogeneity of things, unknown and dynamic network topology, unknown data-point availability, incomplete or inaccurate metadata, and conflict resolution. The scale issue arises with the millions of devices, millions of users, large amounts of data to share and services to request. The heterogeneity of the IoT is due to the fact that the network will be composed of different types of devices from different vendors with varying sensing/actuating characteristics. The unknown dynamic network topology results from the fact that devices will be mostly mobile and their availability is unknown. A related challenge is the unknown data-point availability as things, which provide the desired measurements, may leave the network or malfunction at any time. A data point is measurement of an entity of interest at a specific time. As for metadata inaccuracy, this issue is a direct result of humans, who are prone to making errors, being the main source of metadata specification. Last but not least, conflict resolution is due to the multiple stakeholders involved in the Internet of Things.

- **Middleware Requirements for the Internet of Things**: The middleware we plan on implementing should abstract things (IoT devices) as services and support dynamic service composition. To handle the IoT challenges, the middleware should also support a probabilistic discovery approach where only a subset, instead of a whole set, of devices is selected in a way that provides a good enough answer that satisfies an application’s request [10]. However, and prior to designing the middleware architecture, we extensively surveyed the literature in order to identify research challenges for service-oriented middleware design, therefore investigating service description, discovery, access and composition in the Future Internet of services [7].

- **Ontologies for the Internet of Things**: As part of our middleware architecture, we specified a set of ontologies [20] that model real-world entities as physical concepts, along with things that measure
those entities. Further, to support a smarter service composition, we also modeled mathematical formulas and physics relations as services to substitute missing thing-based services. Those services will instead compute the value of a desired measurement of an entity of interest. Finally, we also specified an ontology that describes estimation models that can be used to estimate the value of a measurement in case of a missing data point or a missing data source. Estimation models can further be used to define probabilistic discovery functions that will be executed by the middleware.

6.5. Composing Applications in the Internet of Things

Participants: Iraklis Leontiadis, Pankesh Patel, Animesh Pathak.

As introduced above, the Internet of Things (IoT) integrates the physical world with the existing Internet, and is rapidly gaining popularity, thanks to the increased adoption of smart phones and sensing devices. Several IoT applications have been reported in recent research, and we expect to see increased adoption of IoT concepts in the fields of personal health, inventory management, and domestic energy usage monitoring, among others.

An important challenge to be addressed in the domain of IoT is to enable domain experts (health-care professionals, architects, city planners, etc.) to develop applications in their fields rapidly, with minimal support from skilled computer science professionals. Similar challenges have already been addressed in the closely related fields of Wireless Sensor and Actuator Networks (WSANs) and Pervasive/Ubiquitous computing. While the main challenge in the former is the extremely large scale of the systems (hundreds to thousands of largely similar nodes, sensing and acting on the environment), the primary concern in the latter has been the heterogeneity of nodes and the major role that the user’s own interaction with these nodes plays in these systems (cf. the classic “smart home” scenario where the user interacts with a smart display which works together with his refrigerator and toaster). The upcoming field of IoT includes both WSANs as well as smart appliances, in addition to the elements of the “traditional” Internet such as Web and database servers, exposing their functionalities as Web services etc. Consequently, an ideal application development abstraction of the IoT will allow (domain expert) developers to intuitively specify the rich interactions between the extremely large number of disparate devices in the future Internet of Things.

The larger goal of our research is to propose a suitable application development framework which addresses the challenges introduced above. This will most likely be achieved by a domain specific language (DSL) that exposes specific functionalities to the domain experts. The first logical step was to construct a domain model. Towards that end, we took advantage of the CRC — Classes, Responsibility, Collaboration — technique, defining the main abstract concepts, their responsibilities, and associations that represent their relationship with each other in the IoT. Specifically, we used this technique to propose a domain model [22] that addresses the following challenges:

- **Creation of common understanding.** The different terms used by different people in the IoT domain can lead to confusion, which can be alleviated by the usage of a common lexicon, as provided by a domain model. This lexicon can then be used by researchers, system programmers, as well as domain experts.

- **Modeling invariant properties.** The domain model represents the invariant properties of the domain — concepts and relationships which do not change from one application to the other. An instance of this in the IoT domain can be the notion of a sensor attached to a device. Depending on the specific applications, the type of sensors and devices can change (e.g., a light sensor attached to a smart phone), but the inherent relationship between the types of entities they represent does not.

- **Enabling modular design.** Application needs often tend to arrive in terms of behavior, which needs to be broken down and divided among the entities in the system. A good domain model aids in this process, since the capabilities of each type of entity are clearly identified. E.g., the application requirement of “the system senses the temperature of a room and keeps it steady” can be easily broken down into an application consisting of temperature sensors, computational components, and HVAC actuators, each performing its well-known role in this sense-compute-actuate loop.
As part of a related effort with a narrower focus on the domain of sensor network macroprogramming — a technique that aims to aid the wide adoption of networked sensing by providing the domain expert the ability to specify their applications at a high level of abstraction — we have explored techniques to bring Web services in the gamut of sensor network macroprogramming. Our research addresses the challenges faced by developers of systems where sensors (e.g., RFID badge sensors in an office) interact with pre-existing larger software components exposed as Web services (e.g., the office personnel access control database). As part of our work, we have proposed extensions to the data-driven ATaG macroprogramming language using which developers can easily incorporate existing Web services in their applications.

We have incorporated our continued research in the above areas into Srijan (§ 5.5), which provides an easy-to-use graphical front-end to the various steps involved in developing an application using the ATaG macroprogramming framework.

6.6. Addressing Middleware Challenges in Large Scale Mobile Social Networks of the Future
Participants: Sara Hachem, Valérie Issarny, Animesh Pathak, Amir Seyedi.

With the increased prevalence of advanced mobile devices (the so-called “smart” phones), interest has grown in Mobile Social Ecosystems (MSE), where users not only access traditional on line Web-based social networks using their mobile devices, but are also able to use the context information provided by these devices to further enrich their interactions. In complex mobile social ecosystems of the future, the heterogeneity of software platforms on constituent nodes, combined with their intrinsic distributed nature and heterogeneity in representation of data and context, as well as user’s privacy and trust concerns, raises the need for middleware support for the development of mobile social applications. We believe that the development of mobile social applications can be greatly simplified by the presence of middleware support. To that end, we have been working on addressing the following challenges:

- **Semantic models for mobile social ecosystems.** In order to enable re-use of data between different social applications run by the same user, we have proposed an expressive and extensible model using semantic techniques to represent MSE and the interactions possible in them. This supports semantic interoperability between separately developed applications and minimizes resource-consuming operations such as data mapping and replication.
- **Efficient decentralized storage of social data.** Instead of storing the social knowledge of the whole world with a single provider — a practice performed today by common social networks such as Facebook — which can lead to privacy issues, our research endeavors to propose a middleware using which users can store their personal knowledge in a distributed manner on the devices owned by them (e.g., smart phone, home desktop, laptop). This also allows users to provide selective access to other users based on semantically defined access control policies.
- **Socially aware policies for access control.** Since social data is private and sensitive in nature, we have proposed a policy framework [21] where the user can specify both the data to be protected as well as the relevant set of peers with access to that data in a socially-aware manner (e.g., “only let my colleagues know my location during weekdays from 9 – 5”). This policy framework can be used as a guard around the user’s knowledge base, allowing access only to authorized peers. We are also working on providing end-users an easy to use editor so as to be able to specify these socially-aware policies easily.
- **Social data extraction from existing sources.** Our research includes work in enabling users to populate their social knowledge base by extracting data from their existing repositories. We have identified two types of sources of such data. The first already contain social links such as “friendship” in addition to general information, while the second do not contain social links, but may contain information which can be correlated to infer social links (e.g., call and SMS logs). We are working on a framework where adapters can be written for the former using their API to import their data; while for the latter, inference algorithms can be used to correlate data and guess/recommend social links.
Inferring trust from proximity. In mobile social network, highly sensitive private data is at risk of being shared with unwanted peers, since users may not have any knowledge about the users they socially connect with. Trust management then appears as a promising decision support for mobile users in establishing social links. However, while the literature is rich in trust models, most approaches lack appropriate trust bootstrapping, i.e., the initialization of trust values. In [24], we address this challenge by introducing proximity-based trust initialization based on the users’ behavioral data available from their mobile devices or other types of social interactions. The proposed approach is further assessed in the context of mobile social networking using users behavioral data collected by the MIT reality mining project. Results show that the inferred trust values correlate with the self-reported survey of users relationships.

We have incorporated our research in the above areas into Yarta [25], a middleware for mobile social applications. Our prototype middleware, as discussed in §5.6, currently supports application development for laptops as well as Android-powered smart phones, providing distributed storage of semantically-modeled social knowledge guarded by a rich policy framework.
6. New Results

6.1. Models and abstractions for distributed systems

This section summarizes the major results obtained by the ASAP team that relate to the foundations of distributed systems.

6.1.1. The weakest failure detector to implement a register in asynchronous systems with hybrid communication

Participants: Damien Imbs, Michel Raynal.

This work introduces an asynchronous crash-prone hybrid system model. The system is hybrid in the way the processes can communicate. On the one side, a process can send messages to any other process. On another side, the processes are partitioned into clusters and each cluster has its own read/write shared memory. In addition to the model, a main contribution of the work concerns the implementation of an atomic register in this system model. More precisely, a new failure detector (denoted $M\Sigma$) is introduced and it is shown that, when considering the information on failures needed to implement a register, this failure detector is the weakest. To that end, the work presents an $M\Sigma$-based algorithm that builds a register in the considered hybrid system model and shows that it is possible to extract $M\Sigma$ from any failure detector-based algorithm that implements a register in this model. The work also (a) shows that $M\Sigma$ is strictly weaker than $\Sigma$ (which is the weakest failure detector to implement a register in a classical message-passing system) and (b) presents a necessary and sufficient condition to implement $M\Sigma$ in a hybrid communication system.

This work has been published in SSS 2011 [38].

6.1.2. The universe of symmetry breaking tasks

Participants: Damien Imbs, Michel Raynal.

Processes in a concurrent system need to coordinate using a shared memory or a message-passing subsystem in order to solve agreement tasks such as, for example, consensus or set agreement. However, coordination is often needed to ‘break the symmetry’ of processes that are initially in the same state, for example, to get exclusive access to a shared resource, to get distinct names or to elect a leader.

This work introduces and studies the family of generalized symmetry breaking (GSB) tasks, that includes election, renaming and many other symmetry breaking tasks. Differently from agreement tasks, a GSB task is “inputless”, in the sense that processes do not propose values; the task only specifies the symmetry breaking requirement, independently of the system’s initial state (where processes differ only on their identifiers). Among various results characterizing the family of GSB tasks, it is shown that (non adaptive) perfect renaming is universal for all GSB tasks.

This work was done in collaboration with Sergio Rajsbaum from the Universidad Nacional Autonoma de Mexico and was published in SIROCCO 2011 [36].

6.1.3. Read invisibility, virtual world consistency and probabilistic permissiveness are compatible

Participants: Tyler Crain, Damien Imbs, Michel Raynal.
The aim of a Software Transactional Memory (STM) is to discharge the programmers from the management of synchronization in multiprocess programs that access concurrent objects. To that end, an STM system provides the programmer with the concept of a transaction. The job of the programmer is to design each process the application is made up of as a sequence of transactions. A transaction is a piece of code that accesses concurrent objects, but contains no explicit synchronization statement. It is the job of the underlying STM system to provide the illusion that each transaction appears as being executed atomically. Of course, for efficiency, an STM system has to allow transactions to execute concurrently. Consequently, due to the underlying STM concurrency management, a transaction commits or aborts.

This work studies the relation between two STM properties (read invisibility and permissiveness) and two consistency conditions for STM systems, namely, opacity and virtual world consistency. Both conditions ensure that any transaction (be it a committed or an aborted transaction) reads values from a consistent global state, a noteworthy property if one wants to prevent abnormal behavior from concurrent transactions that behave correctly when executed alone. A read operation issued by a transaction is invisible if it does not entail shared memory modifications. This is an important property that favors efficiency and privacy. An STM system is permissive (respectively probabilistically permissive) with respect to a consistency condition if it accepts (respectively accepts with positive probability) every history that satisfies the condition. This is a crucial property as a permissive STM system never aborts a transaction “for free”. The work first shows that read invisibility, probabilistic permissiveness and opacity are incompatible, which means that there is no probabilistically permissive STM system that implements opacity while ensuring read invisibility. It then shows that read invisibility, probabilistic permissiveness and virtual world consistency are compatible. To that end the work describes a new STM protocol called IR_VWC_P. This protocol presents additional noteworthy features: it uses only base read/write objects and locks which are used only at commit time; and, in favorable circumstances, the cost of a read operation is $O(1)$.

This work has been published in ICA3PP 2011 [29].

6.1.4. Towards a universal construction for transaction-based multiprocess programs

Participants: Tyler Crain, Damien Imbs, Michel Raynal.

The aim of a Software Transactional Memory (STM) system is to discharge the programmer from the explicit management of synchronization issues. The programmer’s job resides in the design of multiprocess programs in which processes are made up of transactions, each transaction being an atomic execution unit that accesses concurrent objects. The important point is that the programmer has to focus her/his efforts only on the parts of code which have to be atomic execution units without worrying on the way the corresponding synchronization has to be realized.

Non-trivial STM systems allow transactions to execute concurrently and rely on the notion of commit/abort of a transaction in order to solve their conflicts on the objects they access simultaneously. In some cases, the management of aborted transactions is left to the programmer. In other cases, the underlying system scheduler is appropriately modified or an underlying contention manager is used in order that each transaction be (practically always” or with high probability) eventually committed.

This work paper proposed a deterministic STM system in which (1) every invocation of a transaction is executed exactly once and (2) the notion of commit/abort of a transaction remains unknown to the programmer. This system, which imposes restriction neither on the design of processes nor or their concurrency pattern, can be seen as a step towards the design of a deterministic universal construction to execute transaction-based multiprocess programs on top of a multiprocessor. Interestingly, the proposed construction is lock-free (in the sense that it uses no lock).

This work has been published in ICDCN 2012 [30].

6.1.5. A transaction friendly binary search tree

Participants: Tyler Crain, Michel Raynal.
Transactions, which provide optimistic synchronization by avoiding the use of blocking, greatly simplify multithread programming. In fact, the programmer has simply to encapsulate sequential operations or existing critical sections into transactions to obtain a safe concurrent program. Programmers have thus started evaluating transactional memory using data structures originally designed for pessimistic (i.e., non-optimistic) synchronization, whose prominent example is the red-black tree library developed by Oracle Labs that is part of STAMP and microbench distributions. Unfortunately, existing data structures are badly suited for optimistic synchronization as they rely on strong structural invariants, like logarithmic tree depth, to bound the step complexity of pessimistically synchronized accesses. By contrast, this complexity does not apply to optimistically synchronized accesses thus making the invariants overly conservative. More dramatically, guaranteeing such invariants tends to increase the probability of aborting and restarting the same access before it completes. We introduced a concurrent binary search tree that breaks transiently its balance structural invariants for efficiency, a property we call transaction-friendly. This new tree outperforms the existing transaction-based version of the AVL and the red-black trees. Its key novelty stems from the decoupling of update operations: they are split into one transaction that modifies the abstraction state and multiple ones that restructure its tree implementation. The resulting transaction-friendly library trades aborts for few additional access steps and, in particular, it speeds up a transaction-based travel reservation application by up to 3.5X. This work was done in collaboration with Vincent Gramoli from EPFL Lausanne, and is described in [52].

6.1.6. Relations linking failure detectors associated with \( k \)-set agreement in message-passing systems

**Participants:** Achour Mostefaoui, Michel Raynal, Julien Stainer.

The \( k \)-set agreement problem is a coordination problem where each process is assumed to propose a value and each process that does not crash has to decide a value such that each decided value is a proposed value and at most \( k \) different values are decided. While it can always be solved in synchronous systems, \( k \)-set agreement has no solution in asynchronous send/receive message-passing systems where up to \( t \geq k \) processes may crash.

A failure detector is a distributed oracle that provides processes with additional information related to failed processes and can consequently be used to enrich the computability power of asynchronous send/receive message-passing systems. Several failure detectors have been proposed to circumvent the impossibility of \( k \)-set agreement in pure asynchronous send/receive message-passing systems. Considering three of them (namely, the generalized quorum failure detector \( \Sigma_k \), the generalized loneliness failure detector \( \xi_k \) and the generalized eventual leader failure detector \( \Omega_k \)) this work investigates their computability power and the relations that link them. There are three main contributions: (a) it shows that the failure detector \( \Omega_k \) and the eventual version of \( \xi_k \) have the same computational power; (b) it shows that \( \xi_k \) is realistic if and only if \( k \geq n/2 \); and (c) it gives an exact characterization of the difference between \( \xi_k \) (that is too strong for \( k \)-set agreement) and \( \Sigma_k \) (that is too weak for \( k \)-set agreement). This work was published at SSS 2011 [45].

6.1.7. The price of anonymity: optimal consensus despite asynchrony, crash and anonymity

**Participant:** Michel Raynal.

This work [23], done in collaboration with François Bonnet, from JAIST, Japan, addresses the consensus problem in asynchronous systems prone to process crashes, where additionally the processes are anonymous (they cannot be distinguished one from the other: they have no name and execute the same code). To circumvent the three computational adversaries (asynchrony, failures and anonymity) each process is provided with a failure detector of a class denoted \( \psi \), that gives it an upper bound on the number of processes that are currently alive (in a non-anonymous system, the classes \( \psi \) and \( \mathcal{P} \) - the class of perfect failure detectors - are equivalent).

The first part presents a simple \( \psi \)-based consensus algorithm where the processes decide in \( 2t+1 \) asynchronous rounds (where \( t \) is an upper bound on the number of faulty processes). It then shows one of its main results, namely, \( 2t+1 \) is a lower bound for consensus in the anonymous systems equipped with \( \psi \). The second contribution addresses early-decision. The paper presents and proves correct an early-deciding algorithm where the processes decide in \( \min(2f+2,2t+1) \) asynchronous rounds (where \( f \) is the actual number
of process failures). This leads to think that anonymity doubles the cost (wrt synchronous systems) and it is conjectured that \( \min(2f + 2, 2t + 1) \) is the corresponding lower bound.

The work finally considers the \( k \)-set agreement problem in anonymous systems. It first shows that the previous \( \psi \)-based consensus algorithm solves the \( k \)-set agreement problem in \( R_t = 2 \left\lfloor \frac{t}{k} \right\rfloor + 1 \) asynchronous rounds. Then, considering a family of failure detector classes \( \{ \psi_k \}_{0 \leq k \leq n} \) that generalizes the class \( \psi_0 \), the paper presents an algorithm that solves the \( k \)-set agreement in \( R_t = 2 \left\lfloor \frac{t}{k-1} \right\rfloor + 1 \) asynchronous rounds. This last formula relates the cost \( R_t \), the coordination degree of the problem \( (k) \), the maximum number of failures \( (t) \) and the strength \( (\ell) \) of the underlying failure detector.

### 6.1.8. On the road to the Weakest Failure Detector for \( k \)-Set Agreement in Message-passing Systems

**Participant:** Michel Raynal.

In the \( k \)-set agreement problem, each process (in a set of \( n \) processes) proposes a value and has to decide a proposed value in such a way that at most \( k \) different values are decided. While this problem can easily be solved in asynchronous systems prone to \( t \) process crashes when \( k > t \), it cannot be solved when \( k \leq t \). Since several years, the failure detector-based approach has been investigated to circumvent this impossibility. While the weakest failure detector class to solve the \( k \)-set agreement problem in read/write shared-memory systems has recently been discovered (PODC 2009), the situation is different in message-passing systems where the weakest failure detector classes are known only for the extreme cases \( k = 1 \) (consensus) and \( k = n - 1 \) (set agreement).

This work [22], done in collaboration with François Bonnet, from JAIST, Japan, has four contributions whose aim is to help pave the way to discover the weakest failure detector class for \( k \)-set agreement in message-passing systems. These contributions are the following. (a) The first is a new failure detector class, denoted \( \Pi_k \), that is such that \( \Pi_1 = \Sigma \times \Omega \) (the weakest class for \( k = 1 \)), and \( \Pi_{n-1} = \mathcal{L} \) (the weakest class for \( k = n - 1 \)). (b) The second is an investigation of the structure of \( \Pi_k \) that shows that \( \Pi_k \) is the combination of two failures detector classes \( \Sigma_k \) (that is new) and \( \Omega_k \) (they generalize the previous "quorums" and "eventual leaders" failure detectors classes, respectively). (c) The third contribution concerns \( \Sigma_k \) that is shown to be necessary requirement (as far as information on failure is concerned) to solve the \( k \)-set agreement problem in message-passing systems. (d) Finally, the last contribution is a \( \Pi_{n-1} \)-based algorithm that solves the \( (n-1) \)-set agreement problem. This algorithm provides us with a new algorithmic insight on the way the \( (n-1) \)-set agreement problem can be solved in asynchronous message-passing systems. It is hoped that these contributions will help discover the weakest failure detector class for \( k \)-set agreement in message-passing systems.

### 6.1.9. A non-topological proof for the impossibility of \( k \)-set agreement.

**Participant:** Armando Castañeda.

This work was done in collaboration with Hagit Attiya, from Technion, Haifa, Israel. In the \( k \)-set agreement task each process proposes a value, and each correct process has to decide a value which was proposed, so that at most \( k \) distinct values are decided. Using topological arguments it has been proved that \( k \)-set agreement is unsolvable in the asynchronous wait-free read/write shared memory model, when \( k < n \), the number of processes.

This work [34] focuses on a simple, non-topological impossibility proof of \( k \)-set agreement. The proof depends on two simple properties of the immediate snapshot executions, a subset of all possible executions, and on the well known handshaking lemma stating that every graph has an even number of vertices with odd degree.

The paper was presented in the 13th Int’l Symposium on Stabilization, Safety, and Security of Distributed Systems (SSS’11) in Grenoble, France. The journal version of the paper was submitted to Theoretical Computer Science.
6.1.10. Enriching the reduction map of sub-consensus tasks

Participants: Armando Castañeda, Damien Imbs, Michel Raynal.

This work [51] was done in collaboration with Sergio Rajsbaum from the Universidad Nacional Autonoma de Mexico.

Understanding the relative computability power of tasks, in the presence of asynchrony and failures, is a central concern of distributed computing theory. In the wait-free case, where the system consists of \( n \) processes and any of them can fail by crashing, substantial attention has been devoted to understanding the relative power of the subconsensus family of tasks, which are too weak to solve consensus for two processes. The first major results showed that set agreement and renaming (except for some particular values of \( n \)) cannot be solved wait-free in read/write memory. Then it was proved that renaming is strictly weaker than set agreement (when \( n \) is odd).

This work considers a natural family of subconsensus tasks that includes set agreement, renaming and other generalized symmetry breaking (GSB) tasks. It extends previous results, and proves various new results about when there is a reduction and when not, among these tasks. Among other results, the work shows that there are incomparable subconsensus tasks.

6.1.11. Byzantine Consensus Decidability

Participants: Achour Mostefaoui, Michel Raynal.

Solving the consensus problem requires in one way or another that the underlying system satisfies synchrony assumptions. Considering a system of \( n \) processes where up to \( t < n/3 \) may commit Byzantine failures, we proposed in [26] a necessary and sufficient synchrony assumption to solve consensus.

Such a condition is formulated with the notions of a symmetric synchrony property and property ambiguity. A symmetric synchrony property is a set of graphs, where each graph corresponds to a set of bi-directional eventually synchronous links among correct processes. Intuitively, a property is ambiguous if it contains a graph whose connected components are such that it is impossible to distinguish a connected component that contains correct processes only from a connected component that contains faulty processes only. The paper connects then the notion of a symmetric synchrony property with the notion of eventual bi-source, and shows that the existence of a virtual \( \diamondsuit[t+1] \) bi-source is a necessary and sufficient condition to solve consensus in presence of up to \( t \) Byzantine processes in systems with bi-directional links and message authentication. Finding necessary and sufficient synchrony conditions when links are timely in one direction only, or when processes cannot sign messages, still remains open (and very challenging) problems.

6.1.12. Solving \( k \)-set agreement in message-passing systems

Participants: Achour Mostefaoui, Michel Raynal, Julien Stainer.

The \( k \)-set agreement problem is a coordination problem where each process is assumed to propose a value and each process that does not crash has to decide a value such that each decided value is a proposed value and at most \( k \) different values are decided. While it can always be solved in synchronous systems, \( k \)-set agreement has no solution in asynchronous send/receive message-passing systems where up to \( t \geq k \) processes may crash.

A failure detector is a distributed oracle that provides processes with additional information related to failed processes and can consequently be used to enrich the computability power of asynchronous send/receive message-passing systems. Several failure detectors have been proposed to circumvent the impossibility of \( k \)-set agreement in pure asynchronous send/receive message-passing systems. Considering three of them (namely, the generalized quorum failure detector \( \Sigma_k \), the generalized loneliness failure detector \( \mathcal{L}_k \) and the generalized eventual leader failure detector \( \mathcal{O}_k \)), we investigated their computability power and the relations that link them in [45]. It has three main contributions: (a) it shows that the failure detector \( \mathcal{O}_k \) and the eventual version of \( \mathcal{L}_k \) have the same computational power; (b) it shows that \( \mathcal{L}_k \) is realistic if and only if \( k \geq n/2 \); and (c) it gives an exact characterization of the difference between \( \mathcal{L}_k \) (that is too strong for \( k \)-set agreement) and \( \Sigma_k \) (that is too weak for \( k \)-set agreement).
6.1.13. Efficient Implementations of Concurrent Objects  
Participants: Achour Mostefaoui, Michel Raynal.

As introduced by Taubenfeld, a contention-sensitive implementation of a concurrent object is an implementation such that the overhead introduced by locking is eliminated in the common cases, i.e., when there is no contention or when the operations accessing concurrently the object are non-interfering. In [44], we present a methodological construction of a contention-sensitive implementation of a concurrent stack. In a contention-free context a push or pop operation does not rest on a lock mechanism and needs only six accesses to the shared memory. In case of concurrency a single lock is required. Moreover, the implementation is starvation-free (any operation is eventually executed). The paper, that presents the algorithms in an incremental way, visits also a family of liveness conditions and important concurrency-related concepts such as the notion of an abortable object.

6.2. Large-scale and user-centric distributed system  

This section summarizes the major results obtained by the team in 2011 in the context of large-scale distributed systems and social networks. This includes the results obtained within the GOSSPLE ERC project, which encompass two types of social networks: explicit and implicit.

**Explicit networks** connect users based on explicit social connections. In Facebook or MySpace, users issue and accept friendship requests. In Twitter, they decide that they wish to follow the tweets of specific users. In all cases, the topology of the resulting network reflects the choices of users and often consists of links that already exist between real people. Explicit networks are therefore very useful in reinforcing and exploiting existing connections but provide little support for discovering new content.

**Implicit networks** complement explicit ones by providing each user with a set of anonymous acquaintances that share similar interests, that visit similar websites or that have otherwise similar profiles. Different from explicit networks, implicit ones are naturally suited to support the discovery of new content. In previous work [1], we exploited this network to improve web navigation. In the following, we consider additional applications encompassing news dissemination, online transactions, and recommendation.

6.2.1. WhatsUp: P2P news recommender  
Participants: Antoine Boutet, Davide Frey, Anne-Marie Kermarrec.

The main application in the context of GOSSPLE is WhatsUp, an instant news system designed for a large-scale network with no central authority. WhatsUp builds an implicit social network based on the opinions users express about the news items they receive (like-dislike). This is achieved through an obfuscation mechanism that does not require users to ever reveal their exact profiles. WhatsUp disseminates news items through a novel heterogeneous gossip protocol that biases the choice of its targets towards those with similar interests and amplifies dissemination based on the level of interest in every news item. WhatsUp outperforms various alternatives in terms of accurate and complete delivery of relevant news items while preserving the fundamental advantages of standard gossip: namely simplicity of deployment and robustness. This work has been carried out in collaboration with Rachid Guerraoui from EPFL and was demonstrated during the different local events.

6.2.2. Personalized top-k processing  
Participant: Anne-Marie Kermarrec.

Another way to improve the experience of users on the web is to personalize top-k queries. In collaboration with Xiao Bai and Vincent Leroy from Yahoo! Research in Barcelona and Rachid Guerraoui from EPFL Lausanne we, therefore, introduced P4Q, a fully decentralized gossip-based protocol to personalize query processing in social tagging systems. P4Q dynamically associates each user with social acquaintances sharing similar tagging behaviors. Queries are gossiped among such acquaintances, computed on the fly in a collaborative, yet partitioned manner, and results are iteratively refined and returned to the querier. Analytical and experimental evaluations convey the scalability of P4Q for top-k query processing, as well its inherent ability to cope with users updating profiles and departing. The work appeared in the ACM transactions of database systems [12].
6.2.3. Social Market
Participants: Davide Frey, Arnaud Jegou, Anne-Marie Kermarrec.

The ability to identify people that share one’s own interests is one of the most intriguing promises of the Web 2.0 driving user-centric applications such as recommendation systems or collaborative marketplaces. To be truly useful, however, information about other users also needs to be associated with some notion of trust. Consider a user wishing to sell a concert ticket. Not only must she find someone who is interested in the concert, but she must also make sure she can trust this person to pay for it. Social Market (SM) solve this problem by allowing users to identify and build connections to other users that can provide interesting goods or information and that are also reachable through a trusted path on an explicit social network like Facebook. This convergence of implicit and explicit networks yields TAPS, a novel gossip protocol that can be applied in applications devoted to commercial transactions, or to add robustness to standard gossip applications like dissemination or recommendation systems.

This work has been published at SSS 2011 [33], and an extended version bringing better performances and strong privacy guarantees have recently been submitted for publication.

6.2.4. Member classification and party characteristics in Twitter
Participant: Antoine Boutet.

In modern politics, parties and individual candidates must have an online presence and usually have dedicated social media coordinators. In this context, real time member classification and party characterization, taking into account the dynamic nature of social media, are essential to highlight the main differences between parties and to monitor their activities, influences, structures, contents and mood. This work [53] was been done in collaboration with E. Yoneki from Computer Lab, Cambridge, UK.

6.2.5. Graph Drawing and Visual Recommendations
Participants: Anne-Marie Kermarrec, Afshin Moin.

An important aspect of social network is their graph structure. In a collaboration with Vincent Leroy (Yahoo! Research) and Gilles Tredan (TU Berlin) [41], we started from this structure to propose a decentralized gossip-based algorithm called SoCS (Social Coordinate Systems). SoCS achieves efficient distributed social graph embedding using a force-based graph embedding technique to extract communities from a graph. SoCS (i) scales to large dynamic graph, aggregating the computing power of individual nodes and, (ii) avoids a central entity controlling users sensitive data such as relations and preferences. We evaluated SoCS using two different force-based models and compare them in the context of a generated Kleinberg small-world topology. More specifically, we showed that the SoCS graph embedding enables to clearly distinguish between short and long-range links. We also evaluate SoCS against a real DBLP data set, showing that removed links are correctly predicted.

Graph structures are also at the basis of our work on energy/force-based models for graph visualization. We applied visualization both to social network and in the context of recommendation systems. In particular we are working on an SVD-like algorithm for drawing precise 2-dimensional visual recommendations based on Principal Component Analysis (PCA) and Curvilinear Component Analysis (CCA).

6.2.6. Private Similarity Computation in Distributed Systems: from Cryptography to Differential Privacy
Participants: Mohammad Alaggan, Anne-Marie Kermarrec.

The use of personal data in the context of social networks raises important concerns about privacy. In a collaboration [24] with Sébastien Gambs from the CIDre team, we addressed the problem of computing the similarity between two users (a key operation in an implicit social network [1]) while preserving their privacy in a fully decentralized system and for the passive adversary model. First, we introduced a two-party protocol for privately computing a threshold version of the similarity and applied it to well-known similarity measures such as the scalar product and the cosine similarity. The output of this protocol is only one bit
of information telling whether or not two users are similar beyond a predetermined threshold. Afterwards, we explored the computation of the exact and threshold similarity within the context of differential privacy. Differential privacy is a recent notion developed within the field of private data analysis guaranteeing that an adversary that observes the output of the differentially private mechanism, will only gain a negligible advantage (up to a privacy parameter) from the presence (or absence) of a particular item in the profile of a user. This provides a strong privacy guarantee that holds independently of the auxiliary knowledge that the adversary might have. More specifically, we designed several differentially private variants of the exact and threshold protocols that rely on the addition of random noise tailored to the sensitivity of the considered similarity measure. We also analyzed their complexity as well as their impact on the utility of the resulting similarity measure.

6.2.7. Constellation: Programming decentralized social networks

Participants: François Taiani, Anne-Marie Kermarrec.

As they continue to grow, social and collaborative applications (e.g. twitter, Facebook, digg) are increasingly calling for disruptive distributed solutions than can cater for the millions of users these applications serve daily, in hundreds of countries, over a wide variety of devices. To address these challenges, fully decentralized versions of social and collaborative applications are progressively emerging that seek to provide naturally scalable solutions to deliver their services. Gossip protocols in particular appear as a natural solution to implement these decentralized versions, as they intrinsically tend to be highly resilient, efficient, and scalable.

Social applications based on gossip have however been limited so far to relatively homogeneous systems: They typically rely on one similarity measure to self-organize large amount of distributed users in implicit communities, and thus offer powerful means to search, mine, and serve personalized data in a distributed manner.

We posit in this work [54] that we now need to move to more complex gossip-based social applications that can cater for different types of data and similarity, organized in multiple levels of abstraction. Exploring, designing, and evaluating such novel approaches is unfortunately time-consuming and error-prone. To help in this task, we have started to design a new programming language, Constellation, that seeks to simplify the realization and experimentation with social gossip-based applications. Constellation is based on two central observations: (i) future decentralized social applications will need to handle heterogeneous forms of data and self-organization, and (ii) to offer more powerful services, these applications will need to move beyond physical nodes to encompass richer data structures organized in virtualized levels of abstractions.

6.2.8. Leveraging content interconnections for efficient data storage.

Participants: Anne-Marie Kermarrec, Konstantinos Kloudas, François Taiani.

Traffic generated by User Generated Content (UGC) sharing sites, such as YouTube, accounts for a substantial fraction of today’s global Internet load. This success has however brought a number of key technical challenges, crucial for system sustainability and user experience. One of them is the need to place content close to consumers, so that user perceived latency is reduced and bandwidth utilization is minimized. In a joint work with Kevin Huguenin, we try to tackle this problem by leveraging the fact that content hosted by these sites is interconnected, forming a content graph that as shown by former works, has an important impact on a file’s view pattern. In our work titled “Recommended nearby in UGC delivery networks: leveraging geographical and content locality”, we focused on YouTube and we studied how two types of locality previously analyzed in isolation in UGC systems, namely content locality (a.k.a graph locality, induced by the related video feature) and geographic locality, are in fact correlated. Leveraging the above finding, we proposed a novel algorithm for replica placement that tries to predict where future views for a video will come from based on the video’s related videos and places its replicas accordingly. This work has been submitted for publication.

6.2.9. Transparent Componentization: High-level (Re)configurable Programming for Evolving Distributed Systems

Participants: François Taiani, Marin Bertier, Anne-Marie Kermarrec.
This work was done in collaboration Component frameworks and high-level distributed languages have been widely used to develop distributed systems, and provide complementary advantages: Whereas component frameworks foster composability, reusability, and (re)configurability; distributed languages focus on behavior, simplicity and programmability. We argue that both types of approach should be brought together to help develop complex adaptive systems, and we propose an approach to combines both technologies without compromising on any of their benefits. Our approach, termed Transparent Componentization [43], automatically maps a high-level distributed specification onto a underlying component framework. It thus allows developers to focus on the programmatic description of a distributed system’s behavior, while retaining the benefits of a component architecture. As a proof of concept, we present WhispersKit, a programming environment for gossip-based distributed systems. Our evaluation shows that WhispersKit successfully retains the simplicity and understandability of high-level distributed language while providing efficient and transparent reconfigurability thanks to its component underpinnings.

6.2.10. Efficient peer-to-peer backup services through buffering at the edge
Participants: Anne-Marie Kermarrec, Alexandre Van Kempen.

The availability of end devices of peer-to-peer storage and backup systems has been shown critical for usability and for system reliability in practice. This has led to the adoption of hybrid architectures composed of both peers and servers. Such architectures mask the instability of peers thus approaching the performances of client-server systems while providing scalability at a low cost. In this work [31] - done in collaboration with Erwan Le Merrer, Serge DeFrance, Nicolas Le Scouarnec and Gilles Straub from Technicolor, Rennes, France - we advocate the replacement of such servers by a cloud of residential gateways, as they are already present in users’ homes, thus pushing the required stable components at the edge of the network. In our gateway-assisted system, gateways act as buffers between peers, compensating for their intrinsic instability. This enables to offload backup tasks quickly from the user’s machine to the gateway, while significantly lowering the retrieval time of backed up data. We evaluate our proposal using real world traces including existing traces from Skype and Jabber as well as a trace of residential gateways for availability, and a residential broadband trace for bandwidth. Results show that the time required to backup data in the network is comparable to a server-assisted approach, while substantially improving the time to restore data, which drops from a few days to a few hours. As gateways are becoming increasingly powerful in order to enable new services, we expect such a proposal to be leveraged on a short term basis.

6.2.11. Commutative Replicated Data Type for Semantic Stores
Participant: Stéphane Weiss.

This work has been done in collaboration with Khaled Aslan (Université de Nantes - Lina), Pascal Molli (Université de Nantes - Lina) and Hala Skaf-Molli (Université de Nantes - Lina). Web 2.0 tools are currently evolving to embrace semantic web technologies. Blogs, CMS, Wikis, social networks and real-time notifications, integrate ways to provide semantic annotations and therefore contribute to the linked data and more generally to the semantic web vision. This evolution generates a lot of semantic datasets of different qualities, different trust levels and partially replicated. This raises the issue of managing the consistency among these replicas. This issue is challenging because semantic data-spaces can be very large, they can be managed by autonomous participants and the number of replicas is unknown. A new class of algorithms called Commutative Replicated Data Type are emerging for ensuring eventual consistency of highly dynamic content on P2P networks. We define C-Set [25] a CRDT specifically designed to be integrated in Triple-stores. C-Set allows efficient P2P synchronization of an arbitrary number of autonomous semantic stores.

6.2.12. Building large scale platform for chemical program
Participants: Marin Bertier, Achour Mostefaoui.

This work [28] was done in collaboration with the Myriads project team.
Chemical programming is a promising paradigm to design autonomic systems. Within such a paradigm, computations can be seen as chemical reactions controlled by a set of chemical rules. In other words, data are molecules of a chemical solutions, reacting together to produce new data. Reactions take place in an implicitly parallel, and autonomic fashion.

Our objective was to design a distributed chemical platform bringing such concepts. This platform should be adapted to large scale distributed system to benefit at his best the inherent distribution of chemical program.
ASCOLA Project-Team

6. New Results

6.1. Aspects

Participants: Rémi Douence, Abdelhakim Hannousse, Ismael Mejía, Jacques Noyé, Angel Núñez, Nicolas Tabareau, Mario Südholt.

We have provided results on three subject matters in the domain of aspect-oriented software development: the foundations of aspects, aspect languages and their applications, as well as the use of aspects for the manipulation of distributed systems.

As a side note on the form of this document, the reader should be aware that much of our work reported in this section is not exclusive to AOSD but also contributes to software composition issues in a larger sense. This applies, in particular, to the results cited in the subsections on aspect languages and distributed aspects.

6.1.1. Foundations of Aspects

In the domain of foundations of AOSD, we have mainly provided new results on the preservation of formal correctness properties in the presence of aspects and how AOP can be modeled using category theory.

- **Property preservation in the presence of aspects.** In general aspects can arbitrarily change the semantics of the base program. We have identified categories of aspects that preserve class of properties of the base programs \[17\]. This approach makes it possible to prove once and for all that a category of aspects preserve a class of properties. The categories are defined with respect to the semantics of the woven program as well as with restricted aspect languages. In this latter case, languages are defined by grammars hence checking for property preservation boils down to a syntactic check for aspects. Classes of properties are defined using a subset of temporal logic both for sequential and concurrent programs. Our framework is abstract in that it does not depend on the actual programming language but only on conditions on its small step semantics.

- **AOP and category theory.** Aspect-Oriented Programming (AOP) started ten years ago with the remark that modularization of so-called crosscutting functionalities is a fundamental problem for the engineering of large-scale applications. Originating at Xerox PARC, this observation has sparked the development of a new style of programming that is gradually gaining traction. However, AOP lacks theoretical foundations to clarify new ideas showing up in its wake. We have proposed to put a bridge between AOP and the notion of 2-category to enhance the conceptual understanding of AOP \[36\], \[46\]. Starting from the connection between the \(\lambda\)-calculus and the theory of categories, we have provided an internal language for 2-categories and shown how it can be used to define the first categorical semantics for a realistic functional AOP language, called MinAML.

We have later taken advantage of this new categorical framework to introduce the notion of computational 2-monads for AOP \[37\]. We have illustrated their conceptual power by defining a 2-monad for Éric Tanter’s execution levels—which constitutes the first algebraic semantics for execution levels—and then introducing the first exception monad transformer specific to AOP that gives rise to a non-flat semantics for exceptions by taking levels into account.

- **Membranes for AOP.** AOP aims to enhance modularity and reusability in software systems offering an abstraction mechanism to deal with crosscutting concerns. However, in most aspect-oriented languages aspects have a global view of computation that actually introduces a strong coupling between advised code and aspect code, hampering modularity. Proposals that address this problem can be classified in two categories: the first one focuses on controlling aspect scoping, i.e. the visibility of join points to aspects, while the second one focuses on protecting software units from aspects. As a new approach, we have proposed programmable membranes (inspired by the work of Boudol for distributed processes) to control aspect influence over software systems as a uniform framework that unifies and extends previous approaches \[49\].
• **Aspects and invertible program restructurations.** When one chooses a main axis of structural decomposition of a problem, the other axes become secondary. This is known as the tyranny of the dominant decomposition and this hinders program maintenance of secondary concerns. We propose to use automatic program transformations built on top of refactorers in order to solve this tyranny issue \[47\]. This offers a new approach to the expression problem by always providing the right view to the programmers.

6.1.2. Aspect Languages

This year we have provided major results on the integration of programming features for objects, events and aspects, and aspect execution infrastructures. Furthermore, we have investigated the application of aspect languages to the Java security model, component-based software development and software product lines.

• **Integrating OOP, AOP and EBP:** Object-Oriented Programming (OOP) has become the de facto programming paradigm. Event-Based Programming (EBP) and Aspect-Oriented Programming (AOP) complement OOP, covering some of its deficiencies when building complex software. Today’s applications combine the three paradigms. However, OOP, EBP and AOP have not yet been properly integrated. Their underlying concepts are in general provided as distinct language constructs, whereas they are not completely orthogonal. This lack of integration and orthogonality complicates the development of software as it reduces its understandability, its composability and increases the required glue code. \[16\] proposes an integration of OOP, EBP and AOP leading to a simple and regular programming model. This model integrates the notions of class and aspect, the notions of event and join point, and the notions of piece of advice, method and event handler. It reduces the number of language constructs while keeping expressiveness and offering additional programming options. ECaesarJ had previously been developed based on these ideas. \[16\] introduces a simpler variant of it, EJava, a plain extension of Java. \[24\] shows that these ideas can also be easily applied to Scala, leading to EScala, and focuses on the idea that the declarative events provided by the model can be seen as object-oriented events, which, unlike global typed events, obey to the basic OO principles (OO modular reasoning, encapsulation and late-binding). An efficient implementation of EScala is described, based on the idea that an event should not be produced in the absence of at least a consumer. This is equivalent to what could be programmed by hand using variants of the observer pattern, except that all the necessary scaffolding is provided by the language compiler and runtime rather than by the programmer.

• **Prototyping and composing Aspect Languages:** \[11\] presents CALI (Common Aspect Languages Interpreter), a framework for rapid prototyping and composition of aspect languages based on interpreters. In practice the common interpreter is actually a thin interpretative layer built on top of Java and implemented as an AspectJ aspect. This interpreter implements common aspect mechanisms and leaves holes to be defined when developing concrete languages. The approach has been validated by implementing prototypes of significant subsets of well-known general-purpose and domain-specific aspect languages (AspectJ, EAOP, COOL and a couple of other small domain-specific aspect languages) and exploring variants of AspectJ. Languages implemented with CALI, for instance AspectJ and COOL, can then be easily composed.

• **Application to security:** It is inevitable that some concerns crosscut a sizeable application, resulting in code scattering and tangling. This issue is particularly severe for security-related concerns: it is difficult to be confident about the security of an application when the implementation of its security-related concerns is scattered all over the code and tangled with other concerns, making global reasoning about security precarious. In \[21\], we consider the case of access control in Java, which turns out to be a crosscutting concern with a non-modular implementation based on runtime stack inspection. We describe the process of modularizing access control in Java by means of AOP. We first show a solution based on AspectJ that must rely on a separate automata infrastructure. We then put forward a novel solution via dynamic deployment of aspects and scoping strategies. Both solutions, apart from providing a modular specification of access control, make it possible to easily express other useful policies such as the Chinese wall policy. However, relying on expressive scope
control results in a compact implementation, which, at the same time, permits the straightforward expression of even more interesting policies.

- **Aspects and software components:** The relationship of aspects and components, as well as their integration as part of real-world infrastructures and application is still subject to many open issues. We have studied, in particular, how crosscutting concerns naturally appear when several architectural views must be considered at the same time. In this context, we have proposed a domain-specific language to specify these architectures. We have proposed an implementation of composable controllers in Fractal as well as composition operators that makes it possible to solve aspects interferences [25]. We have also shown how to formally model the implementation in Uppaal in order to statically check aspects interference [26]. This work has principally been undertaken in the context of the PhD thesis of Abdelhakim Hannouse [12] that has been defended in Sep. 2011.

Another fundamental issue that relates components and aspects is the question whether behavioral protocols can be used to concisely define crosscutting concerns of component-based systems. Furthermore, the protocols should then be instrumental to enable (automatic) reasoning about composition properties of component systems. As part of her PhD thesis (defended in Oct. 2011) [14], Dong Ha Nguyen has defined a notion of aspects that allow crosscutting concerns to be expressed in terms of a class of non-regular behavioral protocols, so-called visibly-pushdown languages. She has developed, most notably, a constructive way of building correct component-based systems that respect such aspect-modified behavioral protocols. Furthermore, she has shown this year how composition properties can generally be verified using model checking techniques.

- **Aspects and software product lines:** In [41], we take a closer look at the difficulties of feature-oriented modularisation of product lines and demonstrate, using a case study in the domain of home automation, how a better modularisation can be achieved with the ECaesarJ programming language, through a type-safe and stable decomposition of a broad spectrum of software abstractions: classes, methods, events, and state machines, based on late binding and mixin composition. A nice property of this modularisation is that it directly captures, at the implementation level, the high-level decomposition in features, without requiring the user to resort on complex transformation technology (the technology is embedded in the compiler).

Furthermore, ASCOLA members have co-edited and contributed to a comprehensive book on aspects, model-driven engineering and product lines [42] (see Sec. 6.2 for details).

### 6.1.3. Aspects for Distributed Systems

In the field of aspects for distributed systems, we have mainly extended the AWED model for distributed aspects (see Sec. 5.1) in order to define and implement a gray-box distributed composition model. We have also worked on a notion of distributed crosscuts that enable causality relationships to be captured using logical clocks.

- **Gray-box distributed composition.** Existing composition and coordination techniques for distributed applications typically support only interface-level (black-box) composition or arbitrary access to the implementation (gray-box or white-box composition). In [20], we have presented a structured approach to the composition of complex distributed software systems that require invasive modifications. Concretely, we have provided three contributions:
  - We have introduced a small kernel composition language for structured gray-box composition using invasive distributed patterns.
  - We have motivated that gray-box composition approaches should be defined and evaluated in terms of the flexibility and control they provide; a notion of degrees of invasiveness has been introduced to help assess this trade-off.
  - We have applied and evaluated it in the context of several studies involving two real-world software systems: benchmarking of grid algorithms with NASGrid and transactional replication with JBoss Cache. As a main result, we have shown that gray-box composition using
invasive distributed patterns allows the declarative and modular definition of evolutions of real-world applications that need moderate to high degrees of invasive modifications.

We have then provided a first framework-based implementation based on our AWED tool and applied it to evolution tasks in OpenMRS, a health information system [31]. The composition framework supports the definition of expressive pattern-based invasive compositions based on the skeleton paradigm for distributed patterns. Concretely, we have shown that the composition framework enables the concise definition of an evolution scenario of OpenMRS that supports the consolidation of patient data from different distributed instances.

Finally, we have studied a model that integrates distributed aspects and actors and thus improves on the robustness properties of existing models for distributed aspects [32].

- **Causal aspects for distributed applications.** In [45], we have applied logical clocks à la Lamport in order to define causality between distributed join points. This enables us to declaratively define pointcuts in a distributed context such as web-based applications in JavaScript. The definition of advice can then be simplified, because they are executed only when necessary and do not have to maintain and check crosscutting information.

### 6.2. Software Composition

**Participants:** Jean-Claude Royer, Hervé Grall, Christine Louberry, Mario Südholt.

ASCOLA’s work on software composition addresses the foundations of software composition methods, the definition and implementation of concrete composition techniques and their application to different functionalities and application domains. This year we have contributed to the questions of how to identify components in legacy software, how to effectively deploy and reconfigure component-based software in pervasive environments, and how to apply software composition techniques to Cloud security as well as software product lines.

- **Component identification.** The communication integrity property is one of the major principles to implement software architectures, however, there is a lack of tooling for assessing the quality of components codes. To cope with this issue, we defined, in [23], a Java component model and a tool for identifying component types based on the communication integrity property. We apply it to several case studies and compare the result with the SOMOX component recovery tool.

- **QoS-driven deployment and reconfiguration of pervasive applications.** [28] presents the Kalimucho platform, a platform for the dynamic reconfiguration of applications on mobiles and constraints devices. First this article focuses on the heuristics implemented by Kalimucho. They support finding a configuration and a deployment matching two criteria: utility and durability. Moreover, we present a case study to experiment the approach of Kalimucho. It confirms that Kalimucho provides a satisfactory execution time for mobile devices.

A second result we have contributed in this context is a two-dimensional QoS model for Kalimucho that supports the QoS-driven deployment of mobile applications [29]. After presenting the definitions of the context and the quality of service considered in this work, this article describes the QoS model and the process allowing finding the best configuration and deployment. Finally, we present several results concerning the execution time of the deployment process of Kalimucho with different devices.

- **Securing the Cloud: cross domain and multi-level concerns.** The evolution of new deployment architectures, as illustrated by the move towards mobile platforms and the Internet Of Services, and the introduction of new security regulations (imposed by national and international regulatory bodies, such as SOX4 or BASEL5) are important constraints in the design and development of service composition. In such a context, it is not sufficient to apply the corresponding adaptations only at the service orchestration or at the choreography level; there is also the need for controlling the impact of new security requirements on several architectural layers. In [35] we have presented a new service model that supports the clean modularization of such crosscutting security concerns.
• **Software composition and product lines.** As part of the AMPLE project (see [http://www.ample-project.net](http://www.ample-project.net)) we have co-edited a book [42] covering the main scientific solutions and techniques on new methods and techniques for software product lines that have been developed in the project. This project aims at improving traditional software product lines engineering using advanced software engineering namely model-driven and aspect-oriented engineering approaches. Software composition takes an important part in this book and it appears in several chapters with models, events, aspects and components.

Chapter [43] provides an overview of software product lines, model-driven engineering and aspect-oriented software development. The challenges to address and the expected benefits are drawn, this is concluded by an overview of the AMPLE approach and its tool chain.

Chapter [40] reviews the specificities of traceability in a product line context starting by identifying the challenges of maintaining traceability for traditional system development and for software product lines. This work defines the concepts that should guide the adoption of a traceability environment for product lines and illustrates these specifications with a concrete example of a traceability repository. It also provides examples of scenarios that use this traceability environment to solve concrete problems.

### 6.3. Cloud Computing, Virtualization and Data Centers

**Participants:** Frederico Alvares, Gustavo Bervian Brand, Thomas Chavrier, Fabien Hermenier, Adrien Lèbre, Thomas Ledoux, Guillaume Le Louët, Jean-Marc Menaud, Hien Nguyen Van, Rémy Pottier, Flavien Quesnel.

In the context of Cloud computing ASCOLA members have principally worked this year on capacity planning solutions for large scale distributed systems. Capacity planning is the process of planning, analyzing, sizing, managing and optimizing capacity to satisfy demand in a timely manner and at a reasonable cost.

Applied to distributed systems like the Cloud, a capacity planning solution must mainly provide necessary resources for the proper execution of applications and respect service-level agreements in a large distributed environment.

The main challenges in this context are: scalability, fault tolerance and reactivity of the solution in a large-scale distributed system; analyzing, sizing, and optimizing resources to minimize the cost (energy, human, hardware etc.); and profiling, adapting application to ensure the quality of service (throughput, response time, availability etc.).

Our solutions are mainly based on virtualized infrastructures. Our main results concern the management and the execution of applications by leveraging virtualization capabilities on cloud infrastructures and the investigation of solutions that aim at optimizing the trade-off between performance and energy costs of both applications and Cloud resources.

#### 6.3.1. Virtualization and Job Management

This year, in cooperation with the Myriads project-team from INRIA Rennes-Bretagne Atlantique, we have continued to address resource-management issues concerning the federation of very large scale platforms. We have completed our approach aiming at the automatic adaptation of both hardware and software resources to the needs of the applications through a unique method. For each application, scientists describe the requirements in terms of both hardware and software expectations through the definition of a Virtual Platform (VP) and a Virtual System Environment (VSE) [18].

In addition, we started to address the design and the implementation of a fully distributed cloud OS. Designing and implementing such models is a tedious and complex task. However, as well as research studies on OSes and hypervisors are complementary at the node level, we showed that virtualization frameworks can benefit from lessons learned from distributed operating system proposals [34]. Leveraging this preliminary result, we designed and developed a first proof-of-concept of a fully distributed scheduler [33]. This system makes it possible to schedule VMs cooperatively and dynamically in large scale distributed systems. Preliminary results showed that our scheduler was more reactive. This building block is a first element of a more complete
cloud OS, entitled DISCOVERY (DIStributed and COoperative mechanisms to manage Virtual EnviRonments autonomicallY) [30]. The ultimate goal of this system is to overcome the main limitations of the traditional server-centric solutions. The system, currently under development, relies on a peer-to-peer model where each agent can efficiently deploy, dynamically schedule and periodically checkpoint the virtual environments s/he manage.

We have contributed new results on the Entropy software and extended our solution VM that features dynamic consolidation. In [44] and [38] we extended our dynamic consolidation manager to take into account not only resource constraints but also the placement constraints of highly available (HA) applications. In fact, most previous dynamic consolidation systems optimize the placement of the VMs according to their resource usage but do not consider the application placement constraints that are required to achieve both high availability and scalability. Our approach provides efficient dynamic consolidation while guaranteeing to the application administrator that placement requirements will be satisfied and relieving the data center administrator of the burden of considering the constraints of the applications when performing maintenance.

In the same domain, Jean-Marc Menaud has defended his habilitation (HdR - Habilitation à diriger des recherches) [13] on Jun. 22, 2011. One part of this HDR focuses on dynamic adaptation strategies for cluster administration. We have proposed a dedicated language for virtual machines management and one particular feature of our solution is to use a constraint solver to provide an appropriate placement. Based on these results, our recent contributions address the problem of data center energy consumption.

Moreover, we have continued to analyze how energy concerns can be addressed in large scale distributed infrastructures.

6.3.2. Optimization of Energy Consumption in Data Centers

As a direct consequence of the increasing popularity of Cloud Computing solutions, data centers are growing at a fast rate and hence have to face difficult energy consumption issue. Available solutions rely on Cloud Computing models and virtualization techniques to scale up/down applications based on their performance metrics. Although those proposals can reduce the energy footprint of applications and, by transitivity, of cloud infrastructures, they do not consider the internal characteristics of applications to finely define a trade-off between QoS properties of applications and their energy footprint. In [22], we propose a self-adaptation approach that considers both application internals and system properties to reduce the energy footprint in cloud infrastructures. Each application and the infrastructure are equipped with their own control loop, which allows them to autonomously optimize their executions. Simulations show that the approach may lead to appreciable energy savings without interfering on application provider revenues.

6.3.3. Cloud Computing and SLA Management

Cloud computing is a paradigm for enabling remote, on-demand access to a set of configurable computing resources as a service. The pay-per-use model enables service providers to offer their services to customers at different Quality-of-Service (QoS) levels. These QoS parameters are used to compose service-level agreements (SLAs) between a service provider and a service consumer. A main challenge for a service provider is to manage SLAs for its service consumers, i.e., automatically determine the appropriate resources required from the lower layer in order to respect the QoS requirements of the consumers. In [27], we have proposed an optimization framework driven by consumer preferences to address the SLA dependencies problem across the different cloud layers as well as the need of flexibility and dynamicity required by the domain of Cloud computing. Our approach aims at selecting the optimal vertical business process designed using cross-layer cloud services and enforcing SLA dependencies between layers. Experimental results demonstrate the flexibility and effectiveness of our approach.

6.4. Foundations of program semantics

Participants: Nicolas Tabareau, Guilhem Jaber.

ASCOLA team members have contributed several results to the foundations of program semantics.
6.4.1. Program Equivalences

Reasoning about program equivalence is a major problem in semantics. This very old topic has many applications, e.g., program verification, compiler correctness or representation independence. It has been understood since the late 1960s that tools and structures arising in mathematical logic and proof theory can usefully be applied to the development of reasoning principles for program equivalence. In recent years, based on the seminal work of Pitts and Stark, the notion of logical relation appeared to be very fruitful for proving the equivalence of programs in the presence of features like general recursive types, general (higher-order) mutable references, and first-class continuations. We have studied the notion of logical relations for proving program equivalences of low level machine codes [39].

We have also developed a forcing theory inspired by Cohen’s forcing to increase the power of a semantics model just as the latter makes it possible to enrich a set theoretical universe. In this way, we can define a new generation of logical relations—that can be introduced modularly using forcing theory—to be used for proving program equivalence for low level languages, concurrent languages or domain specific languages. [48]
### 6. New Results

#### 6.1. Model Transformation

Model transformation and in particular our ATL model transformation language continues playing a key role in our MDE strategy. During 2011 the new results in this area have been:

- The development of an execution algorithm for the lazy execution of ATL transformations [31]. The increasing adoption of Model-Driven Engineering in industrial contexts highlights scalability as a critical limitation of several MDE tools. When these tools are built around model-to-model (M2M) transformations, the efficiency of the transformation engine risks to become a performance bottleneck for the whole MDE environment. This new execution mode solves this problem by providing on-demand execution of model transformations. The computation required to generate a data element of the target model is only triggered once the user requests to read that data element.

- Towards a general semantics for transformation languages. In the mid-term, we would like to be able to achieve interoperability among existing transformation languages (e.g. to create mixed transformations or simply to compare alternative transformation solutions for the same scenario). In this direction, we have worked on a general composition semantics for rule-based transformation languages [32] and an extensive survey of inheritance support in existing languages [33]. Moreover, to make ATL closer to other transformation languages, in special graph transformations, we have developed a new ATL refining mode [43] that allows executing in-place transformations.

- Also relevant, this year AtlanMod has coorganized the main international conference on model transformations [41].

#### 6.2. Model Representation

As part of our work on core techniques for the specification and internal representation of models we would like to highlight:

- Improved model typing. In order to represent metadata more accurately, we have worked on a functional typing system for megamodeling [18]. The basic idea is to consider transformations as functions, and to give them functional types that can facilitate the reasoning on them.

- Virtual EMF [21] for the transparent Composition, Weaving and Linking of Models. Virtual EMF facilitates the global manipulation of an heterogeneous but related set of models by providing the illusion of having a single and unified view of the modeled system.

- EMF Profiles [25], a mechanism to enable users annotate existing models without polluting them (i.e. without directly changing their content).

- Batch scripting support for the retrieval and manipulation of models from model repositories thanks to our new language MoScript [24].

#### 6.3. Model Quality

Our work on model quality defends the idea that there is not a silver bullet technique to verify the quality of models and that a combined approach offers the best trade-off. Lightweight techniques can provide a quick feedback even if it is partial and, when necessary, more complete ones can be utilized to complement the results. In this sense, this year we have developed:

- Lightweight techniques for the verification of ATL transformations [27] and UML Executable models [26].
• Proposed a method for the automatic generation of correct dynamic models for a given domain [11]. Basic operations to modify the information the software needs to store and manipulate about the domain are generated. The generation process ensures that all generated operations are strongly executable.

• Created a brand new Eclipse Lab project aimed at facilitating the verification of any kind of (EMF-based) model with the EMFtoCSP tool (see the tool description in the tools section of this report)

• And applied some of these ideas to the verification and testing of web applications [15]

6.4. Domain Specific Language

There is a growing interest in the research community on the definition of a new generation of language workbenches to facilitate the definition of non-trivial domain specific languages. In this area, we are working on the definition of a set of quality properties that will help language designers to validate if their language is well-adapted to the needs of the language users. This is done by mining repositories containing a corpus of examples of the language.

To begin with we have been reusing the notion of software clones [16] [30], a well-known property in the programming domain and see if it is also meaningful for languages at the model level [29].

6.5. Reverse Engineering

We have continued our work on MoDisco, specially wrt the dissemination activities around the platform [19] [40] and the extension of the tooling support mainly through the spin-off of the EMFFacet Eclipse project, explained in the tools section of this report

6.6. Systems interoperability

MDE can be used as an intermediate representation between two different technical spaces / platforms / tools to facilitate their interoperability. During this year, we have followed this approach in the following results:

• Automation of the interactions with APIs by automatically discovering and expressing them as models thanks to our API2MOL approach [14]

• Bridging the business and the technical domains, developing among others the Portolan tool [42] for the cartography of Information Systems.

• MDE itself can benefit from the work and techniques available in a different technical space. In this context we have combined MDE and constraint programming to see how the combination improves the solution of classical problems like the configuration of a set of components/plug-ins [22]. We have even organized a workshop on the topic of merging MDE and logic programming [20] to better understand how they can benefit each other.
6. New Results

6.1. Models

6.1.1. Using the Last-mile Model as a Distributed Scheme for Available Bandwidth Prediction

Participants: Olivier Beaumont, Lionel Eyraud-Dubois, Young Won.

Several Network Coordinate Systems have been proposed to predict unknown network distances between a large number of Internet nodes by using only a small number of measurements. These systems focus on predicting latency, and they are not adapted to the prediction of available bandwidth. But end-to-end path available bandwidth is an important metric for the performance optimisation in many high throughput distributed applications, such as video streaming and file sharing networks. In [34], we propose to perform available bandwidth prediction with the last-mile model, in which each node is characterised by its incoming and outgoing capacities. This model has been used in several theoretical works for distributed applications. We design decentralised heuristics to compute the capacities of each node so as to minimise the prediction error. We show that our algorithms can achieve a competitive accuracy even with asymmetric and erroneous end-to-end measurement datasets. A comparison with existing models (Vivaldi, Sequoia, PathGuru, DMF) is provided. Simulation results also show that our heuristics can provide good quality predictions even when using a very small number of measurements.

6.1.2. Divisible Load Scheduling

Participants: Olivier Beaumont, Nicolas Bonichon, Lionel Eyraud-Dubois.

Malleable tasks are jobs that can be scheduled with preemptions on a varying number of resources. In [31], we focus on the special case of work-preserving malleable tasks, for which the area of the allocated resources does not depend on the allocation and is equal to the sequential processing time. Moreover, we assume that the number of resources allocated to each task at each time instant is bounded. We consider both the clairvoyant and non-clairvoyant cases, and we focus on minimizing the weighted sum of completion times. In the weighted non-clairvoyant case, we propose an approximation algorithm whose ratio (2) is the same as in the unweighted non-clairvoyant case. In the clairvoyant case, we provide a normal form for the schedule of such malleable tasks, and prove that any valid schedule can be turned into this normal form, based only on the completion times of the tasks. We show that in these normal form schedules, the number of preemptions per task is bounded by 3 on average. At last, we analyze the performance of greedy schedules, and prove that optimal schedules are greedy for a special case of homogeneous instances. We conjecture that there exists an optimal greedy schedule for all instances, which would greatly simplify the study of this problem. Finally, we explore the complexity of the problem restricted to homogeneous instances, which is still open despite its very simple expression. (Join work with Loris Marchal from ENS Lyon)

6.1.3. Modeling and Practical Evaluation of a Service Location Problem in Large Scale Networks

Participants: Olivier Beaumont, Nicolas Bonichon, Hubert Larchevêque.
In [33], we consider a generalization of a classical optimization problem related to server and replica location problems in networks. More precisely, we suppose that a set of users distributed over a network wish to have access to a particular service proposed by a set of providers. The aim is then to distinguish a set of service providers able to offer a sufficient amount of resources in order to satisfy the requests of the clients. Moreover, a quality of service following some requirements in terms of latencies is desirable. A smart repartition of the servers in the network may also ensure good fault tolerance properties. We model this problem as a variant of Bin Packing, namely Bin Packing under Distance Constraint (BPDC) where the goal is to build a minimal number of bins (i.e. to choose a minimal number of servers) so that (i) each client is associated to exactly one server, (ii) the capacity of the server is large enough to satisfy the requests of its clients and (iii) the distance between two clients associated to the same server is minimized. We prove that this problem is hard to approximate even when using resource augmentation techniques; we compare the number of obtained bins when using polynomial time algorithms allowed to build bins of diameter at most $\beta d_{\max}$, for $\beta > 1$, to the optimal number of bins of diameter at most $d_{\max}$. On the one hand, we prove that (i) if $\beta = (2 - \epsilon)$, BPDC is hard to approximate within any constant approximation ratio, for any $\epsilon > 0$; and that (ii) BPDC is hard to approximate at a ratio lower than $\frac{3}{2}$ even if resource augmentation is used. On the other hand, if $\beta = 2$, we propose a polynomial time approximation algorithm for BPDC with approximation ratio $\frac{7}{3}$ in the general case. We show how to turn an approximation algorithm for BPDC into an approximation algorithm for the non-uniform capacitated $K$-center problem and vice-versa. Then, we present a comparison of the quality of results for BPDC in the context of several Internet latency embedding tools such as Sequoia and Vivaldi, using datasets based on PlanetLab latency measurements.

6.1.4. Use of Internet Embedding Tools for Heterogeneous Resources Aggregation

Participants: Olivier Beaumont, Nicolas Bonichon, Philippe Duchon, Hubert Larchevêque.

In [28], we are interested in large scale distributed platforms like BOINC, consisting of heterogeneous resources and using the Internet as underlying communication network. In this context, we study a resource clustering problem, where the goal is to build clusters having at least a given capacity and such that any two participants to the same cluster are not too far from each other. In this context, the distance between two participants corresponds to the latency of a communication between them. Our goal is to provide algorithms with provable approximation ratios. In such large scale networks, it is not realistic to assume that the whole latency matrix (that gives the latency between any two participants) is known, and we need to rely on embedding tools such as Vivaldi or Sequoia. These tools enable to work on compact descriptions and well described metric spaces in which the distance between two points can be obtained directly from a small amount of information available at each node. We present the Bin Covering under Distance Constraint problem (BCDC for short), and propose dedicated algorithms for this problem for each metric space induced by each of the embedding tools. Then, we propose a comparison of these algorithms based on actual latency measures, that enables to decide which algorithm/embedding tool pair offers in practice for realistic datasets the best balancing between distance prediction and approximation ratios for the resource clustering problem.

6.1.5. Broadcasting on Large Scale Heterogeneous Platforms with Connectivity Artifacts under the Bounded Multi-Port Model

Participants: Olivier Beaumont, Nicolas Bonichon, Lionel Eyraud-Dubois, Przemyslaw Uznanski.

In [32], we consider the classical problem of broadcasting a large message at an optimal rate in a large scale distributed network. The main novelty of our approach is that we consider that the set of participating nodes can be split into two parts: "green" nodes that stay in the open-Internet and "red" nodes that lie behind firewalls or NATs. Two red nodes cannot communicate directly, but rather need to use a green node as a gateway for transmitting a message. In this context, we are interested in both maximizing the throughput (i.e. the rate at which nodes receive the message) and minimizing the degree at the participating nodes, i.e. the number of TCP connections they must handle simultaneously. We consider both cyclic and acyclic solutions for the flow graph. In the cyclic case, our main contributions are a closed form formula for the optimal cyclic throughput and the proof that the optimal solution may require arbitrarily large degrees. In the acyclic case, we propose an algorithm to achieve the optimal throughput with low degree. Then, we prove a worst case ratio between
the optimal acyclic and cyclic throughput and show through simulations that this ratio is on average very close to 1, which makes acyclic solutions efficient both in terms of throughput and of number of connections.

6.2. Overlays and distributed algorithms

6.2.1. Locally Fair Graph Exploration Strategies

Participants: David Ilcinkas, Ralf Klasing, Adrian Kosowski.

In [16], we considered the problem of exploring an anonymous undirected graph using an oblivious robot. The studied exploration strategies are designed so that the next edge in the robot’s walk is chosen using only local information, and so that some local equity (fairness) criterion is satisfied for the adjacent undirected edges. Such strategies can be seen as an attempt to derandomize random walks, and are natural counterparts for undirected graphs of the rotor-router model for symmetric directed graphs. The first of the studied strategies, known as Oldest-First (OF), always chooses the neighboring edge for which the most time has elapsed since its last traversal. Unlike in the case of symmetric directed graphs, we show that such a strategy in some cases leads to exponential cover time. We then consider another strategy called Least-Used-First (LUF) which always uses adjacent edges which have been traversed the smallest number of times. We show that any Least-Used-First exploration covers a graph \( G = (V,E) \) of diameter \( D \) within time \( O(D|E|) \), and in the long run traverses all edges of \( G \) with the same frequency.

6.2.2. Black Hole Search in Directed Graphs

Participant: Adrian Kosowski.

In [21] we considered a team of agents which has to explore a graph \( G \) where some nodes can be harmful. Robots are initially located at the so called home base node. The dangerous nodes are the so called black hole nodes, and once a robot enters in one of them, it is destroyed. The goal is to find a strategy in order to explore \( G \) in such a way that the minimum number of robots is wasted. The exploration ends if there is at least one surviving robot which knows all the edges leading to the black holes. As many variations of the problem have been considered so far, the solution and its measure heavily depend on the initial knowledge and the capabilities of the robots. We assume that \( G \) is a directed graph, the robots are associated with unique identifiers, they know the number of nodes \( n \) of \( G \) (or at least an upper bound on \( n \)), and they know the number of edges \( \Delta \) leading to the black holes. Each node is associated with a white board where robots can read and write information in a mutual exclusive way. A recently posed question (Czyzowicz et al. 2009) is whether some number of robots, expressed as a function of parameter Delta only, is sufficient to detect black holes in directed graphs of arbitrarily large order \( n \). We give a positive answer to this question for the synchronous case, i.e., when the robots share a common clock, showing that \( O(\Delta^2 \Delta) \) robots are sufficient to solve the problem. This bound is nearly tight, since it is known that at least \( 2^\Delta \) robots are required for some instances. Quite surprisingly, we also show that unlike in the case of undirected graphs, for the directed version of the problem, synchronization can sometimes make a difference: for \( \Delta = 2 \), in the synchronous case 4 robots are always sufficient, whereas in the asynchronous case at least 5 robots are sometimes required.

6.2.3. Rendezvous for Location-Aware Agents

Participant: Adrian Kosowski.

In [35] we studied rendezvous of two anonymous agents, where each agent knows its own initial position in the environment. Their task is to meet each other as quickly as possible. The time of the rendezvous is measured by the number of synchronous rounds that agents need to use in the worst case in order to meet. In each round, an agent may make a simple move or it may stay motionless. We consider two types of environments, finite or infinite graphs and Euclidean spaces. A simple move traverses a single edge (in a graph) or at most a unit distance (in Euclidean space). The rendezvous consists in visiting by both agents the same point of the environment simultaneously (in the same round). In [35], we propose several asymptotically optimal rendezvous algorithms. In particular, we show that in the line and trees as well as in multi-dimensional Euclidean spaces and grids the agents can rendezvous in time \( O(d) \), where \( d \) is the distance between the initial
positions of the agents. The problem of location-aware rendezvous was studied before in the asynchronous model for Euclidean spaces and multi-dimensional grids, where the emphasis was on the length of the adopted rendezvous trajectory. We point out that, contrary to the asynchronous case, where the cost of rendezvous is dominated by the size of potentially large neighborhoods, the agents are able to meet in all graphs of at most \( n \) nodes in time almost linear in \( d \), namely, \( O(d \log^2 n) \). We also determine an infinite family of graphs in which synchronized rendezvous takes time \( \Omega(d) \).

6.2.4. Boundary Patrolling by Mobile Agents

**Participant:** Adrian Kosowski.

In the boundary patrolling problem, a set of \( k \) mobile agents are placed on the boundary of a simply connected planar object represented by a cycle of unit length. Each agent has its own predefined maximal speed, and is capable of moving around this boundary without exceeding its maximal speed. The agents are required to protect the boundary from an intruder which attempts to penetrate to the interior of the object through a point of the boundary, unknown to the agents. The intruder needs some time interval of length \( \tau \) to accomplish the intrusion. Will the intruder be able to penetrate into the object, or is there an algorithm allowing the agents to move perpetually along the boundary, so that no point of the boundary remains unprotected for a time period \( \tau \)? Such a problem may be solved by designing an algorithm which defines the motion of agents so as to minimize the idle time \( I \), i.e., the longest time interval during which any fixed boundary point remains unvisited by some agent, with the obvious goal of achieving \( I < \tau \). Depending on the type of the environment, this problem is known as either boundary patrolling or fence patrolling in the robotics literature. The most common heuristics adopted in the past include the cyclic strategy, where agents move in one direction around the cycle covering the environment, and the partition strategy, in which the environment is partitioned into sections patrolled separately by individual agents. We have obtained, to our knowledge, the first study of the fundamental problem of boundary patrolling by agents with distinct maximal speeds. In this scenario, we give special attention to the performance of the cyclic strategy and the partition strategy. In \[ 36 \], we propose general bounds and methods for analyzing these strategies, obtaining exact results for cases with 2, 3, and 4 agents. We show that there are cases when the cyclic strategy is optimal, cases when the partition strategy is optimal and, perhaps more surprisingly, novel, alternative methods have to be used to achieve optimality.

6.2.5. Rendezvous in trees

**Participant:** Adrian Kosowski.

In the rendezvous problem in trees, two identical (anonymous) mobile agents start from arbitrary nodes of an unknown tree and have to meet at some node. Agents move in synchronous rounds: in each round an agent can either stay at the current node or move to one of its neighbors. We consider deterministic algorithms for this rendezvous task. In \[ 51 \] we have presented a tight trade-off between the optimal time of completing rendezvous and the size of memory of the agents. For agents with \( k \) memory bits, we show that optimal rendezvous time is \( \Theta(n + n^2/k) \) in \( n \)-node trees. More precisely, if \( k \geq c \log n \), for some constant \( c \), we design agents accomplishing rendezvous in arbitrary trees of unknown size \( n \) in time \( O(n + n^2/k) \), starting with arbitrary delay. We also show that no pair of agents can accomplish rendezvous in time \( o(n + n^2/k) \), even in the class of lines of known length and even with simultaneous start. Finally, we prove that at least logarithmic memory is necessary for rendezvous, even for agents starting simultaneously in a \( n \)-node line.

6.2.6. How many oblivious robots can explore a line

**Participant:** David Ilcinkas.

In \[ 20 \] we consider the problem of exploring an anonymous line by a team of \( k \) identical, oblivious, asynchronous deterministic mobile robots that can view the environment but cannot communicate. We completely characterize sizes of teams of robots capable of exploring a \( n \)-node line. For \( k < n \), exploration by \( k \) robots turns out to be possible, if and only if either \( k = 3 \), or \( k \geq 5 \), or \( k = 4 \) and \( n \) is odd. For all values of \( k \) for which exploration is possible, we give an exploration algorithm. For all others, we prove an impossibility result.
6.2.7. Asynchronous deterministic rendezvous in bounded terrains
Participant: David Ilcinkas.

Two mobile agents (robots) have to meet in an a priori unknown bounded terrain modeled as a polygon, possibly with polygonal obstacles. Robots are modeled as points, and each of them is equipped with a compass. Compasses of robots may be incoherent. Robots construct their routes, but the actual walk of each robot is decided by the adversary that may, e.g., speed up or slow down the robot. In [18], we consider several scenarios, depending on three factors: (1) obstacles in the terrain are present, or not, (2) compasses of both robots agree, or not, (3) robots have or do not have a map of the terrain with their positions marked. The cost of a rendezvous algorithm is the worst-case sum of lengths of the robots’ trajectories until they meet. For each scenario we design a deterministic rendezvous algorithm and analyze its cost. We also prove lower bounds on the cost of any deterministic rendezvous algorithm in each case. For all scenarios these bounds are tight.

6.2.8. On the Power of Waiting when Exploring Public Transportation Systems
Participants: David Ilcinkas, Ahmed Mouhamadou Wade.

We study the problem of exploration by a mobile entity (agent) of a class of dynamic networks, namely the periodically-varying graphs (the PV-graphs, modeling public transportation systems, among others). These are defined by a set of carriers following infinitely their prescribed route along the stations of the network. Flocchini, Mans, and Santoro [58] (ISAAC 2009) studied this problem in the case when the agent must always travel on the carriers and thus cannot wait on a station. They described the necessary and sufficient conditions for the problem to be solvable and proved that the optimal number of steps (and thus of moves) to explore a n-node PV-graph of k carriers and maximal period p is in $\Theta(k \cdot p^2)$ in the general case.

In [46], we study the impact of the ability to wait at the stations. We exhibit the necessary and sufficient conditions for the problem to be solvable in this context, and we prove that waiting at the stations allows the agent to reduce the worst-case optimal number of moves by a multiplicative factor of at least $\Theta(p)$, while the time complexity is reduced to $\Theta(n \cdot p)$. (In any connected PV-graph, we have $n \leq k \cdot p$.) We also show some complementary optimal results in specific cases (same period for all carriers, highly connected PV-graphs). Finally this new ability allows the agent to completely map the PV-graph, in addition to just explore it.

6.2.9. The impact of edge deletions on the number of errors in networks
Participants: Christian Glacet, Nicolas Hanusse, David Ilcinkas.

In [41], we deal with an error model in distributed networks. For a target t, every node is assumed to give an advice, i.e. to point to a neighbor that take closer to the destination. Any node giving a bad advice is called a liar. Starting from a situation without any liar, we study the impact of topology changes on the number of liars.

More precisely, we establish a relationship between the number of liars and the number of distance changes after one edge deletion. Whenever $\ell$ deleted edges are chosen uniformly at random, for any graph with $n$ nodes, $m$ edges and diameter $D$, we prove that the expected number of liars and distance changes is $O(\ell^2 D n m)$ in the resulting graph. The result is tight for $\ell = 1$. For some specific topologies, we give more precise bounds.

6.2.10. Computations in interconnection networks with a shared whiteboard
Participant: Adrian Kosowski.

In [52], we study the computational power of graph-based models of distributed computing in which each node additionally has access to a global whiteboard. A node can read the contents of the whiteboard and, when activated, can write one message of $O(\log n)$ bits on it. A message is only based on the local knowledge of the node and the current content of the whiteboard. When the protocol terminates, each node computes the output based on the final contents of the whiteboard in order to answer some question on the network’s topology. We propose a framework to formally define several scenarios modelling how nodes access the whiteboard, in a synchronous way or not. This extends the work of Becker et al. where nodes were imposed to create their messages only based on their local knowledge (i.e., with the whiteboard empty). We prove that the four
models studied have increasing power of computation: any problem that can be solved in the weakest one can be solved in the second, and so on. Moreover, we exhibit problems that separate models, i.e., that can be solved in one model but not in a weaker one. These problems are related to Maximal Independent Set and detection of cycles. Finally we investigate problems related to connectivity as the construction of spanning- or BFS-tree in our different models.

6.2.11. Network Verification

Participant: Ralf Klasing.

In [27], we address the problem of verifying the accuracy of a map of a network by making as few measurements as possible on the nodes of the network. In the past, this task has been formalized as an optimization problem that, given a graph \( G = (V, E) \), asks for finding a minimum-size subset \( Q \) of vertices of \( G \) such that the information returned by the queries on \( Q \) uniquely identifies \( G \). Previously, two global query models have been studied. In [27], we propose a query model that uses only local knowledge about the network. Quite naturally, we assume that a query at a given node \( q \) returns the associated routing table, namely a set of entries which provides, for each destination node, a corresponding (set of) first-hop node(s) along an underlying shortest path. First, we show that any network of \( n \) nodes needs \( \Omega(\log \log n) \) queries to be verified. Then, we prove that there is no \( o(\log n) \) approximation algorithm for the problem, unless \( P = NP \), even for graphs with diameter 2. On the positive side, we provide an \( O(\log n) \)-approximation algorithm to verify a network of diameter 2, and we give exact polynomial-time algorithms for paths, trees and cycles of even length.

6.3. Compact and distributed data structures

6.3.1. Query optimization in databases

Participants: Nicolas Hanusse, Sofian Maabout.

Datacubes are data structures designed to query optimization in databases. In [42] we provide some algorithmic solutions in a user centric setting: the request time is guaranteed and the amount of memory space is minimized.

6.3.2. Parallel computations of Borders

Participants: Nicolas Hanusse, Sofian Maabout.

Borders are fundamental building blocks in data mining. They are used to find frequent patterns, dependencies between attributes, ... In [43] we provide an algorithm that computes borders with a speedup of \( p \) (under reasonable hypothesis) for \( p \) cores.

6.3.3. Node-disjoint multipath spanners and their relationship with fault-tolerant spanners

Participants: Cyril Gavoille, Quentin Godfroy.

Motivated by multipath routing, we introduce a multi-connected variant of spanners. For that purpose we introduce in [40] the \( p \)-multipath cost between two nodes \( u \) and \( v \) as the minimum weight of a collection of \( p \) internally vertex-disjoint paths between \( u \) and \( v \). Given a weighted graph \( G \), a subgraph \( H \) is a \( p \)-multipath \( s \)-spanner if for all \( u, v \), the \( p \)-multipath cost between \( u \) and \( v \) in \( H \) is at most \( s \) times the \( p \)-multipath cost in \( G \). The \( s \) factor is called the stretch.

Building upon recent results on fault-tolerant spanners, we show how to build \( p \)-multipath spanners of constant stretch and of \( O(n^{1+1/k}) \) edges, for fixed parameters \( p \) and \( k \), \( n \) being the number of nodes of the graph. Such spanners can be constructed by a distributed algorithm running in \( O(k) \) rounds.

Additionally, we give an improved construction for the case \( p = k = 2 \). Our spanner \( H \) has \( O(n^{3/2}) \) edges and the \( p \)-multipath cost in \( H \) between any two node is at most twice the corresponding one in \( G \) plus \( O(W) \), \( W \) being the maximum edge weight.
6.3.4. On approximate distance labels and routing schemes with affine stretch

**Participant:** Cyril Gavoille.

For every integral parameter $k > 1$, given an unweighted graph $G$, we construct in polynomial time in [25], for each vertex $u$, a distance label $L(u)$ of size $O(n^{2/(2k-1)})$. For any $u, v \in G$, given $L(u), L(v)$ we can return in time $O(k)$ an affine approximation $\hat{d}(u, v)$ on the distance $d(u, v)$ between $u$ and $v$ in $G$ such that $d(u, v) \leq \hat{d}(u, v) \leq (2k-2)d(u, v) + 1$. Hence we say that our distance label scheme has affine stretch of $(2k-2)d + 1$. For $k = 2$ our construction is comparable to the $O(n^{5/3})$ size, $2d + 1$ affine stretch of the distance oracle of Pătraşcu and Roditty (FOCS ’10), it incurs a $o(\log n)$ storage overhead while providing the benefits of a distance label.

For any $k > 1$, given a restriction of $o(n^{1+1/(k-1)})$ on the total size of the data structure, our construction provides distance labels with affine stretch of $(2k-2)d + 1$ which is better than the stretch $(2k-1)d$ scheme of Thorup and Zwick (J. ACM ’05).

Our second contribution is a compact routing scheme with poly-logarithmic addresses that provides affine stretch guarantees. With $O(n^{3/(3k-2)})$-bit routing tables we obtain affine stretch of $(4k-6)d + 1$, for any $k > 1$.

Given a restriction of $o(n^{1/(k-1)})$ on the table size, our routing scheme provides affine stretch which is better than the stretch $(4k-5)d$ routing scheme of Thorup and Zwick (SPAA ’01).

6.3.5. Sparse spanners vs. compact routing

**Participant:** Cyril Gavoille.

Routing with multiplicative stretch 3 (which means that the path used by the routing scheme can be up to three times longer than a shortest path) can be done with routing tables of $\Theta(\sqrt{n})$ bits per node. The space lower bound is due to the existence of dense graphs with large girth. Dense graphs can be sparsified to subgraphs, called spanners, with various stretch guarantees. There are spanners with additive stretch guarantees (some even have constant additive stretch) but only very few additive routing schemes are known.

In [39], we give reasons why routing in unweighted graphs with additive stretch is difficult in the form of space lower bounds for general graphs and for planar graphs. We prove that any routing scheme using routing tables of size $M$ bits per node and addresses of poly-logarithmic length has additive stretch $\Omega(\sqrt{n/M})$ for general graphs, and $\Omega(\sqrt{n}/M)$ for planar graphs. Routing with tables of size $O(n^{1/3})$ thus requires a polynomial additive stretch of $\Omega(n^{1/3})$, whereas spanners with average degree $O(n^{1/3})$ and constant additive stretch exist for all graphs. Spanners, however sparse they are, do not tell us how to route. These bounds provide the first separation of sparse spanner problems and compact routing problems.

On the positive side, we give an almost tight upper bound: we present the first non-trivial compact routing scheme with $o(\log^2 n)$-bit addresses, additive stretch $O(n^{1/3})$, and table size $O(n^{1/3})$ bits for planar graphs. Actually, our scheme applies to general graphs, and the above bound of $n^{1/3}$ on the stretch and the routing table size holds for all graphs of linear local tree-width, a class of graphs including bounded-genus graphs and apex-minor-free graphs.

---

3Tilde-big-$O$ notation is similar to big-$O$ notation up to factors poly-logarithmic in $n$. 
6. New Results

6.1. Intrusion Detection

Metamorphic codes:
In [12], we have proposed a advance code obfuscation technique for metamorphic codes (i.e., codes that automatically recode themselves each time they propagate or are distributed). We have shown that the detection of such codes was a problem for classical nowadays static detection tools. In [25], we focus on a new dynamic detection approach which allows to detect variants produced by our metamorphic engine. In addition, our approach can detect unknown malware as long as their behavior approaches that of a known malware. For this, we propose to use a measure of similarity between program behaviors. This measure is obtained by lossless compression of execution traces in terms of system calls.

Intrusion Detection based on an Analysis of the Flow Control:
In [13], intrusion detection mechanisms based on the construct a model of normal behavior of the supervised entity are studied. Such a model is used during the detection phase to raise an alarm when a deviation is observed. This approach allows to detect unknown attacks.

The most common anomaly detection mechanisms at application level consist in detecting a deviation of the control-flow of a program. A popular method to detect such anomaly is the use of application sequences of system calls. However, such methods do not detect mimicry attacks or attacks against the integrity of the system call parameters. To enhance such detection mechanisms, we propose in [27] an approach to detect in the application the corruption of data items that have an influence on the system calls. This approach consists in building automatically a data-oriented behavior model of an application by static analysis of its source code. The proposed approach is illustrated on various examples, and an injection method is experimented to obtain an approximation of the detection coverage of the generated mechanisms.

Most of today’s MAC implementations can be turned into permissive mode, where no enforcement is performed but alerts are raised instead. This behavior is very close to an anomaly IDS except that the system is configured through a MAC policy. MAC implementations such as SELinux and AppArmor come with a default policy including real life and practical rules ready to be used as is or as a basis for a custom policy. In [30], we first propose an extension of an IDS based on information flow control. We address issues concerning programs execution and improve its expressiveness in terms of security policy. This extended model can be configured to reach a wide variety of different security goals. Particularly, it allows for information flow checking based on users and/or programs dependent policy rules. Furthermore, suspicious modification of binary programs can be detected to avoid malware execution. We also propose an algorithm for deriving an AppArmor MAC policy into an information flow policy, and thus get the advantage of having a ready to use policy offering good security. An integration within Android is described in [37].

Flow based Interpretation of Access Control Policies:
In [32], we introduce a formal property characterizing access control policies for which the interpretations of access control as mechanism over objects and as mechanism over information contained into objects are similar. This leads us to define both a flow based interpretation of access control policies and the information flows generated during the executions of a system implementing an access control mechanism. When these two interpretations are not equivalent, we propose to add a mechanism dedicated to illegal information flow detection to the mechanism of access control over objects. Such a mechanism is parameterized by the access control policy and is proved sound and complete. We also briefly describe two real implementations, at two levels of granularity, of our illegal flow detection mechanism: one for the Linux operating system and one for the Java Virtual Machine.

Intrusion Detection based on Invariants:
RRABIDS (Ruby on Rails Anomaly Based Intrusion Detection System) [40] is an application level intrusion detection system for applications implemented with the Ruby on Rails framework. The goal of this intrusion detection system is to detect attacks against data in the context of web applications. This anomaly based IDS focuses on the modeling of the normal application profile using invariants. These invariants are discovered during a learning phase. Then, they are used to instrument the web application at source code level, so that a deviation from the normal profile can be detected at run-time. On simple examples we show how the approach detects well known categories of web attacks that involve a state violation of the application, such as SQL injections. Finally, an assessment phase is performed to evaluate the accuracy of the detection provided by the proposed approach.

**Alert Correlation:**
Alert correlation is a crucial problem for monitoring and securing computer networks. It consists in analyzing the alerts triggered by intrusion detection systems (IDSs) and other security related tools in order to detect complex attack plans, discover false alerts, etc. The huge amounts of alerts raised continuously by IDSs and the impossibility for security operators to efficiently analyze them requires tools for eliminating false and redundant alerts on the one hand and prioritize them according to the detected activities? dangerousness and preferences of the analysts on the other hand. In [35], we describe an architecture that combines AI-based approaches for representing and reasoning with security operators? knowledge and preferences. Moreover, this architecture allows to combines experts? knowledge with machine learning and classifier based tools. This prototype collects the alerts raised by security related tools and analyzes them automatically.

**Trust-Based IDS for the AODV Protocol:**
Routing in ad hoc networks is based on mutual trust between collaborating nodes. Security problems arise when supposedly honest nodes lie deliberately to maximize their profit. In [11], we are interested in detecting misbehaving nodes within the ad hoc routing protocol AODV. We propose and implement a real-time intrusion detection system based on implicit trust relations: a node implementing this system collects its neighbors? routing messages and reasons on them to decide on their honesty. We also evaluate our implementation, and, based on simulations, show that the system we have developed to detect dishonest behavior is efficient.

**Modelization and Simulation of Zombies Behaviours:**
In [26], we study the modelization and simulation of zombie machines for the evaluation of Network Intrusion Detection Systems (NIDS), used to detect botnets. We propose an automatic method to infer zombies behaviours through the analysis of messages exchanged with their masters. Once computed, a model provides a way to generate realistic and manageable traffic, which is mandatory for an NIDS evaluation. We propose to use a Stochastic Mealy Machine to model zombies behaviours, and an active inference algorithm to learn it. With our approach, it is possible to generate a realistic traffic corresponding to the communications of botnets while ensuring its controllability in the context of an NIDS evaluation.

### 6.2. Privacy

Computer privacy is a domain where the education and information of the general public is paramount. In this perspective, through [44] we have participated to the popularization effort in the area, by exposing a survey of accessible computing tools allowing users to better protect their online privacy.

**Formal Privacy Policies and Logical Tools:**
One of the obstacles to the improvement of the privacy level in distributed applications is the lack of expressiveness, usability and enforceability of the associated policies. This new research track aims at designing better privacy policies for complex systems, more adapted to the specific needs of personal data protection regulations and easier to enforce in a distributed fashion. Logical languages, in particular, are considered as interesting candidates because of the reasoning capabilities attached to the formalisms, allowing autonomous peers to perform efficient, privacy-aware planning. [18] is a contribution to the modal logics used to model formal norms, focusing on specific deadline-related temporal notions often encountered in privacy policies. In [39], we propose an ambitious, collaborative research project based on an epistemic view of
the privacy laws and regulations, which should lead to the design of several tools, including policy writing assistants and validation software. [24] is a generic work in the domain of formal policies, where we propose a logical model of various concepts of responsibility in an organizational framework featuring obligation delegation. This kind of framework is intended to model the handling of complex policies in real-life human institutions.

**Privacy in Social Networking Sites:**

Social Networking Sites (SNS), such as Facebook and LinkedIn, have become the established place for keeping contact with old friends and meeting new acquaintances. As a result, a user leaves a big trail of personal information about him and his friends on the SNS, sometimes even without being aware of it. This information can lead to privacy drifts such as damaging his reputation and credibility, security risks (for instance identity theft) and profiling risks. Another research challenge stems from the fact that in the digital world where it is possible to copy the information as often as desired, it is not easy to control how information is disseminated once it is out on the Internet. In an ongoing collaboration [23] with Ai Thanh Ho and Esma Aïmeur (Université de Montréal), we investigate tools that can help user to maintain the sovereignty of their data on the World Wide Web. We also introduce PrivacyMarker, an approach drawing on the concept of provenance and accountability to protect user privacy on SNS. More precisely, it is possible to imagine that by a combination of logs and techniques such as watermarking and traitor-tracing schemes, the dissemination of information can be (at least partially) controlled and that in case of a privacy breach, it is possible to identify which persons are potentially suspect because they have previously accessed this information.

**Geo-privacy:**

A geolocalised system generally belongs to an individual and as such knowing its location reveals the location of its owner, which is a direct threat against his privacy. To protect the privacy of users, a sanitization process, which adds uncertainty to the data and removes some sensible information, can be performed but at the cost of a decrease of utility due to the quality degradation of the data. In a joint work [16] with Marc-Olivier Killijian and Miguel Nunez del Prado (LAAS-CNRS), we describe GEPETO (for GEoPrivacy-Enhancing TOolkit), a flexible open source software which can be used to visualize, sanitize, perform inference attacks and measure the utility of a particular geolocalised dataset. We also introduce a mobility model that we coin as mobility Markov Chain, which can represent in a compact yet precise way the mobility behaviour of an individual. Finally, we describe an algorithm for learning such a structure from the mobility traces of an individual.

Geosocial networks are relatively new compared to the more “traditional” (i.e. non-geolocated) social networking sites such as Facebook or LinkedIn that have been around since more than 6 years, but they are currently growing relatively fast along with the widespread development of other geolocated applications and technologies. In a study [29] done in cooperation with Olivier Heen (Technicolor) and Christophe Potin, we provide a comparative analysis of some existing geosocial networks with respect to privacy in order to (1) highlight some of privacy issues that are raised by the fast development of these system and (2) propose recommendations that could be integrated in the design of these systems to enhance the privacy of their users based on this analysis.

**Privacy in Distributed Systems:**

In a joint work [19] with Anne-Marie Kermarrec and Mohammad Alaggan (team INRIA ASAP), we address the problem of computing the similarity between two users (according to their profiles) while preserving their privacy in a fully decentralized system and for the passive adversary model. First, we introduce a two-party protocol for privately computing a threshold version of the similarity and apply it to well-known similarity measures such as the scalar product and the cosine similarity. The output of this protocol is only one bit of information telling whether or not two users are similar beyond a predetermined threshold. Afterwards, we explore the computation of the exact and threshold similarity within the context of differential privacy, a recent notion developed that provides a strong privacy guarantee that holds independently of the auxiliary knowledge that the adversary might have. More specifically, we design several differentially private variants of the exact and threshold protocols and we also analyze their complexity as well as their impact on the utility.
of the resulting similarity measure. Finally, we provide experimental results validating the effectiveness of the proposed approach on real datasets.

Other ongoing work tackles the problem of computing an aggregation function in a secure and scalable way in a distributed network [42] (joint work with Rachid Guerraoui, Hamza Harkous, Florian Huc and Anne-Marie Kermarrec).

6.3. Accidental and Malicious Faults in Distributed Systems

**Induced Churn to Face Malicious Behaviors:**

In reputation mechanisms, ensuring durable access to feedbacks is a first barrier against simple attacks. To bias the reputation mechanism, an adversary can create and use several distinct identities. In that case, if the reputation mechanism is solely based on statistical measurements, the trustworthiness can be violated. Our contribution is centered around the study of robust mechanisms that can resist such attacks.

Toward this goal, we have first investigating the problem of uniform sampling in large scale open systems in presence of adversarial nodes. Uniform sampling ensures that any individual in a population has the same probability to be selected as sample. Uniform sampling finds its root in many problems such as data collection, dissemination, load balancing, and data-caching.

By relying on the topological properties of structured peer-to-peer systems, it has been shown that it is possible to guarantee with high probability that any node is equally likely to appear in the local view of each other honest node in a number of rounds polynomial in the size of the system. This is achieved by imposing nodes to frequently depart from their position and move to another random position in the system. Indeed, in [15], we have shown that an adversary can very quickly subvert overlays based on distributed hash tables by simply never triggering leave operations. We have also demonstrated that when all nodes (honest and malicious ones) are imposed on a limited lifetime, the system eventually reaches a stationary regime where the ratio of polluted clusters is bounded, independently from the initial amount of corruption in the system.

In unstructured peer-to-peer systems, nodes cannot rely on the topological nature of structured graphs to detect undesirable behaviors. The sampling has to be uniform and ergodic. Informally, this second property guarantees that each received node id infinitely often has a non-null probability to locally appear as a sample. In [21], we determine necessary and sufficient conditions under which uniform and ergodic sampling is achievable in unstructured peer-to-peer systems potentially populated with a large proportion of Byzantine nodes. Strict restrictions are imposed on the number of messages gossiped by malicious nodes during a given period of time and providing each honest node with a very large memory (in the size of the system).

In [38], we consider the problem of targeted attacks in large scale peer-to-peer overlays. These attacks aimed at exhausting key resources of targeted hosts to diminish their capacity to provide or receive services. To defend the system against such attacks, we rely on clustering and implement induced churn to preserve randomness of nodes identifiers so that adversarial predictions are impossible. We propose robust join, leave, merge and split operations to discourage brute force denial of services and pollution attacks.

**Sequence of Consensus Instances:**

To be able to coordinate efficiently the activities of replicas, a significant body of work on replication techniques, group communication services and agreement problems has been done. The Consensus service has been recognized as a fundamental building block for fault-tolerant distributed systems. Many different protocols to implement such a service have been proposed, however, little effort has been placed in evaluating their performance. During her PhD thesis [14], Izabela Moise has presented a protocol designed to solve several consecutive consensus instances in an asynchronous distributed system prone to crash failures and message omissions. The protocol [31] follows the Paxos approach [49], [47] and integrates two different optimizations to reduce the latency of learning a decision value. As one optimization is risky [48], dynamics triggering criterion are defined to check at runtime if the context seems to be favorable or not. The proposed protocol [34] is adaptive as it tries to obtain the best performance gain depending on the current context. Moreover, it guarantees the persistence of all decision values. Our experimentation results [
focus on the impact of the prediction of collisions (i.e., the cases where the use of the risky optimization is counterproductive).

**Transactional Mobile Agent:**
Mobile devices are now equipped with multiple sensors and networking capabilities. They can gather information about their surrounding environment and interact both with nearby nodes, using a dynamic and self-configurable ad-hoc network, and with distant nodes via the Internet. While the concept of mobile agent is appropriate to explore the ad-hoc network and autonomously discover service providers, it is not suitable for the implementation of strong distributed synchronization mechanisms. Moreover, the termination of a task assigned to an agent may be compromised if the persistence of the agent itself is not ensured. In the case of a transactional mobile agent, we identify two services, Availability of the Sources and Atomic Commit, that can be supplied by more powerful entities located in a cloud. In [33], we propose a solution where these two services are provided in a reliable and homogeneous way. To guarantee reliability, the proposed solution relies on a single agreement protocol that orders continuously all the new actions whatever the related transaction and service.
6. New Results

6.1. Network Economics

Participants: Pierre Coucheney, Hai Tran Hoang, Bruno Tuffin, Jean-Marc Vigne.

While pricing telecommunication networks was one of our main activities for the past few years, we are now dealing with the more general topic of network economics. We have tackled it from different sides: i) investigating how QoS or QoE can be related to users’ willingness to pay, ii) investigating the consequences and equilibrium due to competition among providers in different contexts, iii) studying the economic aspect of interdomain relationships, iv) looking at the economics of applications, for example adword auctions for search engines, v) investigating the economics of security in telecommunications, vi) studying the network neutrality issue.

On the first item, in [70], [29], we have studied how utility functions can be related to QoE recent research. Indeed, a logarithmic version of utility usually serves as the standard example due to its simplicity and mathematical tractability. We argue that there are much more (and better) reasons to consider logarithmic utilities as really paradigmatic, at least when it comes to characterizing user experience with specific telecommunication services. We justify this claim and demonstrate that, especially for Voice-over-IP and mobile broadband scenarios, there is increasing evidence that user experience and satisfaction follows logarithmic laws. Finally, we go even one step further and put these results into the broader context of the Weber-Fechner Law, a key principle in psychophysics describing the general relationship between the magnitude of a physical stimulus and its perceived intensity within the human sensory system.

A notable part of our activity has been related to competition among telecommunication providers, mainly within the framework of the ANR CAPTURES project. The goal is to improve most of the pricing models analysis which only deal with a single provider while competition (that is observed in the telecommunication industry) can drive to totally different outcomes. A general view of some of our results is summarized in [67]. A general model of competition in loss networks is described and analyzed in [22] as a two-levels game: at the smallest time scale, users’ demand is split among providers according to the Wardrop principle, depending on the access price and available QoS (depending itself on the level of demand at the provider), and at the largest time scale, providers play a pricing game, trying non-cooperatively to maximize their revenue. A striking result is that this game leads to the same outcome than if providers were cooperatively trying to maximize social welfare: the so-called price of anarchy is equal to one. An additional (higher) level of game is analyzed in [23] (but using another type of negative externality for users, based here on delay), at which providers play on the technologies to implement, based on the infrastructure and licence (if any) costs, anticipating what would be the resulting price war outcome and revenue for given profiles of sets of technologies. This type of study may help a regulator to decide a licence cost, in order to drive the resulting Nash equilibrium to a better point in terms of social or user welfare. A specific situation we have analyzed is the case for a competitive market operated by a Mobile Network Operator (MNO) and a Mobile Virtual Network Operator (MVNO) [46]. The resource that is leased by the MNO to the MVNO is spectrum. MNO and MVNO compete posting subscription prices and the mobile users may choose to subscribe to one operator, or not to subscribe. The scenario is modeled by a three-level game comprising: a bargaining game, which models the spectrum leasing by the MNO; a competition game, which models the price competition between the MNO and the MVNO; and a subscription game, which models the subscription choice by the mobile users, and the outcome of which may be either not to subscribe, to subscribe to the MNO or to subscribe to the MVNO. We assess which conditions lead to an equilibrium where the competition does take place and the amount of the spectrum that should be leased to maximize user or social welfare.
Another important activity is around interdomain issues, with a network like the Internet being made of thousands of autonomous systems. Intermediate domains need some (economic in our case) incentives for forwarding the traffic of other domains. In [33], we have described the problem, provided a state of the art and highlighted the difficulties that must be solved. In [32], we have designed a decentralized algorithm based on double-sided auctions to allocate (and charge) the resource usage.

But network economics is not only about ISPs, it also deals with the application side. In order to make money many service providers base their revenue on advertisement. Search engines for example get revenue thanks to adword auctions, where commercial links are proposed and charged to advertisers as soon as the link is clicked through. Most search engines have chosen (or switched to) a revenue-based ranking and charging scheme instead of a bid-based one. In [53] we investigate the relevance of that scheme when advertisers’ valuation comes from a random distribution, showing that depending on the search engine’s click-through-rate, revenue-based does not always outperform bid-based in terms of revenue to the search engine. But in this adword auction context too, there exist very few works dealing with serach engines in competition for advertisers. We have developed a two-level game where at the largest time scale search engines decide which allocation rule to implement, between revenue-based and bid-based; and at the lowest time-scale advertisers decide how to split their advertising budget between the two search engines, depending on the benefits this will bring to them. The game at the largest time scale is solved using backward induction, the search engines anticipating the reactions of advertisers [54], [52]. We describe the advertisers best strategies and show how to determine, depending on parameters, an equilibrium on the ranking rule strategy for search engines; this may explain Yahoo!’s move to switch from bid-based to revenue-based ranking to follow Google’s strategy.

We similarly have looked at the competition aspects linked to security. We have reviewed the interactions and strategies of attackers and defenders [68]. But we have also looked at the economics of network security, when network users can choose among different security solutions to protect their data, offered by competitive security providers [51]. The interactions among users are modeled as a noncooperative game, with a negative externality coming from the fact that attackers target popular systems to maximize their expected gain.

A new issue we are investigating is the network neutrality debate coming from the increasing asymmetry between Internet Service Providers (ISPs), mainly due to some prominent and resource consuming content providers which are usually connected to a single ISP. We have described and analyzed in [69] the respective arguments of neutrality proponents and opponents, and are currently completing the analysis of several promising game-theoretic models on this issue.

6.2. Dependability and extensions

Participants: Raymond Marie, Gerardo Rubino, Samira Saggadi, Bruno Tuffin.

We maintain a permanent research activity in different domains related to dependability, performability and vulnerability analysis of communication systems. Our focus is on evaluation techniques using both the Monte Carlo and the Quasi-Monte Carlo approaches. Monte Carlo (and Quasi-Monte Carlo) methods often represent the only available tool to solve complex problems in the area, and rare event simulation requires a special attention, in order to be able to efficiently analyze the model, that is, to be able to use good estimators having, in particular, a sufficiently small relative variance. Novel results in simulation can be decomposed into two subsets: results on rare event simulation, and those on Randomized Quasi-Monte Carlo methods.

The effectiveness of randomized quasi-Monte Carlo (RQMC) techniques is examined in [26] to estimate the integrals that express the discrete choice probabilities in a mixed logit model, for which no closed form formula is available. These models are used extensively in travel behavior research. We consider popular RQMC constructions, but our main emphasis is on randomly-shifted lattice rules, for which we study how to select the parameters as a function of the considered class of integrands. We compare the effectiveness of all these methods and of standard Monte Carlo (MC) to reduce both the variance and the bias when estimating the log-likelihood function at a given parameter value.
The main part of our activity in this simulation area in 2011 has been on rare event simulation though. The two major simulation families or rare event estimations are importance sampling and splitting. In [63], we have provided a recent view of those methods, while in [64] we have overviewed how the zero-variance importance sampling can be approximated in classical reliability problems.

The problem of estimating the probability that a given set of nodes is connected in a graph (or network) where each link is failed with a given probability has received a lot of attention from us in 2011. We have proposed in [21] a new Monte Carlo method, based on dynamic Importance Sampling. The method generates the link states one by one, using a sampling strategy that approximates an ideal zero-variance importance sampling scheme. The approximation is based on minimal cuts in subgraphs. In an asymptotic rare-event regime where failure probability becomes very small, we prove that the relative error of our estimator remains bounded, and even converges to 0 under additional conditions, when the unreliability of individual links converges to 0. The empirical performance of the new sampling scheme is illustrated by examples. The method is even sped up in [50] by applying series-parallel reductions at each step of the algorithm.

The same problem is also analyzed in [15] by novel method that exploits a generalized splitting (GS) algorithm. We show that the proposed GS algorithm can accurately estimate extremely small unreliabilities and we exhibit large examples where it performs much better than existing approaches. Remarkably, it is also flexible enough to dispense with the frequently made assumption of independent edge failures. In [17], another splitting approach is explored for the same problem, with very good results. It consists of a standard splitting procedure applied to the so-called Creation Process that can be associated with the initial static model. The paper discusses both a method for splitting this process, and an experimental analysis of the covering of the resulting estimator, showing its good behavior on different classes of test problems. Last, in [16], always for the same static reliability problem, we proposed a new procedure belonging to the RVR family (Recursive Variance Reduction) where a new estimator based both in computed minpaths and mincuts of the graph, together with series-parallel reductions, allows to obtain very good accuracy in many rare events situations.

Finally, a versatile Monte Carlo method for estimating multidimensional integrals, with applications to rare-event probability estimation is presented in [39], [75]. The method uses two distinct and popular Monte Carlo simulation techniques, namely Markov chain Monte Carlo (MCMC) and Importance Sampling, combined into a single algorithm. We show that for some illustrative and applied numerical examples the proposed Markov Chain Importance Sampling algorithm performs better than methods based solely on Importance Sampling or solely on MCMC.

Concerning the risk on spares for life-time maintenance purposes which is due to uncertainties on the mean up time, an extended version of a presentation made in 2010 has been published in [24].

6.3. Performance evaluation

Participants: Laura Aspirot, Raymond Marie, Gerardo Rubino, Bruno Sericola.

An important problem arising when dimensioning a P2P system is to understand the evolution of the peers’ population with time. The number of units being usually large, the standard stochastic models used to represent this kind of system (e.g. a Markovian stochastic process) are difficult to use in practice. Instead, it is popular today to move to deterministic continuous-state (fluid) models whose dynamics is governed by differential equations. It is then of interest to analyze the conditions under which the latter are the limit, in some sense, of the former. We started to develop this program in [36] by focusing on some popular models of P2P systems, and analyzed when and how the deterministic model is the limit of the stochastic one when the number of peers goes to infinity.

In [60], we continued to explore the concept of power of a queueing model proposed by Kleinrock in the 80s. Kleinrock’s idea was to build a metric combining two “competing” ones, the mean throughput and the mean response time, for the system in equilibrium. The power is defined as the ratio of normalized versions of those metrics. We discuss different ways of adapting this concept to more general queueing systems such as queueing networks. In this research line, [60] opens the way for a definition of efficiency, which is currently analyzed in the team.
In [30], we expose a clear methodology to analyze maximum level and hitting probabilities in a Markov driven fluid queue for various initial condition scenarios and in both cases of infinite and finite buffers. Step by step we build up our argument that finally leads to matrix differential Riccati equations for which there exists a unique solution. The power of the methodology resides in the simple probabilistic argument used that permits to obtain analytic solutions. We illustrate our results by a comprehensive fluid model that we solve exactly.

In [65], we analyze the transient behavior of a fluid queue driven by a general ergodic birth and death process using spectral theory in the Laplace transform domain. These results are applied to the stationary regime and to the busy period analysis of that fluid queue.

Finally, in [71] we present a global view of the performance evaluation area in computer and communication systems, an extended and reviewed version of a talk given in 2010.

### 6.4. Quantitative aspects of distributed systems

**Participants:** Bruno Sericola, Romaric Ludinard.

This work is a collaboration with the Inria team-project Asap. We proposed in [20] a fully decentralized algorithm to provide each of the nodes of a distributed system with a value reflecting its connectivity quality. Comparing these values between nodes, enables to have a local approximation of a global characteristic of the graph. Our algorithm relies on an anonymous probe visiting the network in a unbiased random fashion. Each node records the time elapsed between visits of the probe which is called the return time of the random walk. Computing the standard deviation of such return times enables to approximate the conductance of the graph. Typically, this information may be used by nodes to assess their position, and therefore the fact that they are critical, in a graph exhibiting low conductance.

We continue our collaboration with the Inria team-projects Adept and Ipso. It is well-known that peer-to-peer overlays networks can only survive Byzantine attacks if malicious nodes are not able to predict what will be the topology of the network for a given sequence of join and leave operations. In [13] and in [35], we investigate adversarial strategies by following specific games. Our analysis demonstrates first that an adversary can very quickly subvert DHT-based overlays by simply never triggering leave operations. We then show that when all nodes (honest and malicious ones) are imposed on a limited lifetime, the system eventually reaches a stationary regime where the ratio of polluted clusters is bounded, independently from the initial amount of corruption in the system. These results have been obtained using Markov models. In [14] and [34], we consider the behavior of a stochastic system composed of several identically distributed, but non independent, discrete-time absorbing Markov chains competing at each instant for a transition. The competition consists in determining at each instant, using a given probability distribution, the only Markov chain allowed to make a transition. We analyze the first time at which one of the Markov chains reaches its absorbing state. We obtain its distribution and its expectation and we propose an algorithm to compute these quantities. We also exhibit the asymptotic behavior of the system when the number of Markov chains goes to infinity. Actually, this problem comes from the analysis of large-scale distributed systems and we show how our results apply to this domain.

### 6.5. QoE (Quality of Experience)

**Participants:** Sebastián Basterrech, Yassine Hadjadj-Aoul, Sofiene Jelassi, Adlen Ksentini, Gerardo Rubino, Kamal Singh, César Viho.

We continue the development of the PSQA technology (Pseudo-Subjective Quality Assessment) in the area of Quality of Experience (QoE). PSQA is today a stable technology allowing to build measuring modules capable of quantifying the quality of a video or an audio sequence, as perceived by the user, when received through an IP network. It provides an accurate and efficiently computed evaluation of quality. Accuracy means that PSQA gives values close to those than can be obtained from a panel of human observers, under a controlled subjective testing experiment, following an appropriate standard (which depends on the type of sequence or application). Efficiency means that our measuring tool can work in real time, if necessary. Observe that perceived quality is the main component of QoE. PSQA works by analyzing the networking environment of the communication
and some the technical characteristics of the latter. It works without any need to the original sequence (as such, it belongs to the family of no-reference techniques).

It must be pointed out that a PSQA measuring or monitoring module is network dependent and application dependent. Basically, for each specific networking technology, application, service, the module must be built from scratch. But once built, it works automatically and very efficiently, allowing if necessary to use it in real time.

At the heart of the PSQA approach there is the statistical learning process necessary to develop measuring modules. So far we have been using Random Neural Networks (RNNs) as our learning tool (see [82] for a general description), but recently, we have started to explore other approaches. For instance, in the last ten years a new computational paradigm was presented under the name of Reservoir Computing (RC) [78] covering the main limitations in training time for recurrent neural networks while introducing no significant disadvantages. Two RC models have been developed independently and simultaneously under the name of Liquid State Machine (LSM) [81] and Echo State Networks (ESN) [78] and constitute today one of the basic paradigms for Recurrent Neural Networks modeling [79]. The main characteristic of the RC model is that it separates two parts: a static sub-structure called reservoir which involves the use of cycles in order to provide dynamic memory in the network, and a parametric part composed of a function such as a multiple linear regression or a classical single layer network. The reservoir can be seen as a dynamical system that expand the input stream in a space of states. The learning part of the model is the parametric one. In a recent collaboration with the Applied Computational Intelligence Research Unit, Artificial Neural Networks Group of the University of the West of Scotland during the first half of the year, we developed an algorithm based on a combination of topology preserving maps such as the Self-Organising Map [80] and the Scale Invariant Map [77] to improve the performance of RC models. The obtained results are presented in two papers: [37] and [38].

In [42] we developed a PSQA version for evaluating the perceived quality in the context of SVC video coding. The tool is based on the use of the RNN model. The main difficulties in defining this tool is regarding the relation between the SVC layers, since the enhanced layers require the information of the base layer in order to be decoded.

In [61], we developed a tool for evaluating the perceived quality of an application distributing streamed video using HTTP (and thus, TCP). The difficulties here are focused around the possible playout interruptions and the quality variations due to the use of adaptive bitrate techniques. Our procedure belongs to the no-reference family of learning ones, and it is also based on the use of the RNN tool.

In [41] we compared PSQA used for the video evaluation to other no-reference tools as well as two objective evaluation tools. We showed that PSQA outperforms the majority of the other tools, in terms of high correlation with human evaluation. This version will be used as the main metric for evaluating the QoE in the future internet architecture proposed by the FP7 Alicante project.

We have also been developing single-ended parametric-model speech-quality assessors of VoIP conversations over future networks. To do that, a careful identification and accurate characterization of quality-degrading factors over next-generation networks has been done. The recent progress and challenges for accurate assessment of voice quality over evolving VoIP systems has been detailed in the survey paper [19]. In [18], we study the perceived effects of packet loss processes, which are the principal source of quality degradation over IP networks. In reality, the perceived effect of a given packet loss process is highly related to the distribution of missing packets. Basically, the higher the burstiness of packet loss processes, the greater the perceived quality degradation. Recently, several assessors of speech quality sensitive to packet loss burstiness have been proposed in the literature. A comprehensive comparison study of bursty-packet-loss-aware artificial assessors has been conducted in [18]. An extended and more elaborated version has been published in [47]. Moreover, novel artificial quality assessors that consider transient loss of connectivity incurred by mobile users over mobile transport system have been developed. A paper describing our developed tools and performance results is under preparation. Recently, we started to work on new analytical models of packet losses and delays of packet-based voice conversations over wireless ad-hoc networks. The developed models will be used to design specialized artificial quality assessors of multimedia services over wireless ad-hoc networks. Moreover,
we are working on the enhancement of a voice quality assessor version of PSQA, by considering the features of removed speech signals.

6.6. Wireless networks

Participants: Nizar Bouabdallah, Yassine Hadjadj-Aoul, Adlen Ksentini, Raymond Marie, Bruno Sericola, César Viho.

We continue working on wireless networking. The focus mainly concerns wireless distribution of audio and video, which require strict Quality of Service (QoS) support.

In [27], we investigated the main challenges when the goal is to constitute an efficient Radio Resource Management (RRM) framework. The existing solutions of RRM were classified based on the considered decision-making technique. Moreover, we investigated in [28] how QoE can help for designing efficient RRM for wireless networks. A resource allocation mechanism is proposed in [62]. In [59] we proposed a novel network selection mechanism for heterogeneous wireless networks that take QoE into consideration for decision-making. The main idea is to use QoE of ongoing users in candidate networks as an indicator to select the best network for connection. Besides, in order to provide efficient interworking between the different access players, we first defined some issues related to the interworking operation between the satellite and terrestrial domains. We suggested some solutions and discussed their potential in [31].

We also investigated in [74] solutions that ensure the scalability of mobile networks, which are facing a rapid increase of data traffic. We devise methods that enable User Equipments, both in idle and active mode and while being on the move, to always have optimal Packet Data Network (PDN) connections (i.e., IP addresses) in such decentralized networks. We demonstrated the effectiveness of such approach in current mobile and wireless networks. In these systems, minimizing energy consumption is becoming more and more crucial. In [66], we devised a PID (Proportional Integral Derivative)-based controller permitting to reduce the amount of wasted energy by determining an optimal schedule between the sleep and wakeup periods of the wireless interface during the VoIP communication while keeping the perceived quality at the desired level.

Based on our previous research on proactive routing for wireless ad-hoc networks, we have published a book chapter in [73], focusing on modeling the resilience of routing information for ad hoc networks where topology information is uncertain.

We continue our collaboration with the Inria team-projects Pops (Lille), D-Net (Lyon), Reso (Lyon) and the NPA (Networks and Performance Analysis) research group of LIP6 (Paris) on fast self-stabilization in large scale wireless networks. In these systems, distributed self-organization is more convenient than centralized planification. Self-stabilization protocols are a useful technique to provide self-organization but their stabilizing time is related to the size of the network. In [25], we show that a clustering algorithm, known for its good robustness properties, is actually self-stabilizing. We propose several enhancements to the scheme in order to reduce the stabilization time and thus improved the stability in a dynamic environment. The key technique to these enhancements is a localized self-stabilizing algorithm for directed acyclic graph construction. We provide extensive studies (both theoretical and experimental) that show that our approach enables efficient yet adaptive clustering in wireless multihop networks.

6.7. Sensor networks

Participants: Nizar Bouabdallah, Sofiane Moad.

Wireless Sensor Networks (WSNs) are composed of tiny sensor nodes, which are capable of sensing and processing data from inaccessible environments and communicating them to the end-user for further analysis. WSNs are characterized by the limited capacity of their sensor node batteries, making energy efficiency a critical issue. Once a WSN is deployed, sensor nodes must self-organize and live as long as possible, based only on their initial energy stores. Consequently, techniques minimizing energy consumption are required to improve network lifetime. Our research on WSNs [72] revolves around two main directions: 1) clustering, and 2) radio diversity. Regarding clustering, we first developed a Connectivity Degree-Based Energy Efficient
Networks, Systems and Services, Distributed Computing - New Results - Project-Team DIONYSOS

Clustering Protocol for WSNs (CDEEC) \[55\], resulting in better topology management and decreased energy consumption compared to the well-known clustering protocol HEED. Then, we integrated a compression mechanism within a cluster-based architecture to develop a Compression Cluster-based scheme in a Spatial Correlated Region protocol (CC_SCR) \[56\], with the goal of further decreasing the energy consumption. In the direction of radio diversity, we first proposed the WETX metric \[58\] which uses a minimum-energy radio while routing, then we proposed the BL metric \[57\], on top of WETX, which allows energy-balancing inside a network in order to further extend its lifetime. The validation of our contributions was carried out with analytical analysis, and simulation using TOSSIM.

6.8. Scalable Video Coding (SVC) transmission over IP and Broadcast networks

Participants: Majd Ghareeb, Adlen Ksentini, César Viho, Yassine Hadjadj-Aoul.

One of the multimedia market trends is audiovisual service (TV or VoD) anywhere, at any time. To support such service, a Video Service Provider has to manage, store, and distribute content towards multiple kinds and scales of terminals, and over different and transient access technologies to reach the end user. To solve such issues, video scalability seems to be the most relevant solution. It encodes the video in multiple separated layers, which enable a large number of users with heterogeneous capability to view any desired video stream, at anytime, and from anywhere. One of the most well known scalable standards is the Scalable Video Coding (SVC) extension of H.264/MPEG-4 AVC video compression. Our researches in this topic are related to how to optimize and enhance SVC transmission over IP and broadcast networks.

With the aim at keeping a high perceived video quality using SVC, MultiPath Video Streaming (MPVS) over Video Distribution Network (VDN) comes as a promising solution to overcome the limitations of the classical single path and IP-level video streaming approaches. In \[45\] and \[43\] we proposed different approaches that couple the three SVC scalability modes (Spatial, Temporal, SNR), with the path diversity provided by VDN. Our method adapts to both the heterogeneity of end-users using the scalable video coding as well as to network bandwidth fluctuations by observing the changes of the available bandwidth over the multiple overlay paths, and updating the streaming strategy accordingly. In \[44\] we enhanced the precedent solutions by using the PSQA-SVC version \[42\] in order to get the end-user feedback in terms of QoE, which helps adapting the streaming strategy. In \[48\] we designed a new protocol optimizing the energy consumption when transmitting video streams. We propose to exploit the SVC coding to adapt dynamically the received video quality to the instantaneous wireless nodes’ characteristics. This is achieved through determining the number of the transmitted/received enhancements layers of an SVC video based on the wireless node context.

In \[49\] we proposed to support SVC over DVBT2 networks, by associating the layering architecture of both technologies in order to tackle users mobility. This association allows mobile receivers with good physical channels to decode all the SVC layers and benefit from high video quality. Meanwhile, users with poor channel conditions can at least decode the base layer and benefit from acceptable video quality. Further, we introduced a novel QoE-based adaptive mechanism for SVC layers decoding. The proposed approach selects dynamically the number of layers to decode, at the receiver side, so as to maximize the users’ perceived quality. Thus, no feedbacks or signaling messages are needed to implement the proposed algorithm. This makes it compliant with unidirectional technologies such as DVB-T2.
6. New Results

6.1. Fundamental results and algorithms: distributed planning

**Participants:** Eric Fabre, Loig Jézéquel.

A planning problem consists in organizing some actions in order to reach an objective. Formally, this is equivalent to finding a path from an initial state to a goal/marked state in a huge automaton. The latter is specified by a collection of resources, that may be available or not (which defines a state), and actions that consume and produce resources (which defines a transition). In the case of optimal planning, actions have a cost, and the objective is to find a path of minimal cost to the goal.

Our interest in this problem is threefold. First, it is naturally an instance of a concurrent system, given that actions have local effects on resources. Secondly, it is a weak form of an optimal control problem for a concurrent/distributed system. Finally, we are interested in distributed solutions to such problems, which is an active topic in the planning community under the name of “factored planning.”

Our previous contributions to the domain was the first optimal factored planning algorithm [61]. The main idea is to represent a planning problem as a network of interacting weighted automata, the objective being to jointly drive all of them to a target state, while minimizing the cost of their joint trajectory. We have developed and tested [68] a distributed algorithm to solve this problem, based on a weighted automata calculus, and that takes the shape of a message passing procedure. Components perform local computations, exchange messages with their neighbors, in an asynchronous manner, and the procedure converges to the path that each component should follow. The optimal global plan is thus given as a tuple of (compatible) local plans, i.e. a partial order of actions.

In 2011, we have extended this framework in two directions. In terms of modelling, first. In most planning problems, some actions consume/produce resources, but also are enabled by the presence of other resources, that they only read but not consume. We have proposed to model this feature under the form of networks of automata with read arcs. Interactions then take the form of synchronous actions, as previously, but also the form of readings: a component may only be allowed to fire some local transition if another component is in a specific state. Our distributed planning approach has then been extended to this new model of distributed systems [38].

The second improvement is algorithmic. So far, our distributed optimal planning algorithm computes all possible distributed plans, in a factored form. This contrasts with the philosophy of planning algorithms, that look for one plan only, and organize the computations to quickly reach the best plan. In other words, most planning algorithms are based on a common ground known as the A-star algorithm, a depth-first search procedure in a graph, guided by a heuristic function that estimates the remaining cost to reach the goal. We have developed a truly distributed version of this algorithm, to perform a search on a product graph. Each component runs an A-star procedure to find a path to its goal, taking into account the costs of its neighbors in order to guarantee that all components converge to local plans that are compatible and jointly optimal.

6.2. Fundamental results and algorithms: communication with messages and scenarios

**Participants:** Loïc Hélouët, Rouwaida Abdallah, Claude Jard, Blaise Genest.

In this paragraph, we collect our fundamental results regarding the models and algorithms we use for communicating systems, and in particular, scenarios.
A major challenge with models communicating with messages (e.g.: scenarios) is to exhibit good classes of models allowing users to specify easily complex distributed systems while preserving the decidability of some key problems, such as diagnosis, equality and intersection. Furthermore, when these problems are decidable for the designed models, the second challenge is to design algorithms to keep the complexity low enough to allow implementation in real cases.

This year, we have considered analysis for a timed extension of scenarios called Time-constrained MSCs and implementation techniques that take scenarios as an input model and output an equivalent distributed implementation.

The first part of our work is the study of Time-Constrained MSC graphs (TC-MSGs for short). Time-constrained MSCs (TC-MSCs) are simply MSCs decorated with constraints on the respective occurrence dates of events. The semantics of a TC-MSC $T$ is a dated MSC, that is a MSC where events are associated with an occurrence date. For a given TC-MSC, there can be an infinite set $L(T)$ of dated MSCs satisfying its constraints. Note however that some time-constraints in a TC-MSC may not be satisfiable, and hence $L(T)$ can simply be empty. TC-MSCs can be extended by composition mechanisms such as TC-MSC graphs. TC-MSC graphs are simply automata labeled by TC-MSC. Each path $\rho$ of a TC-MSC $G$ is associated with a TC-MSC $T_\rho$ obtained by concatenation of TC-MSC along $\rho$. The language $L(G) = \bigcup_{\rho \text{ path of } G} L(T_\rho)$ of a TC-MSC Graph is then the union of all dated MSCs associated to paths of $G$. Because of inconsistent timing constraints, some path may have no possible realization (i.e $L(T_\rho = \emptyset)$). One can even design a MSC Graph $G$ such that $L(G) = \emptyset$ - such TC-MSC graph is clearly inconsistent-. It has been shown [64] that checking whether $L(G) = \emptyset$ is an undecidable problem in general, but can be decided for the restricted subclass of regular TC-MSC graphs (that have the expressive power of event-count timed automata). We have proposed two restrictions allowing for the decision of emptiness. The first one is $K$-drift boundedness, which imposes for a fixed integer $K$ that for every $T_\rho$ there exists one dated realization such that for every pair of events $e, f$ appearing in the same transition of $G$, the dates of $e$ and $f$ differ by at most $K$. We have shown that $K$-drift boundedness is decidable in a symbolic and efficient way, and that for $K$-drift bounded TC-MSC graphs, emptiness is decidable [52]. This extends decidability results beyond regular specifications. The second restriction is $K$-non-zenoness, which imposes that for a fixed $K$, for every path $\rho$ of $G$, there exists one realization such that at every date $d$, at most $K$ events occur between date $d$ and $d+1$. When a TC-MSC graph is $A$-drift bounded and $B$-non-zeno, then $L(G)$ has a regular set of representants, which opens the way for more involved model-checking applications [51].

The second part of our work is the study of realistic implementation of scenarios. The main idea is to propose distributed implementation (communicating state machines) of High-level MSCs that do not contain deadlocks, and behave exactly as the original specification. It is well known that a simple projection of a HMSC on each of its process to obtain communicating finite state machines results in an implementation with more behaviors than the original specification. An implementation of a HMSC $H$ is considered as consistent if and only if is exhibits the same prefix closed set of behaviors as $H$. We have studied how such projection with additional local controllers allows the distributed synthesized behavior to remain consistent with the original specification. This work has been implemented in our scenario prototype (see the Software section). As usually for scenarios, the synthesis algorithm works for a particular syntactic class of scenarios, namely the class of local HMSCs. Roughly speaking, in local HMSC, a decision to behave according to a scenario or another is always taken by a single participant. The deciding process need not be the same at each choice. This class is a sensible restriction of HMSCs, as distributed choices can not be implemented without additional synchronization among processes [53].

Last, we have extended existing results on diagnosis from scenarios [15]. We have shown that when a distributed implementation is instrumented with software probes that publish their observations while the system is running, and when the system is modeled as a High-level MSC, then diagnosis can be expressed as a new HMSC the executions of which are all explanations of the observation. The construction of diagnosis can be performed offline or online, and we have considered the conditions under which online diagnosis can run with finite memory.
6.3. Fundamental results and algorithms: timed models

Participants: Claude Jard, Aurore Junier, Akshay Sundararaman.

Our works in that subject concern Time Petri Nets (TPNs) and Network Calculus. With TPNs, we are particularly interested in symbolic unfoldings (extended with parameters). Possible applications are supervision of distributed timed systems [8] and testing of concurrent systems (work done in collaboration with Stefan Haar, INRIA-LSV in Cachan).

The article [26] was made during the internship of the master degree of Aurore Junier under the supervision of Anne Bouillard (ENS Paris). It uses a (min, plus)-algebra to define a worst-case delay bound for networks where flows have fixed priorities.

After that, we studied a well-known problem: detection of congestion and failure in networks. The idea was to find an efficient and deterministic method that is very reactive and takes little memory space. Such a method does not exists for now and is an important issue for Alcatel-Lucent. We achieved a solution that solves this problem based on the analysis of flows behaviour. This work is part of the work done within the Alcatel-Lucent-Inria joint lab and a patent is being established.

We are also studying the way buffers of routers can increase. The objective is to find a method that can detect if sizes of buffers can dangerously increase on a defined topology. We start by looking at Link State Advertisements (LSA) in the OSPF protocol. We represent the topology and a part of the protocol by a Time Petri Net and try to infer parameters ensuring stability.

6.4. Fundamental results and algorithms: dynamic epistemic logic

Participants: Guillaume Aucher, François Schwarzentruber.

Dynamic Epistemic Logic (DEL) deals with the representation and the study of knowledge and belief change in a multi-agent setting. The core representative task of this logical framework can be split up in three parts: 1/ the initial global state of the distributed system, 2/ an event occurring in this system, 3/ a product update taking as argument these two representations and yielding a new representation of the new global state of the distributed system. Therefore, we can express uniformly within the DEL framework epistemic statements about:

(i) what is true about an initial state
(ii) what is true about an event occurring in this initial state
(iii) what is true about the resulting state after the event has occurred.

We axiomatized within the DEL framework what we can infer about (iii) given (i) and (ii), what we can infer about (ii) given (i) and (iii), and what we can infer about (i) given (ii) and (iii). Given three logical formulas \( \phi \), \( \phi' \) and \( \phi'' \) describing respectively (i), (ii) and (iii), we also showed how to build three formulas that capture respectively all the information which can be inferred about (iii) from \( \phi \) and \( \phi' \), all the information which can be inferred about (ii) from \( \phi \) and \( \phi'' \), and all the information which can be inferred about (i) from \( \phi' \) and \( \phi'' \).

We showed how our results extend to other modal logics than the minimal modal logic \( K \). These results are to appear in [9] and [10]. In [19], we also provided a tableau method deciding whether such inferences are valid. We implemented it in LOTRECscheme and showed that this decision problem is NEXPTIME-complete. This work contributes to the proof theory and the study of the computational complexity of DEL which have rather been neglected so far.

Application to fault localization in IMS network (see the UNIVERSELF project) has started. The various agents involved in an IMS network (clients, assistance, administrators...) have a partial view of the network and so need to communicate their partial knowledge of the network to each other in order to localize the fault in the network (each communication having possibly a different cost). One of the main problems is to determine which communication should occur and which agent should be queried so that the fault is eventually localized. This problem can naturally be expressed in the DEL framework. We have shown how the initial state of an IMS network representing the knowledge of each agent can be represented by a particular kind of epistemic
model (i) and how the desired state where the fault is localized can be expressed by a logical formula (iii). The problem amounts to determining which communication or sequence of communications should occur (ii) so that one passes from the initial epistemic model (i) to another epistemic model where the fault is localized (iii), and also to determine if such a communication or sequence of communications is possible. We have focused so far on the case of a single communication, but we plan to extend it to a sequence of communications. Further theoretical work still needs to be done to address the issue of sequential communication.

In parallel to this work, we also axiomatized different notions of knowledge and belief which are defined by means of a ‘sphere’ semantics. This work is the result of an invited contribution and is to appear in [48].

6.5. Fundamental results and algorithms: statistical model checking

Participants: Sean Sedwards, Cyrille Jégourel, Axel Legay.

Our work on statistical model checking (SMC) avoids an explicit representation of the state space by building a statistical model of the executions of a system and giving results within confidence bounds. The key challenges of this approach are to reduce the length (simulation steps and cpu time) and number of simulation traces necessary to achieve a result with given confidence. Rare properties pose a particular problem in this respect, since they are not only difficult to observe but their probability is difficult to bound. A further goal is to make a tool where the choice of modelling language and logic are flexible.

We have developed the prototype of a compact, modular and efficient SMC platform which we have named PLASMA (PLatform for Statistical Model checking Algorithms). PLASMA incorporates an efficient discrete event simulation algorithm and features an importance sampling engine that can reduce the necessary number of simulation runs when properties are rare. We have found that PLASMA performs significantly better than PRISM (the de facto reference probabilistic model checker) when used in a similar mode: PLASMA’s simulation algorithm scales with a lower order and can handle much larger models. When using importance sampling, PLASMA’s performance with rare properties is even better.

6.6. Fundamental results and algorithms: quantitative model checking and quantitative specification Theories

Participants: Uli Fahrenberg, Axel Legay.

Model checking of systems deals with the question whether a given model of a computer system satisfies the properties one might want to require of it. This is a well-established and successful approach to formal verification of safety-critical computer systems.

When the models of the systems contain quantitative information, the model checking problem becomes complicated by the fact that in most cases, quantitative properties of the systems do not need to be satisfied exactly. Indeed, the model or the properties might be subject to measurement error, or probabilistic information might only be an approximation. In this case, it is of little use to know whether or not a model satisfies a specification precisely; what is needed instead is a notion of satisfaction distance: a measure which can assess to which extent a quantitative model satisfies a quantitative specification.

In other words, what is needed is a notion of satisfaction which is robust in the sense that small deviations in the model or the specification only lead to small changes in the outcome of the model checking question. We have published work on such distances in the papers [37], [34].

For more elaborate reasoning about distributed systems or systems-of-systems, an important role is played by specification theories. Such systems are often far too complex to reason about, or model-check, as a whole, and additionally they might be composed of a large number of components which are implemented by different vendors. Hence one needs methods for compositional reasoning, which allow to infer properties of a system from properties of its components, and for incremental design, which allow to synthesize and refine specifications in a step-wise manner.
Such specification theories are by now well-established e.g. in the incarnations of interface theories and modal transition systems. Additionally to defining a formalism for describing and model-checking specifications, they provide notions of refinement of specifications, logical conjunction of specifications, and structural composition and quotient.

When the models and specifications contain quantitative information, all the above notions need to be made robust. One needs to introduce a quantitative version of refinement, and the operations on specifications need to be continuous with respect to refinement distance: compositions of specifications with small refinement distance need themselves to have small refinement distance. We have published work on these issues in the papers [21], [35]; additionally, two other papers within this research area are currently under submission.

6.7. Specific studies: Web services orchestrations

Participants: Ajay Kattepur, Albert Benveniste, Claude Jard.

Web services orchestrations and choreographies refer to the composition of several Web services to perform a co-ordinated, typically more complex task. We decided to base our study on a simple and clean formalism for WS orchestrations, namely the Orc formalism proposed by Jayadev Misra and William Cook [71].

Main challenges related to Web services QoS (Quality of Service) include: 1/ To model and quantify the QoS of a service. 2/ To establish a relation between the QoS of queried Web services and that of the orchestration (contract composition); 3/ To monitor and detect the breaching of a QoS contract, possibly leading to a reconfiguration of the orchestration. Typically, the QoS of a service is modeled by a contract (or Service Level Agreement, SLA) between the provider and consumer of a given service. To account for variability. In previous years, we proposed soft probabilistic contracts specified as probabilistic distributions involving the different QoS parameters; we studied contract composition for such contracts; we developed probabilistic QoS contract monitoring; and we studied the monotonicity of orchestrations; an orchestration is monotonic if a called service improves its performance, then so does the overall orchestration.

This year, in the framework of the Associated Team FOSSA with the University of Texas at Austin (John Thywissen (PhD), Jayadev Misra and William Cook), we have extended our approach to general QoS parameters, i.e., beyond response time. We now encompass composite parameters, which are thus only partially, not totally, ordered. We have developed a general algebra to capture how QoS parameters are transformed while traversing the orchestration and we have extended our study of monotonicity. Finally, we have developed corresponding contract composition procedures. John Thywissen (from UT Austin) and Ajay Kattepur have started extending the Orc language and execution engine to support QoS according to our theory. This extension mainly consists in 1/ providing a rich type system to declare QoS domains and related algebra, and 2/ providing a new operator for Orc that allows for selecting competing returns from different sites on the basis of their QoS. A journal paper is under revision.

A key task in extending Orc for QoS was to extend the Orc engine so that causalities between the different site calls are made explicit at run time while execution progresses. This benefits from our previous work on Orc semantics, but a new set of rules has been proposed to generate causalities in an efficient way, by covering new features of the language. This is joint work of Claude Jard, Ajay Kattepur and John Thywissen from Austin. A publication is in preparation.

Besides this main line of work, other topics have been addressed by Ajay Kattepur as part of his thesis.

- In [41], we study variability of composite services. We model variability as a feature diagram (FD) that captures all valid configurations of its orchestration. Then, we apply pair-wise testing to sample the set of all possible configurations to obtain a concise subset. Finally, we test the composite service for selected pairwise configurations for a variety of QoS metrics such as response time, data quality, and availability. Using two case studies, Car crash crisis management and e-Health management, we demonstrate that pairwise generation effectively samples the full range of QoS variations in a dynamic orchestration. The pairwise sampling technique eliminates over 99% redundancy in configurations, while still calling all atomic services at least once.
• Web services orchestrations conventionally employ exhaustive comparison of runtime quality of service (QoS) metrics for decision making. The ability to incorporate more complex mathematical packages is needed, especially in case of workflows for resource allocation and queuing systems. By modeling such optimization routines as service calls within orchestration specifications, techniques such as linear programming can be conveniently invoked by non-specialist workflow designers. Leveraging on previously developed QoS theory, we propose the use of a high-level flexible query procedure for embedding optimizations in languages such as Orc. The Optima site provides an extension to the sorting and pruning operations currently employed in Orc. Further, the lack of an objective technique for consolidating QoS metrics is a problem in identifying suitable cost functions. We use the analytical hierarchy process (AHP) to generate a total ordering of QoS metrics across various domains. With constructs for ensuring consistency over subjective judgements, the AHP provides a suitable technique for producing objective cost functions. Using the Dell Supply Chain example, we demonstrate the feasibility of decision making through optimization routines, specially when the control flow is QoS dependent. This work was published in [39].

• With web services quality of service (QoS) modeled as random variables, the accuracy of sampled values for precise service level agreements (SLAs) come into question. Samples with lower spread are more accurate for calculating contractual obligations, which is typically not the case for web services QoS. Moreover, the extreme values in case of heavy-tailed distributions (eg. 99.99 percentile) are seldom observed through limited sampling schemes. To improve the accuracy of contracts, we propose the use of variance reduction techniques such as importance sampling. We demonstrate this for contracts involving demand and refuel operations within the Dell supply chain example. Using measured values, efficient forecasting of future deviation of contracts may also be performed. A consequence of this is a more precise definition of sampling, measurement and variance tolerance in SLA declarations. This work was published in [40].

6.8. Specific studies: active documents and web services

Participants: Albert Benveniste, Loïc Hélouët, Benoît Masson.

Active Documents have been introduced by the GEMO team at INRIA Futurs, headed by Serge Abiteboul, mainly through the language Active XML (or AXML for short). AXML is an extension of XML which allows to enrich documents with service calls or sc’s for short. These sc’s point to web services that, when triggered, access other documents; this materialization of sc’s produces in turn AXML code that is included in the calling document. One therefore speaks of dynamic or intentional documents. In the past years, we have collaborated with the GEMO team to study a distributed version of their language.

This year, we have addressed the problem of distributed documents from a different point of view. Starting from our knowledge of distributed active XML (DAXML), we have first proposed a Petri Net semantics for a subset of DAXML [43], and then considered compositionality issues [54]. Compositionality in services can be addressed in several ways: first one have to ensure that modules that provide services an modules that use them agree on the data that they exchange. This notion is called composability of modules. However, composability does not ensure that a service always terminates (i.e. it returns a result to the caller) when it is invoked with appropriate data. Composability plus termination of services is called compatibility. We have shown that under some restrictions on the recursion in active documents, on the data, and upon the assumption that services use positive guards, composability is decidable. This work has also helped us isolate the core idea behind active documents, and propose a model for them called Docnets. Docnets are dynamic Petri nets which places are typed, which transitions are guarded computable type transformations, and which can receive new tokens from their environment. Docnet modules compose well, and if their closure by type transformation is finite, their compatibility is decidable. This work led to a publication [43].

Within the context of the DST associated team, we have proposed a new model, that combines arbitrary numbers of finite workflows, hence allowing for the definition of sessions. Sessions is a central paradigm in web-based systems. As messages exchange between two sites need not follow the same route over the net, a site can not rely on the identity of machines to uniquely define a transaction. This unique identification is
essential, as commercial site, for instance, need to manage several interactions at a given time. The current trend, as in BPEL, is to associate a unique identifier to each session. Modeling realistic session hence often forces to include session counters, and hence render most of properties undecidable. The session formalism studied in 2011 can be seen as a mix of BPEL and ORC elements, but was designed to keep several properties decidable. The strength of this formalism is to allow designing systems that use sessions without the obligation to provide identifiers. The formalism has the expressive power of reset Petri nets for which coverability is decidable. This is sufficient to decide whether a set of agent can be found in some bad configuration during the lifetime of a system. This joint work with Ph. Darondeau from the S4 Team, and with M. Mukund from the Chennai Mathematical Institute led to a publication in the ATVA conference [28].

Our last work on Web-services was the development of an experimental platform. During his post-doc, Benoît Masson has designed a distributed Active XML engine, which can be distributed over a network. We have built a lightweight experimentation platform, made of four linux machines, that run DAXML services and communicate with one another. Simultaneously, R. Abdallah has designed a synthesis tool to generate REST services from High-level Message Sequence Charts. These services were successfully tested on the platform.

6.9. Specific studies: security and privacy

Participants: Guillaume Aucher, Blaise Genest.

We have worked on three parallel lines of research related to security and privacy. The first line deals with problems of delegation and revocation in distributed systems. The second line deals with problems of compliance of a system with respect to a privacy regulation expressed in a language combining epistemic, deontic and dynamic modalities. The third line tackles the minimal information needed at runtime to e.g. break in a (stochastic) system.

6.9.1. Delegation and revocation in distributed systems

Together with Steve Barker from King’s College London, Guido Boella from the University of Torino, Valerio Genovese and Leon van der Torre from the University of Luxembourg, we defined a (sound and complete) propositional dynamic logic to specify and reason about delegation and revocation schemes in distributed systems. This logic describes formally a family of delegation and revocation models that are based on the work of [65]. We extended our logic to accommodate an epistemic interpretation of trust. What emerges from this work is a rich framework of formally well-defined delegation and revocation schemes that accommodates an important trust component. In particular, we showed how to automatically reason about whether an agent is authorized to do an operation on an object and about the authorization policy resulting from the execution of a sequence of actions. We used our logical framework to give a formal account of eight different types of revocation schemes informally introduced in previous literature. This work is published in [18].

6.9.2. Privacy policy with modal logic: the dynamic turn

As explained in Section 6.4, we want to define a logical language to specify privacy policies which is close to the natural language. In general, privacy policies can be defined either in terms of permitted and forbidden knowledge, or in terms of permitted and forbidden actions. For example, it may be forbidden to know the medical data of a person, or it may be forbidden to disclose these data. Implementing a privacy policy based on permitted and forbidden actions is relatively easy, since we can add a filter on the system checking the outgoing messages. Such a filter is an example of a security monitor. If the system attempts to send a forbidden message, then the security monitor blocks the sending of that message. However, the price to pay for this relatively straightforward implementation is that it is difficult to decide which actions are permitted or forbidden so that a piece of information is not disclose. We are therefore interested in privacy policies expressed in terms of permitted and forbidden knowledge. Expressing a privacy policy in terms of permitted and forbidden knowledge is relatively easy, since it lists the situations, where, typically, it may not be permitted to know some sensitive information. Implementing a privacy policy based on permitted and forbidden knowledge is quite difficult, since the system has to reason about the relation between permitted knowledge and actions. The challenge is that the exchange of messages changes the knowledge, and the
security monitor therefore needs to reason about these changes. This inference problem is already non trivial with a static privacy policy, and becomes challenging when privacy policies can change over time. Together with Guido Boella and Leon van der Torre, we therefore introduced a dynamic modal logic that permits not only to reason about permitted and forbidden knowledge to derive the permitted actions, but also to represent explicitly the declarative privacy policies together with their dynamics. The logic can be used to check both regulatory and behavioral compliance, respectively by checking that the permissions and obligations set up by the security monitor of an organization are not in conflict with the privacy policies, and by checking that these obligations are indeed enforced. We also showed that the complexity of the model checking problem is quadratic in the size of the model and the formula and provided the corresponding model-checking algorithms. This work is published in [11].

6.9.3. **Minimal information needed**

Together with Nathalie Bertrand from Vertecs, we tackle the problem of the *minimal information* a user needs *at runtime* to achieve a simple goal, modeled as reaching an objective with probability one [25]. The natural question is then to minimize the additional information the user needs to fulfill her objective. This optimization question gives rise to two different problems, whether we consider to minimize the *worst case cost*, or the *average cost*. On the one hand, concerning the worst case cost, we show that efficient techniques from the model checking community can be adapted to compute the optimal worst case cost and give optimal strategies for the users. On the other hand, we show that the optimal average price (a question typically considered in the AI community) cannot be computed in general, nor can it be approximated in polynomial time even up to a large approximation factor. Following this negative results, we investigate with P.S. Thiagarajan’s group at NUS, Singapore basic algorithms of the AI community to infer the exact probability in (compact) stochastic systems. We proposed in [45] a simple parametrized extension of the usual Factored Frontier algorithm in order to choose the desired accuracy of the algorithm, at the cost of additional but manageable computations. We showed its benefit when dealing with biological pathways.

6.10. **Specific studies: network maintenance**

**Participants:** Eric Fabre, Carole Hounkonnou.

This work represents part of our activities within the research group “High Manageability,” supported by the common lab of Alcatel-Lucent Bell Labs (ALBLF) and INRIA. It concerns a methodology for the graceful shut down and restart of routers in OSPF networks, one of the core protocols of IP networks. A methodology has been proposed to safely switch off the software layer of a router while still maintaining this router in the forwarding plane: the router still forwards packets, but is not able to adapt its routing table to changes in network conditions or topology. Nevertheless, it is possible to check whether this frozen router is harmless or can cause packet losses, through a centralized or distributed algorithm. And if ever it puts the network at risk, minimal patches can be set up temporarily until the router comes back to normal activity. This avoids running twice a global OSPF update at all nodes (once for shutdown of the equipment, one for restart). There is a patent project on this activity, that we don’t detail more here.

6.11. **Specific studies: network and service diagnosis**

**Participants:** Eric Fabre, Carole Hounkonnou.

This work represents part of our activities within the research group “High Manageability,” supported by the common lab of Alcatel-Lucent Bell Labs (ALBLF) and INRIA. It is also supported by the UniverSelf EU integrated project, and conducted in relation with Orange Labs.

The objective is to develop a framework for the joint diagnosis of networks and of the supported services. We are aiming at a model-based approach, in order to tailor the methods to a given network instance and to follow its evolution. We also aim at active diagnosis methods, that collect and reason on alarms provided by the network, but that can also trigger tests or the collection of new observations in order to refine a current diagnosis.
In 2011, the main effort was dedicated to a key and difficult part of this approach: the definition of a methodology for self-modelling. This consists in automatically building a model of the monitored system, by instantiating generic network elements. There are several difficulties to address:

- The model must capture several layers, from the physical architecture up to the service architecture and its protocols. As a case-study, we have chosen VoIP services on an IMS network, deployed over a wired IP network.
- The model should be hierarchical, to allow for multiscale reasoning, and to reflect the intrinsic hierarchical nature of the managed network.
- The model should be generic, i.e. obtained by assembling component instances coming from a reduced set of patterns, just like a text is obtained by assembling words.
- The model should be adaptive, to capture the evolving part of the network (e.g. introduction of new elements) but also its intrinsically dynamic nature (e.g. opened/closed connections).
- The model should display the hierarchical dependency of resources, specifically the fact that lower-level resources are assembled to provide a support to a higher level resource or functionality.
- The model should allow progressive discovery and refinement: for a matter of size, it is not possible to first build a model of the complete network and then monitor it; one must adopt an approach where the model is build on-line, and where the construction is guided by the progress of the diagnosis algorithms.

The first elements of a methodology achieving these objectives have been designed in 2011. The next efforts will aim at refining the grammar of this model, for our specific case study, and at developing the dedicated diagnosis algorithms. For the latter, we envision a new setting of hierarchical and generic Bayesian networks, in order to capture the dependencies between network elements at different granularities.
6. New Results

6.1. Exposure to diffusion in dynamic networks

In many contexts, complex networks are subject to diffusion phenomenon, like spreading of epidemics in human groups or the diffusion of information in social networks. Often, the underlying network is dynamic, that is, his links change along time. Clearly, the dynamics of links has an influence on the diffusion phenomenon taking place over the network. A first step to understand these relationship is to determine which nodes of the dynamic network are more likely to be reached by a diffusion process. We designed new notions of exposure in order to do it, based either on contacts, paths or flows in a dynamic network. In particular, the notion of dynamic flow, which we introduced, has given interesting preliminary results. We computed the exposition scores of nodes of real world dynamic networks and showed that it is correlated to the likelihood of nodes to be affected by a diffusion in the classical SI model.

6.2. Aggregation of temporal graph series

A very natural and extensively used way to represent a dynamic network, where links change along time, is to build a graph series : the series of snapshots of the network taken at different time of its evolution. The way to do so is to aggregate all the contact information on a time window into a single graph : that is, we put an edge between u and v in the graph if they are in contact at least once during the considered time window. Doing so for disjoint windows of equal length which cover the whole period of study, we obtain a series of graphs representing the dynamics of the network. A question remain : how one should choose the length of the aggregation window? The problem is critical since depending on the choice made, the properties of the dynamic network are different and the conclusion derived from its analysis may change. We design a systematic method to estimate the maximum possible aggregation length. Up to our knowledge, this is the first method addressing the problem. It is based on activity rate of dynamic paths in the dynamics. On a dynamic path, only some time steps are used to move within the network. When the aggregation time is short, the activity rate of paths is close to zero and it tends to 1 when this time grow until the whole period of experiments. Between the two behaviors, we showed that there is a phase transition that we interpret as the moment when the properties of the dynamics are distorted because of the too long aggregation time.

6.3. Characterizing changes in dynamic networks

Very often, dynamic networks are described as time series of graphs. Many works focus on analyzing or capturing into models the properties of the graphs of the series. This approach has a clear limitation : it looses the relationships between the different graphs of the series, which however contain a key information on the dynamics. In order to get more insight in the relationships between the graphs of the series, we analyzed the structure of we call the difference graphs. The difference graph of two consecutive graphs $G_1, G_2$ of the series is the graph whose edge set is the symmetric difference of the edge sets of $G_1$ and $G_2$. In other words, this is exactly the graph of the pairs whose adjacency relationship changed from $G_1$ to $G_2$. We showed that the structure of difference graphs is very particular : their edges are concentrated around a small number of vertices. This shows that the changes between two graphs of the series are not spread everywhere in the network, but are due to changes of the neighborhood of only a small number of nodes of the network. We could show this fact by computing a graph parameter called Minimum Vertex Cover (MVC), which is NP-complete to compute. Using a preprocessing step, we could compute the exact value of this parameter for all difference graphs of real world series. We obtained that the value of the MVC on difference graphs is very small compared to the expected value on a random graph with same density. We also showed that the most common models of dynamic networks do not capture this property of concentration of edges in the difference graphs of the series. Our result shed light on the way dynamic networks evolve and open the way to significant improvement of existing models.
6.4. Community detection: dynamic and overlapping

Overlapping community detection is a popular topic in complex networks. Comparing to disjoint community structure, overlapping community structure is more reasonable to describe networks at a macroscopic level. Overlaps shared by communities play an important role in combining different communities. We propose two different approaches to detect overlaps: fuzzy community detection and overlapping community detection. The former estimates membership degree of node belonging to community, and the latter allows node to be shared by communities. In this paper, a fuzzy detection and a clique optimization are introduced. Experimental studies in synthetic networks show fuzzy detection yields meaningful information in stability and hierarchy of communities. And clique optimization is efficient in capturing overlapping nodes. Applications in real networks whose community structure is not well-known find that overlapping clusters found by our fuzzy detection can provide different views than general overlapping nodes in characterize overlaps.

Although community detection has drawn tremendous amount of attention across the sciences in the past decades, no formal consensus has been reached on the very nature of what qualifies a community as such. We take an orthogonal approach by introducing a novel point of view to the problem of overlapping communities. Instead of quantifying the quality of a set of communities, we choose to focus on the intrinsic community-ness of one given set of nodes. To do so, we propose a general metric on graphs, the cohesion, based on counting triangles and inspired by well established sociological considerations. The model has been validated through a large-scale online experiment called Fellows in which users were able to compute their social groups on Facebook and rate the quality of the obtained groups. By observing those ratings in relation to the cohesion we assess that the cohesion is a strong indicator of users subjective perception of the community-ness of a set of people.

6.5. Cross-Layer Optimization for Software Layer to Physical Device layer Mapping

We develop a generic method for mapping software state machines used in protocol stacks and communication layers directly to hardware communication devices using their specifications. The proposed method can handle power modes and timing constraints imposed by hardware devices in order to optimize the software code running on top of the device. This property allows the use of the hardware device in its lowest power consumption mode while making sure that real time constraints are met. To validate the merit of the proposed method, the generated code and power consumption gain, we evaluate the optimizations that can be done on a BMAC medium access control layer used in wireless sensor networks using a large scale experimental testbed. The results show that an average energy consumption gain of up to 60% at the radio level can be achieved.
FOCUS Project-Team

6. New Results

6.1. Contracts and sessions

Participants: Ugo Dal Lago, Ornela Dardha, Maurizio Gabbielli, Elena Giachino, Claudio Guidi, Jacopo Mauro, Fabrizio Montesi.

Contracts are descriptions of the functionalities offered by a component or a service, and of the way these functionalities may be accessed by clients. A contract may include a description of the component capabilities, place constraints on their usage, as well as declare preferences, entitlements and credentials. When a client wants to use one of the functionalities offered, it engages a dialogue (e.g., a sequence of interactions) with the servers; this is usually called a session.

The expected dialogue in a session can be specified by means of types, the session types.

We have studied [11] the integration of union and session types in a class-based language for building network applications, which amalgamates sessions and methods in order to flexibly exchange data according to communication protocols. We have established type safety properties guaranteeing that, after a session has started, computation cannot get stuck on a communication deadlock, and studied type inference. On a similar topic is the paper [23].

We have used types [32] to guarantee bounds on sessions. These are polynomial bounds, on both time and space needed by the interacting processes to carry out the interactions in their sessions. This is the first example of a refinement of session types guaranteeing quantitative properties beyond the usual safety property, and builds on earlier work on soft linear logic.

In service-oriented architectures, the mechanism that allows to manage sessions and, in particular, to assign incoming messages to the correct sessions, is critical for efficiency and performance. A well-known solution to this problem, first introduced by BPEL, makes use of correlation sets. Intuitively these distinguish different sessions by means of the values for some specific variables which are present also in messages, thus allowing for their routing to sessions on the basis of these values. We have studied [36] a typed language for programming services based on correlation sets, that takes into account key aspects of service-oriented systems, such as distribution, loose coupling, open-endedness and integration. We have provided an implementation of the language as an extension of the Jolie language and applied it to a nontrivial real-world example of a fully-functional distributed user authentication system.

In current SOC languages based on correlation sets, a message can be assigned to a unique session. In another line of work on correlation sets [35], we have studied the possibility – useful in many practical examples – of broadcasting messages to more than one session. We have investigated a data structure, based on radix trees, and an algorithm for managing a correlation mechanism, that support the broadcast primitive without degrading performances.

6.2. Fault Handling, compensations and transactions

Participants: Mila Dalla Preda, Maurizio Gabbielli, Ivan Lanese, Jacopo Mauro, Gianluigi Zavattaro.

One of the predominant properties of CBUS is the loose coupling among the components. In fact, components can dynamically connect/disconnect and can be modified/updated at run time. It is thus important to support unexpected events, called faults.

In [30] we have studied the problem of fault handling in the kind of object-oriented languages developed in the EU Hats project; notably these languages have asynchronous method calls whose results are returned inside futures. We present an extension for those languages where futures are used to return fault notifications and to coordinate error recovery between the caller and callee. This can be exploited to ensure that invariants involving many objects are restored after faults.
Traditional fault handling mechanisms, including those based on try-catch operators, do not seem sufficient to deal with the non-local errors and failures of distributed systems. At the application level, more advanced transactional models and primitives are needed to guarantee integrity and continuity of the whole system. We study approaches based on long running transactions and compensations. A long running transaction is a computation that either successfully terminates, or it aborts. In case of abort, a compensation is executed to take the system to a consistent state. In [53], extending work started last year, we make a thorough comparison among different approaches to the specification of compensations, in particular static forms of recovery where the compensation is statically defined together with the transaction, and dynamic forms where the compensation is progressively built along with a computation.

We have also continued our study on faults and compensations in Service Oriented Computing. The approach to the interplay between bi-directional request-response interaction and faults, proposed in our past works on the Jolie language, supported the idea that the bi-directional pattern should not be interrupted in case of faults. However, this may cause long delays or even deadlocks if the communicating partner disappears. On the contrary, the approach of WS-BPEL causes no delay, but it does not allow to compensate the remote activity. We have investigated [38] an intermediate approach in which it is not necessary for the fault handler to wait for the response, but it is still possible on response arrival to gracefully close the conversation with the remote service.

A related work, but mainly developed in 2010, is [21].

6.3. Service orchestration and choreography

Participants: Mila Dalla Preda, Maurizio Gabbielli, Claudio Guidi, Ivan Lanese, Jacopo Mauro, Fabrizio Montesi, Marco Pistore, Gianluigi Zavattaro.

Orchestration has to do with the definition of services that should obey given behaviours. The services may be realised by composing services already available. Orchestration is often discussed in relationship with “choreography”, which refers to global descriptions of the intended behaviour of a system of components, stating the role of each participant and the set of coordination requirements.

In [55] we have studied the basic linguistic constructs and a reference implementation for aggregation, a mechanism for composing services that abstracts from the order of their communications. Aggregation is widely used in practice. However, since it is not natively supported by service-oriented languages, it is mostly implemented by means of ad-hoc solutions which typically exploit middleware technology.

A critical aspect for pervasive computing is the possibility to discover and use process knowledge at run time depending on the specific context. This can be achieved by using an underlying service-based application and exploiting its features in terms of dynamic service discovery, selection, and composition. Pervasive process fragments represent a service-based tool that allows to model incomplete and contextual knowledge. In [40] we provide a solution to automatically compose such fragments into complete processes, according to a specific context and specific goals. We compute the solution by encoding process knowledge, domain knowledge and goals into an AI planning problem.

Concerning choreography languages, two main approaches have been followed in their design: the interaction-oriented approach at the basis of WS-CDL and the process-oriented approach of BPEL4Chor. In [52] we investigate the relationships between the two approaches. In particular, we point out several possible interpretations for interaction-oriented choreographies: one synchronous and various asynchronous. Under each of these possible interpretations we characterize the class of interaction-oriented choreographies which have a direct process-oriented counterpart, and we formalize the corresponding notion of equivalence between the initial interaction-oriented choreography and the corresponding process-oriented counterpart.

In [50] we study the issue of checking a multiparty choreography against formal protocol specifications, and then projecting it onto a description of the individual service orchestrators. Contributions are also the definition of a multiparty choreography model, and the correctness proof for the projection.
6.4. Primitives for adaptable and evolvable components

Participants: Mario Bravetti, Ivan Lanese, Michael Lienhardt, Jacopo Mauro, Marco Pistore, Davide Sangiorgi, Gianluigi Zavattaro.

In Focus we study linguistic primitives for components, and models for them following the process calculus approach. A special emphasis is given to the adaptability and evolvability of the components — important issue in complex software systems. Components indeed are often used in contexts that had not been predicted at the time when the components were built. Moreover, the needs and the requirements on a system may change over time: one may find that the original specification was incomplete or ambiguous, or new needs may arise that had not been predicted at design time. As designing and deploying a system is costly, it is important that the system be capable of evolving and adapting itself to changes in the surrounding environment.

Models and linguistic constructs for adaptability and evolvability of components are studied in [34] and [19]. The key features of the component model in [34] are: a hierarchical structure of components; the capacity to move, update, wrap components; method interfaces for components; and capacities to isolate and distribute components. In the model in [19], adaptable processes have a location and can be subject to dynamic update actions at runtime (related to this paper is also [20]).

In [22] we provide an adaptation approach that can automatically adapt business processes to run-time context changes that impede achievement of a business goal. We define a formal framework that adopts planning techniques to automatically derive necessary adaptation activities on demand. The adaptation consists in identifying recovery activities that guarantee that the execution of a business process can be successfully resumed and, as a consequence, the business goals are achieved. The solution proposed is evaluated on a real-world scenario from the logistics domain.

Adaptability and evolvability are major concerns in Software Product Lines. The EU Hats project has developed the idea of delta-oriented programming (DOP) as a technique for implementing Software Product Lines based on modifications (add, remove, modify) to a core object-oriented program. Such modifications can introduce errors into a program, when type signatures of classes are modified. To overcome this problem we have introduced [54] a type system for delta-oriented programs. The system is based on row polymorphism, a well-known method in type systems for records.

6.5. Resource Control

Participants: Ugo Dal Lago, Marco Gaboardi, Daniel Hirschkoff, Simone Martini, Paolo Parisen Toldin, Giulio Pellitta, Davide Sangiorgi.

In Focus, we study both foundations and methodologies for controlling the amount of resources programs and processes make use of. The employed techniques mainly come from the world of type theory and proof theory, and as such have been used extensively in the context of sequential computation. Interesting results have been obtained recently indicating that those techniques can be quite useful in the concurrent context too, thus being potentially interesting for CBUS.

We have continued our work on techniques for ensuring termination of programs. On the one hand we have refined [25] previous techniques, enhancing them by taking into account input/output capabilities of channels and subtyping. On the other hand we have studied [28] how to transport techniques initially devised for processes onto sequential higher-order languages with imperative features (e.g., λ-calculi with references). The method employed makes it possible to combine term rewriting measure-based techniques for termination of imperative languages with traditional approaches to termination in purely functional languages, such as logical relations.

In [31], a type system of linear and dependent types, called dIPCFl, has been proved to be a sound, but also relatively complete, way to prove intensional (but also extensional) properties of PCF programs. In other words, not only all properties of programs proved by typing are operationally valid, but all true properties of programs can be proved so by way of dIPCFl. This holds not only for terms of base type, but also for (first-order) functions: this makes the type system more expressive than intersection type disciplines.
A characterization of probabilistic polynomial time classes by way of linear typing systems for a variant of Gödel’s T called RSLR, has been proposed [51]. Classes like BPP and PP are characterized by RSLR once appropriate constraints on the probability of error are imposed.

A unifying methodology for the study of resource consumption of processes has been presented in [13]: it is a refinement of realizability, in which not only termination but also concrete resource bounds can be obtained by showing a function to be realized by a program.

On the same topic is the paper [26], that polishes previous work.

### 6.6. Verification of extensional properties

**Participants:** Elena Giachino, Cosimo Laneve, Tudor Alexandru Lascu, Davide Sangiorgi, Gianluigi Zavattaro.

*Extensional* refers to properties that have to do with behavioral descriptions of a system (i.e., how a system looks like from the outside). Examples of such properties include classical functional correctness and deadlock freedom. We mainly employ techniques based on behavioral equivalences (and preorders), and on types and logics. Type systems offer a good trade-off between expressiveness and efficiency of the techniques. A substantial amount of the work carried out this year has to do with the transfer of techniques between the areas of concurrency theory and object-oriented languages.

We have developed [29] a technique for the deadlock analysis of systems of concurrent object groups. The technique makes use of types in the form of contracts, that is, abstract descriptions of method’s behaviours. Object groups are collections of objects that perform collective work. Within a group, there can be only one running thread at a time; the scheduling of threads is cooperative.

We have studied [37] the concept of ownership types, originally introduced for (sequential) object-oriented languages, in the setting of pure message-passing concurrency. Ownership types have the effect of statically preventing certain communications, and can block the accidental or malicious leakage of secrets. Intuitively, a channel defines a boundary and forbids access to its inside from outer channels, thus preserving the secrecy of the inner names from malicious outsiders.

In a different line of work, we have analyzed ad hoc networks, intended as networks of devices connected by wireless links and communicating via broadcast. We have considered [27], [18] models in which the communication topology of a network is represented as a graph. Nodes represent states of individual processes, and adjacent nodes represent single-hop neighbors. Processes are finite state automata that communicate via selective broadcast messages. Reception of a broadcast is restricted to single-hop neighbors. In these systems we have studied various forms of reachability (example: the existence of an initial topology in which the execution of the protocol can lead to a configuration with at least one node in a certain state).

Induction is a pervasive tool in Computer Science and Mathematics for defining structures and reasoning on them. Coinduction is the dual of induction, and as such it brings in tools that are quite different from those provided by induction. The best known instance of coinduction is bisimulation, mainly employed to define and prove equalities among potentially infinite objects: processes, streams, non-well-founded sets, and so on. Sangiorgi has completed [47], [49] two comprehensive textbooks on bisimulation and coinduction (in [49], Sangiorgi is an editor, and author of two chapter contributions [48], [46]). The books explain the fundamental concepts and techniques, and the duality with induction. A special emphasis is put on bisimulation as a behavioural equivalence for processes. Thus the books also serve as an introduction to models for expressing processes, and to the associated techniques of operational and algebraic analysis.

### 6.7. Tutorial papers on Service-Oriented Computing

**Participants:** Mario Bravetti, Ivan Lanese, Davide Sangiorgi, Gianluigi Zavattaro.

We have contributed to a few tutorial papers that summarise the work on languages and tools for Service-Oriented Computing that has been carried out within the EU project Sensoria in Focus and elsewhere. The papers appear as chapters of a book dedicated to the topic.
The chapters [45] and [43] present and contrast the primitives and the behavioural theories of the main process calculi designed for modeling services.

Languages and models for service-oriented applications usually include primitives and constructs for exception and compensation handling. Exception handling is used to react to unexpected events while compensation handling is used to undo previously completed activities. In [44] we investigate the impact of exception and compensation handling in service-oriented process calculi.

The chapter [42] deals with contracts: descriptions intended to provide support for the automatic on-demand discovery of functionalities offered by a service. The approach followed is to describe such contracts as process calculi expressions. We show how, in certain cases, service contracts can be automatically extracted out of service behaviour, and how they can be used to formally check compliance among the communication protocols followed by the interacting services.

Finally, in [41] we present different tools that have been developed for verifying properties of service implementations with respect to their formal specifications in an automated, or semi-automated, way.

### 6.8. Expressiveness of computational models

**Participants:** Mila Dalla Preda, Ugo Dal Lago, Ivan Lanese, Cosimo Laneve, Davide Sangiorgi.

Expressiveness refers to the study of the expressive power of computational models. In 2011 we have addressed four main aspects.

First, we have continued our investigation of reversible computations. Reversibility is a main ingredient in the study of programming abstractions for reliable systems, e.g. for exception handling. In fact, reversibility can be used for going back to some consistent state after an exception has occurred. In previous years we had defined $\rho\pi$, a higher-order calculus where processes can both go forward and backward in the computation. This year we have studied [33] fine-grained rollback primitives to control reversibility. The definition of a proper semantics for such a primitive is a surprisingly delicate matter because of the potential interferences between concurrent rollbacks. We have also considered lower-level distributed semantics, which are closer to an actual implementation of the rollback primitives, and their relationship with the high-level semantics.

A thread of research close to that of $\rho\pi$ is the study of the properties and the expressive power of a simple calculus with reversible transitions, called reversible structures. In [24], we have demonstrated a standardization theorem for these structures. When terms in reversible structures have unique id, the standardization theorem may be strengthened in a form that bears a quadratic algorithm for reachability, a problem that is EXPSPACE-complete for generic structures (as in Petri Nets). The expressive power of reversible structures has been studied in [17], [12] and a compilation of asynchronous Reversible CCS has been provided.

A second aspect has been motivated by the analogy between malicious software and biological infections [39]. In the paper, we have used a formalism originally developed for the analysis of biological systems — the kappa calculus by Danos and Laneve — for the formalization and analysis of malicious software. In particular we have modeled the different actors involved in a malicious code attack in the kappa-calculus. Then, by simulating the behavior, we have shown how to extract relevant information that can drive the choice of the defense technique to apply.

A third aspect has been the refinement [14] of some previous work on the expressiveness and decidability of higher-order concurrent languages. — formalisms for concurrency in which processes can be passed around in communications.

A fourth aspect has been the study of properties of a simple calculus for quantum computation. In [16], we have demonstrated a confluence property both for finite and infinite computations using a novel technique.
4. New Results

4.1. Understanding graph representations

4.1.1. Distributed algorithms without knowledge of global parameters

Participants: Amos Korman, Jean-Sébastien Sereni, Laurent Viennot.

Many fundamental local distributed algorithms are non-uniform, that is, they assume that all nodes know good estimations of one or more global parameters of the network, e.g., the maximum degree $\Delta$ or the number of nodes $n$. In [28], we introduce a rather general technique for transforming a non-uniform algorithm into a uniform one with same asymptotic complexity.

4.1.2. Asymptotic modularity

Participants: Fabien de Montgolfier, Mauricio Soto, Laurent Viennot.

Modularity has been introduced as a quality measure for graph partitioning by Newman and Girvan. It has received considerable attention in several disciplines, especially complex systems. In order to better understand this measure from a graph theoretical point of view, we study in [32], [31] the asymptotic modularity of a variety of graph classes.

4.1.3. Internet Structure

Participants: Fabien de Montgolfier, Mauricio Soto, Laurent Viennot.

In [33], [1], we study the measurement of the Internet according to two graph parameters: treewidth and hyperbolicity.

4.1.4. Multipath Spanners

Participants: Cyril Gavoille, Quentin Godfroy, Laurent Viennot.

Motivated by multipath routing, we introduce in [23], [39] a multi-connected variant of spanners.

4.1.5. $\delta$–hyperbolicity

Participants: Victor Chepoi [CNRS LIF, University of Marseille, France], Feodor Dragan [University of Ohio, USA], Bernard Estellon [CNRS LIF, University of Marseille, France], Michel Habib [CNRS LIAFA, University of Paris Diderot, France], Yann Vaxes [University of Florence, Italy], Yang Xiang [University of Ohio, USA].

$\delta$–Hyperbolic metric spaces have been defined by M. Gromov in 1987 via a simple 4-point condition: for any four points $u, v, w, x$, the two larger of the distance sums $d(u, v) + d(w, x), d(u, w) + d(v, x), d(u, x) + d(v, w)$ differ by at most $2\delta$. They play an important role in geometric group theory, geometry of negatively curved spaces, and have recently become of interest in several domains of computer science, including algorithms and networking. In [5] paper, we study un-weighted $\delta$–hyperbolic graphs. Using the Layering Partition technique, we show that every $n$–vertex $\delta$–hyperbolic graph with $\delta \geq 1/2$ has an additive $O(\delta \log n)$–spanner with at most $O(\delta n)$ edges and provide a simpler, in our opinion, and faster con- struction of distance approximating trees of $\delta$-hyperbolic graphs with an additive error $O(\delta \log n)$. The construction of our tree takes only linear time in the size of the input graph. As a consequence, we show that the family of $n$–vertex $\delta$–hyperbolic graphs with $\delta \geq 1/2$ admits a routing labeling scheme with $O(\delta \log^2 n)$ bit labels, $O(\delta \log n)$ additive stretch and $O(\log^2 (4\delta))$ time routing protocol, and a distance labeling scheme with $O(\log^2 n)$ bit labels, $O(\delta \log n)$ additive error and constant time distance decoder.
4.1.6. Perfect Phylogeny

4.1.6.1. Perfect Phylogeny Is $NP$–Hard

**Participants:** Michel Habib [CNRS LIAFA, University of Paris Diderot, France], Juraj Stacho [University of Haifa, Israel].

We answer in the affirmative [24], to the question proposed by Mike Steel as a $100$ challenge: “Is the following problem $NP$–hard? Given a ternary 1 phylogenetic $X$-tree $T$ and a collection $Q$ of quartet subtrees on $X$, is $T$ the only tree that displays $Q$?” As a particular consequence of this, we show that the unique chordal sandwich problem is also $NP$–hard.

4.1.6.2. Compatibility of Multi-states Characters

**Participants:** Michel Habib [CNRS LIAFA, University of Paris Diderot, France], Thu-Hien To [CNRS LIAFA, University of Paris Diderot, France].

Perfect phylogeny consisting of determining the compatibility of a set of characters is known to be $NP$–complete. We propose in [25], a conjecture on the necessary and sufficient conditions of compatibility: Given a set $C$ of $r$–states full characters, there exists a function $f(r)$ such that $C$ is compatible if $f$ every set of $f(r)$ characters of $C$ is compatible. According to numerous references, $f(2) = 2$, $f(3) = 3$ and $f(r) \geq r$. Some conjectured that $f(r) = r$ for any $r \geq 2$. In this paper, we present an example showing that $f(4) \geq 5$. Therefore it could be the case that for $r \geq 4$ characters, the problem behavior drastically changes. In a second part, we propose a closure operation for chordal sandwich graphs. The later problem is a common approach of perfect phylogeny.

4.1.7. Graph sandwich

**Participants:** Arnaud Durand [CNRS LIAFA, University of Paris Diderot, France], Michel Habib [CNRS LIAFA, University of Paris Diderot, France].

Graph sandwich problems were introduced by Golumbic et al. (1994) in [12] for DNA physical mapping problems and can be described as follows. Given a property $\Pi$ of graphs and two disjoint sets of edges $E_1 \cup E_2$ on a vertex set $V$, the problem is to find a graph $G$ on $V$ with edge set $E_s$ having property $\Pi$ and such that $E_1 \subseteq E_s \subseteq E_2$. In [8] paper, we exhibit a quasi-linear reduction between the problem of finding an independent set of size $k \geq 2$ in a graph and the problem of finding a sandwich homogeneous set of the same size $k$. Using this reduction, we prove that a number of natural (decision and counting) problems related to sandwich homogeneous sets are hard in general. We then exploit a little further the reduction and show that finding efficient algorithms to compute small sandwich homogeneous sets would imply substantial improvement for computing triangles in graphs.

4.1.8. Diameter of Real-World Undirected Graphs

**Participants:** Pierluigi Crescenzi [University of Florence, Italy], Roberto Grossi [University of Pisa, Italy], Michel Habib [CNRS LIAFA, University of Paris Diderot, France], Lorenzo Lanzi [University of Florence, Italy], Andrea Marino [University of Florence, Italy].

In [16], we propose a new algorithm for computing the diameter of undirected unweighted graphs. Even though, in the worst case, this algorithm has complexity $O(nm)$, where $n$ is the number of nodes and $m$ is the number of edges of the graph, we experimentally show (on almost 200 real-world graphs) that in practice our method works in linear time. Moreover, we show how to extend our algorithm to the case of undirected weighted graphs and, even in this case, we present some preliminary very positive experimental results.

4.1.9. Parsimonious flooding in dynamic graphs

**Participants:** Hervé Baumann [CNRS LIAFA, University of Paris Diderot, France], Pierluigi Crescenzi [University of Florence, Italy], Pierre Fraigniaud [CNRS LIAFA, University of Paris Diderot, France].
An edge-Markovian process with birth-rate $p$ and death-rate $q$ generates infinite sequences of graphs $(G_0, G_1, G_2, \ldots)$ with the same node set $[n]$ such that $G_t$ is obtained from $G_{t-1}$ as follows: if $e \notin E(G_{t-1})$ then $e \in E(G_t)$ with probability $p$, and if $e \in E(G_{t-1})$ then $e \notin E(G_t)$ with probability $q$. In [2], we establish tight bounds on the complexity of flooding in edge-Markovian graphs, where flooding is the basic mechanism in which every node becoming aware of an information at step $t$ forwards this information to all its neighbors at all forthcoming steps $t' > t$. These bounds complete previous results obtained by Clementi et al. Moreover, we also show that flooding in dynamic graphs can be implemented in a parsimonious manner, so that to save bandwidth, yet preserving efficiency in terms of simplicity and completion time. For a positive integer $k$, we say that the flooding protocol is $k$-active if each node forwards an information only during the $k$ time steps immediately following the step at which the node receives that information for the first time. We define the reachability threshold for the flooding protocol as the smallest integer $k$ such that, for any source $s[n]$, the $k$-active flooding protocol from $s$ completes (i.e., reaches all nodes), and we establish tight bounds for this parameter. We show that, for a large spectrum of parameters $p$ and $q$, the reachability threshold is by several orders of magnitude smaller than the flooding time. In particular, we show that it is even constant whenever the ratio $p/(p + q)$ exceeds $\log n/n$. Moreover, we also show that being active for a number of steps equal to the reachability threshold (up to a multiplicative constant) allows the flooding protocol to complete in optimal time, i.e., in asymptotically the same number of steps as when being perpetually active. These results demonstrate that flooding can be implemented in a practical and efficient manner in dynamic graphs. The main ingredient in the proofs of our results is a reduction lemma enabling to overcome the time dependencies in edge-Markovian dynamic graphs.

4.2. Distributed computational complexities

4.2.1. Local Distributed Decision

Participants: Pierre Fraigniaud [CNRS LIAFA, University of Paris Diderot, France], Amos Korman [CNRS LIAFA, University of Paris Diderot, France], David Peleg [Weizmann Institute of Science, Israel].

Inspired by sequential complexity theory, in [20] we focus on a complexity theory for distributed decision problems. We first study the intriguing question of whether randomization helps in local distributed computing, and to what extent. Our main result provides a sharp threshold for the impact of randomization on decision hereditary problems. In addition, we investigate the impact of non-determinism on local decision, and establish some structural results inspired by classical computational complexity theory. Specifically, we show that non-determinism does help, but that this help is limited, as there exist languages that cannot be decided nondeterministically. Perhaps surprisingly, it turns out that it is the combination of randomization with non-determinism that enables to decide all languages in constant time. Finally, we introduce the notion of local reduction, and establish some completeness results.

4.2.2. Asynchronous Wait-free Decision

Participants: Pierre Fraigniaud [CNRS LIAFA, University of Paris Diderot, France], Sergio Rajsbaum [Maths. Institute, University of Mexico, Mexico], Corentin Travers [Technion, Israel].

In order to capture the core of asynchronous distributed decision model, we address in [22] the wait-free model with crash failures. The set of tasks whose input is a pair $(s, t)$ and deciding whether $t \in \Delta(s)$, i.e., whether $t$ is a valid output for $s$, has been proven to be decidable in this model.

4.2.3. Mobile Distributed Decision

Participants: Pierre Fraigniaud [CNRS LIAFA, University of Paris Diderot, France], Andrzej Pelc [UQO, University of Quebec, Canada].

In [21], we partially answer the question of decidability of any language for mobile agents in a 2D environment like telecom networks or robots. It is proven that, for every agent, verifying whether (i) he/she is alone or not and (ii) he/she is able to capture the environment, is associated with the question of pertaining to an equivalence class of a map. A positive answer helps in the non-deterministic decision for any language for mobile agent.
4.2.4. Approximating the Statistics of various Properties in Randomly Weighted Graphs

**Participants:** Yuval Emek [University of Tel Aviv, Israel], Amos Korman [CNRS LIAFA, University of Paris Diderot, France], Yuval Shavitt [University of Tel Aviv, Israel].

In [19], we consider the setting of randomly weighted graphs. Under this setting, weighted graph properties typically become random variables and we are interested in computing their statistical features. Unfortunately, this turns out to be computationally hard for some weighted graph properties albeit the problem of computing the properties per se in the traditional setting of algorithmic graph theory is tractable. For example, there are well known efficient algorithms that compute the diameter of a given weighted graph, yet, computing the expected diameter of a given randomly weighted graph is \#P-hard even if the edge weights are identically distributed. In this paper, we define a family of weighted graph properties and show that for each property in this family, the problem of computing the \( k \)’th moment (and in particular, the expected value) of the corresponding random variable in a given randomly weighted graph \( G \) admits a fully polynomial time randomized approximation scheme (FPRAS) for every fixed \( k \). This family includes fundamental weighted graph properties such as the diameter of \( G \), the radius of \( G \) (with respect to any designated vertex) and the weight of a minimum spanning tree of \( G \).

4.2.5. New bounds for the controller problem

**Participants:** Yuval Emek [University of Tel Aviv, Israel], Amos Korman [CNRS LIAFA, University of Paris Diderot, France].

In [10], we establish two new lower bounds on the message complexity of the controller problem. We first prove a simple lower bound stating that any \((M, W)\)-controller must send \( \Omega(N \log \frac{M}{W+1}) \) messages. Second, for the important case when \( W \) is proportional to \( M \) (this is the common case in most applications), we use a surprising reduction from the (centralized) monotonic labeling problem to show that any \((M, W)\)-controller must send \( \Omega(N \log N) \) messages. In fact, under a long lasting conjecture regarding the complexity of the monotonic labeling problem, this lower bound is improved to a tight \( \Omega(N \log^2 N) \).

4.2.6. Online computation with advice

**Participants:** Yuval Emek [University of Tel Aviv, Israel], Pierre Fraigniaud [CNRS LIAFA, University of Paris Diderot, France], Amos Korman [CNRS LIAFA, University of Paris Diderot, France], Adi Rosén [CNRS LIAFA, University of Paris Diderot, France].

In [9], we consider a model for online computation in which the online algorithm receives, together with each request, some information regarding the future, referred to as advice. We are interested in the impact of such advice on the competitive ratio, and in particular, in the relation between the size \( b \) of the advice, measured in terms of bits of information per request, and the (improved) competitive ratio. In this paper we propose the above model and illustrate its applicability by considering two of the most extensively studied online problems, namely, metrical task systems (MTS) and the \( k \)-server problem.

4.2.7. Tight Bounds For Distributed MST Verification

**Participants:** Liah Kor [Weizmann Institute of Science, Israel], Amos Korman [CNRS LIAFA, University of Paris Diderot, France], David Peleg [Weizmann Institute of Science, Israel].

In [26], we establish tight bounds for the Minimum-weight Spanning Tree (MST) verification problem in the distributed setting. Specifically, we provide an MST verification algorithm that achieves simultaneously \( O(\sqrt{n} + D) \) time, where \( |E| \) is the number of edges in the given graph \( G \) and \( D \) is \( G \)'s diameter. On the negative side, we show that any MST verification algorithm must send \( \Omega(|E|) \) messages and incur \( \tilde{O}(\sqrt{n} + D) \) time in worst case. Our upper bound result appears to indicate that the verification of an MST may be easier than its construction.

4.2.8. Distributed verification and hardness of distributed approximation

**Participants:** Atish Das Sarma [Google research, USA], Stephan Holzer [ETH, Zurich, Switzerland], Liah Kor [Weizmann Institute of Science, Israel], Amos Korman [CNRS LIAFA, University of Paris Diderot, France], Danupon Nanongkai [Nanyang Technological University, Singapore], David Peleg [Weizmann Institute of Science, Israel], Roger Wattenhofer [ETH, Zurich, Switzerland].
In [30], we initiate a systematic study of distributed verification, and give almost tight lower bounds on the running time of distributed verification algorithms for many fundamental problems such as connectivity, spanning connected subgraph, and $s - t$ cut verification. We then show applications of these results in deriving strong unconditional time lower bounds on the hardness of distributed approximation for many classical optimization problems including minimum spanning tree, shortest paths, and minimum cut. Many of these results are the first non-trivial lower bounds for both exact and approximate distributed computation and they resolve previous open questions. Moreover, our unconditional lower bound of approximating minimum spanning tree (MST) subsumes and improves upon the previous hardness of approximation bound of Elkin [STOC 2004] as well as the lower bound for (exact) MST computation of Peleg and Rubinovich [FOCS 1999]. Our result implies that there can be no distributed approximation algorithm for MST that is significantly faster than the current exact algorithm, for any approximation factor. Our lower bound proofs show an interesting connection between communication complexity and distributed computing which turns out to be useful in establishing the time complexity of exact and approximate distributed computation of many problems.

### 4.3. Peer to Peer Networks Performance

**Participants:** Fabien Mathieu, François Baccelli.

In [3], we present and discuss possible architectures for P2P systems to manage overlays that try to cope with the underlying network.

In [40], [29], we discuss theoretical performance issues that arise from using “Live Seeding”, a technique that can be employed to leverage the capacity of a P2P/Hybrid Live Streaming Systems by utilizing the capacities of idle peers.

In [38], we propose a new paradigm for P2P networks, where the bandwidth bottleneck is not the access node anymore. This new model is versatile enough to be used in the context of classical networks with congestion control, wireless networks, or semantic networks.

### 4.4. Fault Tolerance in Distributed Networks

#### 4.4.1. Verification of population protocols

**Participants:** Hugues Fauconnier, Carole Gallet-Delporte.

In [15], we address the problem of verification by model-checking of the basic population protocol (PP) model of Angluin et al. This problem has received special attention in the last two years and new tools have been proposed to deal with it. We show that the problem can be solved by using the existing model-checking tools, e.g., Spin and Prism. In order to do so, we apply the counter abstraction to get an abstraction of the PP model which can be efficiently verified by the existing model-checking tools. Moreover, this abstraction preserves the correct stabilization property of PP models. To deal with the fairness assumed by the PP models, we provide two new recipes. The first one gives sufficient conditions under which the PP model fairness can be replaced by the weak fairness implemented in Spin. We show that this recipe can be applied to several PP models. In the second recipe, we show how to use probabilistic model-checking and, in particular, Prism to take completely in consideration the fairness of the PP models. The correctness of this recipe is based on existing theorems involving finite discrete Markov chains. An abstract of this paper has been also published in [34].

#### 4.4.2. Failure Detection

**Participants:** Hugues Fauconnier, Carole Gallet-Delporte.

What does it mean to solve a distributed task? In Paxos, Lamport proposed a definition of solvability in which every process is split into a proposer that submits commands to be executed, an acceptor that takes care of the command execution order, and a learner that receives the outcomes of executed commands. The resulting perspective of computation in which every proposed command can be executed, be its proposer correct or faulty, proved to be very useful when processes take steps on behalf of each other, i.e., in simulations.
Most interesting tasks cannot be solved asynchronously, and failure detectors were proposed to circumvent these impossibilities. Alas, when it comes to solving a task using a failure detector, we cannot leverage simulation-based techniques. A process cannot perform steps of failure detector-based computation on behalf of another process, since it cannot access the remote failure-detector module.

In [17], we propose a new definition of solving a task with a failure detector in which computation processes that propose inputs and provide outputs are treated separately from synchronization processes that coordinate using a failure detector. In the resulting framework, any failure detector is shown to be equivalent to the availability of some \( k \)-set agreement. As a corollary, we obtain a complete classification of tasks, including ones that evaded comprehensible characterization so far, such as renaming.

Shared objects like atomic register, test-and-set, cmp-and-swap are classical hardware primitives that help to develop fault-tolerant distributed applications. In order to compare shared objects, in [41], we consider their implementations in message passing models. With the minimal failure detector for each object, we get a new hierarchy that has only two levels. This paper summarizes recent works and results on this topic.

In [7], we first define the basic notions of local and non-local tasks for distributed systems. Intuitively, a task is local if, in a system with no failures, each process can compute its output value locally by applying some local function on its own input value (so the output value of each process depends only on the process' own input value, not on the input values of the other processes); a task is non-local otherwise. All the interesting distributed tasks, including all those that have been investigated in the literature (e.g., consensus, set agreement, renaming, atomic commit, etc.) are non-local.

In this paper we consider non-local tasks and determine the minimum information about failures that is necessary to solve such tasks in message-passing distributed systems. As part of this work, we also introduces weak set agreement — a natural weakening of set agreement — and show that, in some precise sense, it is the weakest non-local task in message-passing systems.

4.4.3. Adversary disagreement and Byzantine agreement

Participants: Hugues Fauconnier, Carole Gallet-Delporte.

At the heart of distributed computing lies the fundamental result that the level of agreement that can be obtained in an asynchronous shared memory model where \( t \) processes can crash is exactly \( t + 1 \). In other words, an adversary that can crash any subset of size at most \( t \) can prevent the processes from agreeing on \( t \) values. But what about all the other \( 2^{2^{n-1}} - (n + 1) \) adversaries that are not uniform in this sense and might crash certain combination of processes and not others? In [6], we present a precise way to classify all adversaries. We introduce the notion of disagreement power: the biggest integer \( k \) for which the adversary can prevent processes from agreeing on \( k \) values. We show how to compute the disagreement power of an adversary and derive \( n \) equivalence classes of adversaries.

So far, the distributed computing community has either assumed that all the processes of a distributed system have distinct identifiers or, more rarely, that the processes are anonymous and have no identifiers. These are two extremes of the same general model: namely, \( n \) processes use \( \ell \) different authenticated identifiers, where \( 1 \leq \ell \leq n \). In [18], we ask how many identifiers are actually needed to reach agreement in a distributed system with \( t \) Byzantine processes.

We show that having \( 3\ell + 1 \) identifiers is necessary and sufficient for agreement in the synchronous case but, more surprisingly, the number of identifiers must be greater than \( \frac{n + 3t}{2} \) in the partially synchronous case. This demonstrates two differences from the classical model (which has \( \ell = n \)): there are situations where relaxing synchrony to partial synchrony renders agreement impossible; and, in the partially synchronous case, increasing the number of correct processes can actually make it harder to reach agreement. The impossibility proofs use the fact that a Byzantine process can send multiple messages to the same recipient in a round. We show that removing this ability makes agreement easier: then, \( t + 1 \) identifiers are sufficient for agreement, even in the partially synchronous model.
4.4.4. Fast and compact self stabilizing verification, computation, and fault detection of an MST

**Participants:** Amos Korman [CNRS LIAFA, University of Paris Diderot, France], Shay Kutten [Technion, Israel], Toshimitsu Masuzawa [Osaka University, Japan].

In [27], we address the impact of optimizing the memory size on the time complexity, and show that this carries at most a small cost in terms of time in the context of MST. Specifically, we present a self stabilizing distributed verification algorithm whose time complexity is $O(\log^2 n)$ in synchronous networks, or $O(\Delta \log^2 n)$ in asynchronous networks, where $\Delta$ denotes the largest degree of a node. More importantly, the memory size at each node remains optimal - $O(\log n)$ bits throughout the execution. This answers an open problem posed by Awerbuch and Varghese (FOCS 1991). We also show that $\Omega(\log n)$ time is necessary if the memory size is restricted to $O(\log n)$ bits, even in synchronous networks. We demonstrate the usefulness of our verification scheme by using it as a module in a new self stabilizing MST construction algorithm. This algorithm has the important property that, if faults occur after the construction ended, they are detected by some nodes within $O(\log^2 n)$ time in synchronous networks, or within $O(\Delta \log^2 n)$ time in asynchronous networks. The rest of the nodes detect within $O(D \log n)$ time, where $D$ denotes the diameter. Moreover, if a constant number of faults occur, then, within the required detection time above, they are detected by some node in the $O(\log n)$ locality of each of the faults. The memory size of the self stabilizing MST construction is $O(\log n)$ bits per node (optimal), and the time complexity is $O(n)$. This time complexity is significantly better than the best time complexity of previous self stabilizing MST algorithms, that was $\Omega(n^2)$ even when using memory of $\Omega(\log^2 n)$ bits, and even without having the above localized fault detection property. The time complexity of previous algorithms that used $O(\log n)$ memory size was $O(n|E|)$.

4.5. Discrete Optimization Algorithms

4.5.1. Estimating Satisfiability

**Participants:** Yacine Boufkhad, Thomas Hugel.

In [4], the problem of estimating the proportion of satisfiable instances of a given CSP (constraint satisfaction problem) can be tackled through weighting. It consists in putting onto each solution a non-negative real value based on its neighborhood in a way that the total weight is at least 1 for each satisfiable instance. We define in this paper a general weighting scheme for the estimation of satisfiability of general CSPs. First we give some sufficient conditions for a weighting system to be correct. Then we show that this scheme allows for an improvement on the upper bound on the existence of non-trivial cores in 3-SAT obtained by Maneva and Sinclair (2008) to 4.419. Another more common way of estimating satisfiability is ordering. This consists in putting a total order on the domain, which induces an orientation between neighboring solutions in a way that prevents circuits from appearing, and then counting only minimal elements. We compare ordering and weighting under various conditions.

4.5.2. Eigenvectors of three term recurrence Toeplitz matrices and Riordan group

**Participant:** Dominique Fortin.

Eigenvalues of tridiagonal (including main) Toeplitz matrices are analytically known under some regular distance to the main diagonal. Any eigenvector may be easily computed then, through a backward process; instead, in [11], we give an analytical form for each component through the reciprocation of the underlied trinomial. More generally, the connection to the Riordan group follows some bilinear iterative process.

4.5.3. Piecewise Convex Maximization problems and algorithms

**Participants:** Dominique Fortin, Ider Tseveendorj.

In [14], we provide a global search algorithm for maximizing a piecewise convex function $F$ over a compact $D$. We propose to iteratively refine the function $F$ at local solution $y$ by a virtual cutting function $p_y(\cdot)$ and to solve
\[ \max \{ \min \{ F(x) - F(y), p_y(x) \} \mid x \in D \} \]

Instead. We call this function either a patch, when it avoids returning back to the same local solutions, or a pseudo patch, when it possibly yields a better point. It is \textit{virtual} in the sense that the role of cutting constraints is played by additional convex pieces in the objective function. We report some computational results, that represent an improvement on previous linearization based techniques.

It is well known that maximization of any difference of convex functions could be turned into a convex maximization; in [13], we aim at a piecewise convex maximization problem instead. Despite, it may seem harder, sometimes the dimension may be reduced by 1 and the local search improved by using extreme points of the closure of the convex hull of better points. We show that it is always the case for both binary and permutation problems and give, as such instances, piecewise convex formulations for the maximum clique problem and the quadratic assignment problem.

In [12], we consider mathematical programming problems with the so-called piecewise convex objective functions. A solution method for this interesting and important class of nonconvex problems is presented. This method is based on Newton’s law of universal gravitation, multicriteria optimization and Helly’s theorem on convex bodies. Numerical experiments using well known classes of test problems on piecewise convex maximization, convex maximization as well as the maximum clique problem show the efficiency of the approach.
6. New Results

6.1. Scheduling Strategies and Algorithm Design for Heterogeneous Platforms


6.1.1. Virtual Machine Resource Allocation for Service Hosting on Heterogeneous Distributed Platforms

We proposed algorithms for allocating multiple resources to competing services running in virtual machines on heterogeneous distributed platforms. We developed a theoretical problem formulation, designed algorithms, and compared these algorithms via simulation experiments based in part on workload data supplied by Google. Our main finding is that vector packing approaches proposed in the homogeneous case can be extended to provide high-quality solutions in the heterogeneous case, and combined to provide a single efficient algorithm. We also considered the case when there may be errors in estimates of performance-related resource needs. We provided a resource sharing algorithm and proved that for the single-resource, single-node case, when there is no bound on the error, its performance ratio relative to an omniscient optimal algorithm is $\frac{2J}{J+1}$, where $J$ is the number of services. We also provided a heuristic approach for compensating for bounded errors in resource need estimates that performs well in simulation.

6.1.2. Dynamic Fractional Resource Scheduling vs. Batch Scheduling

We finalized this work in which we proposed a novel job scheduling approach for homogeneous cluster computing platforms. Its key feature is the use of virtual machine technology to share fractional node resources in a precise and controlled manner. Other VM-based scheduling approaches have focused primarily on technical issues or extensions to existing batch scheduling systems, while we take a more aggressive approach and seek to find heuristics that maximize an objective metric correlated with job performance. We derived absolute performance bounds and developed algorithms for the online, non-clairvoyant version of our scheduling problem. We further evaluated these algorithms in simulation against both synthetic and real-world HPC workloads and compared our algorithms to standard batch scheduling approaches. We found that our approach improves over batch scheduling by orders of magnitude in terms of job stretch, while leading to comparable or better resource utilization. Our results demonstrated that virtualization technology coupled with lightweight online scheduling strategies can afford dramatic improvements in performance for executing HPC workloads.

6.1.3. Greedy algorithms for energy minimization

This year, we have revisited the well-known greedy algorithm for scheduling independent jobs on parallel processors, with the objective of energy minimization. We have assessed the performance of the online version, as well as the performance of the offline version, which sorts the jobs by non-increasing size before execution. We have derived new approximation factors, as well as examples that show that these factors cannot be improved, thereby completely characterizing the performance of the algorithms.

6.1.4. Energy-aware mappings on chip multiprocessors

This year, in collaboration with Rami Melhem at Pittsburgh University (USA), we have studied the problem of mapping streaming applications that can be modeled by a series-parallel graph, onto a 2-dimensional tiled CMP architecture. The objective of the mapping is to minimize the energy consumption, using dynamic and voltage scaling techniques, while maintaining a given level of performance, reflected by the rate of processing the data streams. This mapping problem turned out to be NP-hard, but we identified simpler instances, whose
optimal solution can be computed by a dynamic programming algorithm in polynomial time. Several heuristics were proposed to tackle the general problem, building upon the theoretical results. Finally, we assessed the performance of the heuristics through comprehensive simulations using the StreamIt workflow suite and various CMP grid sizes.

We are pursuing this work by investigating the routing of communications in chip multiprocessors (CMPs). The goal is to find a valid routing in the sense that the amount of data routed between two neighboring cores does not exceed the maximum link bandwidth while the power dissipated by communications is minimized. Our position is at the system level: we assume that several applications, described as task graphs, are executed on a CMP, and each task is already mapped to a core. Therefore, we consider a set of communications that have to be routed between the cores of the CMP. We consider a classical model, where the power consumed by a communication link is the sum of a static part and a dynamic part, with the dynamic part depending on the frequency of the link. This frequency is scalable and it is proportional to the throughput of the link. The most natural and widely used algorithm to handle all these communications is XY routing: for each communication, data is first forwarded horizontally, and then vertically, from source to destination. However, if it is allowed to use all Manhattan paths between the source and the destination, the consumed power can be reduced dramatically. Moreover, some solutions may be found while none existed with the XY routing. We have compared XY routing and Manhattan routing, both from a theoretical and from a practical point of view. We considered two variants of Manhattan routing: in single-path routing, only one path can be used for each communication, while multi-paths routing allows to split a communication between different routes. We established the NP-completeness of the problem of finding a Manhattan routing that minimizes the dissipated power, we exhibited the minimum upper bound of the ratio power consumed by an XY routing over power consumed by a Manhattan routing, and finally we performed simulations to assess the performance of Manhattan routing heuristics that we designed.

6.1.5. Power-aware replica placement

We have investigated optimal strategies to place replicas in tree networks, with the double objective to minimize the total cost of the servers, and/or to optimize power consumption. The client requests are known beforehand, and some servers are assumed to pre-exist in the tree. Without power consumption constraints, the total cost is an arbitrary function of the number of existing servers that are reused, and of the number of new servers. Whenever creating and operating a new server has higher cost than reusing an existing one (which is a very natural assumption), cost optimal strategies have to trade-off between reusing resources and load-balancing requests on new servers. We provide an optimal dynamic programming algorithm that returns the optimal cost, thereby extending known results without pre-existing servers. With power consumption constraints, we assume that servers operate under a set of $M$ different modes depending upon the number of requests that they have to process. In practice $M$ is a small number, typically 2 or 3, depending upon the number of allowed voltages. Power consumption includes a static part, proportional to the total number of servers, and a dynamic part, proportional to a constant exponent of the server mode, which depends upon the model for power. The cost function becomes a more complicated function that takes into account reuse and creation as before, but also upgrading or downgrading an existing server from one mode to another. We have shown that with an arbitrary number of modes, the power minimization problem is NP-complete, even without cost constraint, and without static power. Still, we have provided an optimal dynamic programming algorithm that returns the minimal power, given a threshold value on the total cost; it has exponential complexity in the number of modes $M$, and its practical usefulness is limited to small values of $M$. Still, experiments conducted with this algorithm showed that it can process large trees in reasonable time, despite its worst-case complexity.

6.1.6. Reclaiming the energy of a schedule

In this work, we consider a task graph to be executed on a set of processors. We assume that the mapping is given, say by an ordered list of tasks to execute on each processor, and we aim at optimizing the energy consumption while enforcing a prescribed bound on the execution time. While it is not possible to change the allocation of a task, it is possible to change its speed. Rather than using a local approach such as backfilling, we have considered the problem as a whole and studied the impact of several speed variation models on its
complexity. For continuous speeds, we gave a closed-form formula for trees and series-parallel graphs, and we cast the problem into a geometric programming problem for general directed acyclic graphs. We showed that the classical dynamic voltage and frequency scaling (DVFS) model with discrete modes leads to a NP-complete problem, even if the modes are regularly distributed (an important particular case in practice, which we analyzed as the incremental model). On the contrary, the VDD-hopping model leads to a polynomial solution. Finally, we provided an approximation algorithm for the incremental model, which we extended for the general DVFS model.

6.1.7. Workload balancing and throughput optimization

We have investigated the problem of optimizing the throughput of streaming applications for heterogeneous platforms subject to failures. The applications are linear graphs of tasks (pipelines), and a type is associated to each task. The challenge is to map tasks onto the machines of a target platform, but machines must be specialized to process only one task type, in order to avoid costly context or setup changes. The objective is to maximize the throughput, i.e., the rate at which jobs can be processed when accounting for failures. For identical machines, we have proved that an optimal solution can be computed in polynomial time. However, the problem becomes NP-hard when two machines can compute the same task type at different speeds. Several polynomial time heuristics have been designed, and simulation results have demonstrated their efficiency.

6.1.8. Comparing archival policies for BlueWaters

In this work, we focus on the archive system which will be used in the BlueWaters supercomputer. We have introduced two new tape archival policies that can improve tape archive performance in certain regimes, compared to the classical RAIT (Redundant Array of Independent Tapes) policy. The first policy, PARALLEL, still requires as many parallel tape drives as RAIT but pre-computes large data stripes that are written contiguously on tapes to increase write/read performance. The second policy, VERTICAL, writes contiguous data into a single tape, while updating error correcting information on the fly and delaying its archival until enough data has been archived. This second approach reduces the number of tape drives used for every user request to one. The performance of the three RAIT, PARALLEL and VERTICAL policies have been assessed through extensive simulations, using a hardware configuration and a distribution of I/O requests similar to these expected on the BlueWaters system. These simulations have shown that VERTICAL is the most suitable policy for small files, whereas PARALLEL must be used for files larger than 1 GB. We have also demonstrated that RAIT never outperforms both proposed policies, and that a heterogeneous policy mixing VERTICAL and PARALLEL performs 10 times better than any other policy.

6.1.9. Using Virtualization and Job Folding for Batch Scheduling

In this work we study the problem of batch scheduling within a homogeneous cluster. In this context, the problem is that the more processors the job requires the more difficult it is to find an idle slot to run it on. As a consequence the resources are often inefficiently used as some of them remain unallocated in the final schedule. To address this issue we propose a technique called job folding that uses virtualization to reduce the number of processors allocated to a parallel job and thus allows to execute it earlier. Our goal is to optimize the resource use. We propose several heuristics based on job folding and we compare their performance with classical on-line scheduling algorithms as FCFS or backfilling. The contributions of this work are both the design of the job folding algorithms and their performance analysis.

6.1.10. A Genetic Algorithm with Communication Costs to Schedule Workflows on a SOA-Grid

We propose in this work the problem of scheduling a collection of workflows, identical or not, on a SOA (Service Oriented Architecture) grid. A workflow (job) is represented by a directed acyclic graph (DAG) with typed tasks. All of the grid hosts are able to process a set of typed tasks with unrelated processing costs and are able to transmit files through communication links for which the communication times are not negligible. The goal of our study is to minimize the maximum completion time (makespan) of the workflows. To solve this problem we propose a genetic approach. The contributions of this paper are both the design of a Genetic Algorithm taking the communication costs into account and its performance analysis.
6.1.11. Checkpointing policies for post-petascale supercomputers

In this work, we provided an analysis of checkpointing strategies for minimizing expected job execution times in an environment that is subject to processor failures. In the case of both sequential and parallel jobs, we gave the optimal solution for exponentially distributed failure inter-arrival times, which, to the best of our knowledge, is the first rigorous proof that periodic checkpointing is optimal. For non-exponentially distributed failures, we developed a dynamic programming algorithm to maximize the amount of work completed before the next failure, which provides a good heuristic for minimizing the expected execution time. Our work considers various models of job parallelism and of parallel checkpointing overhead. We first performed extensive simulation experiments assuming that failures follow Exponential or Weibull distributions, the latter being more representative of real-world systems. The obtained results not only corroborate our theoretical findings, but also show that our dynamic programming algorithm significantly outperforms previously proposed solutions in the case of Weibull failures. We then performed simulation experiments that use failure logs from production clusters. These results confirmed that our dynamic programming algorithm significantly outperforms existing solutions for real-world clusters.

We have also showed an unexpected result: in some cases, when (i) the platform is sufficiently large, and (ii) the checkpointing costs are sufficiently expensive, or the failures are frequent enough, then one should limit the application parallelism and duplicate tasks, rather than fully parallelize the application on the whole platform. In other words, the expectation of the job duration is smaller with fewer processors! To establish this result we have derived and analyzed several scheduling heuristics.

6.1.12. Scheduling parallel iterative applications on volatile resources

In this work we study the efficient execution of iterative applications onto volatile resources. We studied a master-worker scheduling scheme that trades-off between the speed and the (expected) reliability and availability of enrolled workers. A key feature of this approach is that it uses a realistic communication model that bounds the capacity of the master to serve the workers, which requires the design of sophisticated resource selection strategies. The contribution of this work is twofold. On the theoretical side, we assess the complexity of the problem in its off-line version, i.e., when processor availability behaviors are known in advance. Even with this knowledge, the problem is NP-hard. On the pragmatic side, we proposed several on-line heuristics that were evaluated in simulation while a Markovian model of processor availabilities.

We have started this study with the simple case of iterations composed of independent tasks that can execute asynchronously. Then we have investigated a much more challenging scenario, that of a tightly-coupled application whose tasks steadily communicate throughout the iteration. In this latter scenario, if one processor computing some task fails, all the work executed for current iteration is lost, and the computation of all tasks has to be restarted. Similarly, if one processor of the current configuration is preempted, the computation of all tasks is interrupted. Changing the configuration within an iteration becomes a much riskier decision than with independent tasks.

6.1.13. Tiled QR factorization algorithms

In this work, we have revisited existing algorithms for the QR factorization of rectangular matrices composed of $p \times q$ tiles, where $p \geq q$. We target a shared-memory multi-core processor. Within this framework, we study the critical paths and performance of algorithms such as FIBONACCI and GREEDY, and those found within PLASMA. Although neither is optimal, both are shown to be asymptotically optimal for all matrices of size $p = q^2 f(q)$, where $f$ is any function such that $\lim_{q \to \infty} f = 0$. This novel and important complexity result applies to all matrices where $p$ and $q$ are proportional, $p = \lambda q$, with $\lambda \geq 1$, thereby encompassing many important situations in practice (least squares). We provide an extensive set of experiments that show the superiority of the new algorithms for tall matrices.

We have then extended this work to a distributed-memory environment, that corresponds to clusters of multi-core processors. These platforms make the present and the foreseeable future of high-performance computing. In the context of a cluster of multicores, in order to minimize the number of inter-processor communications (aka, “communication-avoiding” algorithm), it is natural to consider two-level hierarchical reduction trees
composed of an “inter-node” tree which acts on top of “intra-node” trees. At the intra-node level, we propose a hierarchical tree made of three levels: (0) “TS level” for cache-friendliness, (1) “low level” for decoupled highly parallel inter-node reductions, (2) “coupling level” to efficiently resolve interactions between local reductions and global reductions. Our hierarchical algorithm and its implementation are flexible and modular, and can accommodate several kernel types, different distribution layouts, and a variety of reduction trees at all levels, both inter-cluster and intra-cluster. Numerical experiments on a cluster of multicore nodes confirm that each of the four levels of our hierarchical tree contributes to build up performance and build insights on how these levels influence performance and interact within each other. Our implementation of the new algorithm with the Dague scheduling tool significantly outperforms currently available QR factorization softwares for all matrix shapes, thereby bringing a new advance in numerical linear algebra for petascale and exascale platforms.

### 6.1.14. Scheduling malleable tasks and minimizing total weighted flow

Malleable tasks are jobs that can be scheduled with preemptions on a varying number of resources. In this work, we have focused on the special case of work-preserving malleable tasks, for which the area of the allocated resources does not depend on the allocation and is equal to the sequential processing time. Moreover, we have assumed that the number of resources allocated to each task at each time instant is bounded. Although this study concerns malleable task scheduling, we have shown that this is equivalent to the problem of minimizing the makespan of independent tasks distributed among processors, when the data corresponding to tasks is sent using network flows sharing the same bandwidth.

We have considered both the clairvoyant and non-clairvoyant cases, and we have focused on minimizing the weighted sum of completion times. In the weighted non-clairvoyant case, we have proposed an approximation algorithm whose ratio is the same as in the unweighted non-clairvoyant case. In the clairvoyant case, we have provided a normal form for the schedule of such malleable tasks, and proved that any valid schedule can be turned into this normal form, based only on the completion times of the tasks. We have shown that in these normal form schedules, the number of preemptions per task is bounded by 3 on average. At last, we have analyzed the performance of greedy schedules, and proved that optimal schedules are greedy for a special case of homogeneous instances. We conjecture that there exists an optimal greedy schedule for all instances, which would greatly simplify the study of this problem.

### 6.1.15. Parallelizing the construction of the ProDom database

ProDom is a protein domain family database automatically built from a comprehensive analysis of all known protein sequences. ProDom development is headed by Daniel Kahn (INRIA project-team BAMBOO, formerly HELIX). With the protein sequence databases increasing in size at an exponential pace, the parallelization of MkDom2, the algorithm used to build ProDom, has become mandatory (the original sequential version of MkDom2 took 15 months to build the 2006 version of ProDom).

When protein domain families and protein families are built independently, the result may be inconsistent. In order to solve this inconsistency problem, we designed a new algorithm, MPI_MkDom3, that simultaneously builds a clustering in protein domain families and one in protein families. This algorithm mixes the principles of MP_MkDom2 and that of the building of Hogenom. As a proof of concept, we successfully processed all the sequences included in the April 2010 version of the UniProt database, namely 6 118 869 sequences and 2 194 382 846 amino-acids.

### 6.2. Algorithms and Software Architectures for Service Oriented Platforms

**Participants:** Daniel Balouek, Nicolas Bard, Julien Bigot, Yves Caniou, Eddy Caron, Florent Chuffart, Simon Delamare, Frédéric Desprez, Gilles Fedak, Sylvain Gault, Haiwu He, Cristian Klein, Georges Markomanolis, Adrian Muresan, Christian Pérez, Vincent Pichon, Jonathan Rouzaud-Cornabas, Anthony Simonet, José Saray, Bing Tang.
6.2.1. Parallel constraint-based local search

Constraint Programming emerged in the late 1980’s as a successful paradigm to tackle complex combinatorial problems in a declarative manner. It is somehow at the crossroads of combinatorial optimization, constraint satisfaction problems (CSP), declarative programming language and SAT problems (boolean constraint solvers and verification tools). Up to now, the only parallel method to solve optimization problems being deployed at large scale is the classical branch and bound, because it does not require much information to be communicated between parallel processes (basically: the current bound).

Adaptive Search was proposed by [86], [87] as a generic, domain-independent constraint-based local search method. This meta-heuristic takes advantage of the structure of the problem in terms of constraints and variables and can guide the search more precisely than a single global cost function to optimize, such as for instance the number of violated constraints. A parallelization of this algorithm based on threads realized on IBM BladeCenter with 16 Cell/BE cores show nearly ideal linear speed-ups for a variety of classical CSP benchmarks (magic squares, all-interval series, perfect square packing, etc.).

We parallelized the algorithm using the multi-start approach and realized experiments on the HA8000 machine, an Hitachi supercomputer with a maximum of nearly 16000 cores installed at University of Tokyo, and on the Grid’5000 infrastructure, the French national Grid for the research, which contains 8612 cores deployed on 11 sites distributed in France. Results show that speedups may surprisingly be architecture and problem dependant. Work in progress considers communications between each computing resource, and a new problem (costa) has been tested for its capability to have an exponential distribution of its time to complete on a sequential resolution.

6.2.2. Service Discovery in Peer-to-Peer environments

In 2010 we experimentally validated the scalability of the Spades Based Middleware (SBAM). SBAM is an auto-stabilized P2P middleware designed for the service discovery. The context of this development is the ANR SPADES project (see Section 7.2.2). In 2011, we wanted to guaranty truthfulness of information exchanged between SBAM-agents. In this context, the implementation of an efficient mechanism ensuring quality of large scale service discovery became a challenge. In collaboration with LIP6 team we developed a self stabilized model called CoPIF and we implemented it in SBAM using synchronous message exchange between agents. Indeed, when a node has to read its neighbor states, it sends a message to each and wait all response. Despite the fact that this kind of implementation is expensive, especially on a large distributed data structure, experiment shown that our model implementation stay efficient, even on a huge prefix tree. We use this broadcast mechanism not only to check the truthfulness of the distributed data structure but also to propagate activation of services on the entire SPADES platform. For the end of 2011 and the beginning of 2012 we plan to work on experimental evaluation of a self-stabilization inspired fault tolerance mechanism. We do this through a collaboration with Myriads team at Rennes.

Moreover, in the occasion of demonstration session of IEEE P2P’2011, we introduced the feasibility of multisite resources aggregation, thanks to SBAM, we ran SBAM on up to 200 peers (we generated machine volatility in order to show the self-stabilization) on 50 physical nodes of Grid’5000 to demonstrate the scalability of multi sites, self-stabilization good performance of our P2P middleware SBAM.

6.2.3. Décrypthon

In 2011, The DIET WebBoard (a web interface to manage the Décrypthon Grid through the DIET middleware) only received bugfixes and a few new features: the possibility to use a totally customized command to call the DIET client, improved support for multiprocessor tasks, and a basic support for replication of tasks (possibility to launch “clones” of an important task, in order to increase the probability of having a successful result). We deployed the new versions of the DIET Webboard on the Décrypthon university grid whenever we made changes to it.

In 2011, we started to port the Rhénovia application (a neuron simulation program in Java and python) on the Décrypthon grid.
The “Help cure muscular dystrophy, phase 2” program that we submitted to the world community grid was still in progress, we received large amounts of result files every day. We had to do the sorting of these files, checking, compressing and moving them to a long term storage space on a regular basis. We also made statistics for the internet users: http://graal.ens-lyon.fr/~nbard/WCGStats/. The last update was on 2011 June 27th: 76.67%.

6.2.4. Scheduling Applications with a Complex Structure

Non-predictably evolving applications are applications that change their resource requirements during execution. These applications exist, for example, as a result of using adaptive numeric methods, such as adaptive mesh refinement and adaptive particle methods. Increasing interest is being shown to have such applications acquire resources on the fly. However, current HPC Resource Management Systems (RMSs) only allow a static allocation of resources, which cannot be changed after it started. Therefore, non-predictably evolving applications cannot make efficient use of HPC resources, being forced to make an allocation based on their maximum expected requirements.

In 2011, we have revisited CooRM, an RMS targeting moldable application, and extended it to CooRMv2, an RMS which supports efficient scheduling of non-predictably evolving applications. An application can make “pre-allocations” to specify its peak resource usage. The application can then dynamically allocate resources as long as the pre-allocation is not outgrown. Resources which are pre-allocated but not used, can be filled by other applications. Results show that the approach is feasible and leads to a more efficient resource usage while guaranteeing that resource allocations are always satisfied.

As future work, we plan to extend CooRMv2 for non-homogeneous clusters, for example, for supercomputers that feature a non-homogeneous network. Moreover, we would like to apply the concepts proposed by CooRMv2 to large scale resource managers such as XtremOS.

6.2.5. High Level Component Model

Most software component models focus on the reuse of existing pieces of code called primitive components. There are however many other elements that can be reused in component-based applications. Partial assemblies of components, well defined interactions between components and existing composition patterns (a.k.a. software skeletons) are examples of such reusable elements. It turns out that such elements of reuse are important for parallel and distributed applications. Therefore, we have designed High Level Component Model (HLCM), a software component model that supports the reuse of these elements thanks to the concepts of hierarchy, genericity and connectors—and in particular the novel concepts of open connection.

In 2011, we have developed two specific implementations of HLCM: L2C for for C++, MPI and CORBA based applications and GLUON++ for CHARM++ based applications in collaboration with Prof. Kale’s team at the University of Illinois at Urbana-Champaign. L2C was used to study how HLCM may simplify the development of domain decomposition applications. GLUON++ was in particular used to study the performance portability of FFT library on various kind of machines. Moreover, on going work includes the study of the benefit of HLCM for MapReduce applications.

6.2.6. Simplifying Code-Coupling in the SALOME platform

The SALOME platform is a generic platform for pre- and post-processing for numerical simulations. It is made of modules which are themselves a set of components. YACS is the module responsible for coupling applications, based on spatial and temporal relationships. The coupling of domain decomposition code, such as the coupling of several instances of Code_Aster, a thermomechanical calculation code from EDF R&D, turns out to be a complex task because of the lack of abstraction of current SALOME model.

In 2011, we have proposed and implemented some extensions to the SALOME model and platform to remove this limitation. The main extension is the ability to express the cloning of a service, which generates also the cloning of connections. The actual semantic of the cloning operation has been specified in function of the nature of the service (sequential, parallel) and of the ports (data or control flow). It has greatly simplified the expression of the coupling of several instances of Code_Aster without generating any measurable overhead at runtime: no more recompilation is needed when varying the number of coupled instances.
6.2.7. Towards Data Desktop Grid

Desktop Grids use the computing, network and storage resources from idle desktop PC’s distributed over multiple-LAN’s or the Internet to compute a large variety of resource-demanding distributed applications. While these applications need to access, compute, store and circulate large volumes of data, little attention has been paid to data management in such large-scale, dynamic, heterogeneous, volatile and highly distributed Grids. In most cases, data management relies on ad-hoc solutions, and providing a general approach is still a challenging issue.

We have proposed the BitDew framework which addresses the issue of how to design a programmable environment for automatic and transparent data management on computational Desktop Grids. BitDew relies on a specific set of meta-data to drive key data management operations, namely life cycle, distribution, placement, replication and fault-tolerance with a high level of abstraction.

Since July 2010, in collaboration with the University of Sfax, we are developing a data-aware and parallel version of Magik, an application for arabic writing recognition using the BitDew middleware. We are targeting digital libraries, which require distributed computing infrastructure to store the large number of digitized books as raw images and at the same time to perform automatic processing of these documents such as OCR, translation, indexing, searching, etc.

In 2011, we have surveyed P2P strategies (replication, erasure code, replica repair, hybrid storage), which provides reliable and durable storage on top of hybrid distributed infrastructures composed of volatile and stable storage. Following this simulation studies, we are implementing a prototype of the Amazon S3 storage on top of BitDew, which will provide reliable storage by using both Desktop free disk space and volunteered remote Cloud storage.

6.2.8. MapReduce programming model for Desktop Grid

MapReduce is an emerging programming model for data-intense application proposed by Google, which has recently attracted a lot of attention. MapReduce borrows from functional programming, where programmer defines Map and Reduce tasks executed on large sets of distributed data. In 2010, we have developed an implementation of the MapReduce programming model based on the BitDew middleware. Our prototype features several optimizations which make our approach suitable for large scale and loosely connected Internet Desktop Grid: massive fault tolerance, replica management, barriers-free execution, latency-hiding optimization as well as distributed result checking. We have presented performance evaluations of the prototype both against micro-benchmarks and real MapReduce applications. The scalability test shows that we achieve linear speedup on the classical WordCount benchmark. Several scenarios involving lagger hosts and host crashes demonstrate that the prototype is able to cope with an experimental context similar to real-world Internet.

In collaboration with the Huazhong University of Science & Technology, we have developed an emulation framework to assess MapReduce on Internet Desktop Grid. We have made extensive comparison on BitDew-MapReduce and Hadoop using Grid5000 which show that our approach has all the properties desirable to cope with an Internet deployment, whereas Hadoop fails on several tests.

In collaboration with the Babes-Bolyai University of Cluj-Napoca, we have proposed a distributed result checker based on the Majority Voting approach. We evaluated the efficiency of our algorithm by computing the aggregated probability with which a MapReduce computation produces an erroneous result.

We have published two chapters in collective books around Cloud and Desktop Grid technologies. The first one, in collaboration with University of Madrid is an introduction to MapReduce and Hadoop, the second one, in collaboration with Virginia Tech is a presentation of two alternative implementations of MapReduce for Desktop Grids : Moon and Bitdew.

6.2.9. SpeQuloS: Providing Quality-of-Service to Desktop Grids using Cloud resources

EDGI is an FP7 European project, following the successful FP7 EDGEs project, whose goal is to build a Grid infrastructure composed of "Desktop Grids", such as BOINC or XtremWeb, where computing resources are...
provided by Internet volunteers, and "Service Grids", where computing resources are provided by institutional Grid such as EGEE, gLite, Unicore and "Clouds systems" such as OpenNebula and Eucalyptus, where resources are provided on-demand. The goal of the EDGI project is to provide an infrastructure where Service Grids are extended with public and institutional Desktop Grids and Clouds.

The main limitation with the current infrastructure is that it cannot give any QoS support for applications running in the Desktop Grid (DG) part of the infrastructure. For example, a public DG system enables clients to return work-unit results in the range of weeks. Although there are EGEE applications (e.g. the fusion community’s applications) that can tolerate such a long latency most of the user communities want much smaller latencies.

In 2011, we have developed the SpeQuloS middleware to solve this critical problem. Providing QoS features even in Service Grids is hard and not solved yet satisfactorily. It is even more difficult in an environment where there are no guaranteed resources. In DG systems, resources can leave the system at any time for a long time or forever even after taking several work-units with the promise of computing them. Our approach is based on the extension of DG systems with Cloud resources. For such critical work-units the SpeQuloS system is able to dynamically deploy fast and trustable clients from some Clouds that are available to support the EDGI DG systems. It takes the right decision about assigning the necessary number of trusted clients and Cloud clients for the QoS applications. At this stage, the prototype is fully developed and validated. It supports the XtremsWeb and BOINC Desktop Grid and OpenNebula, StratusLab, OpenStack and Amazon EC2 Clouds. The first versions have been delivered to the EDGI production infrastructure. We have conducted extensive simulations to evaluate various strategies of Cloud resources provisioning. Results show that SpeQuloS improve the QoS of BoTs on three aspects : it reduces the makespan by removing the tail effect, it improves the execution stability and it allows to accurately predicts the BoT completion time.

6.2.10. Performance evaluation and modeling

Simulation is a popular approach to obtain objective performance indicators of platforms that are not at one’s disposal. It may for example help the dimensioning of compute clusters in large computing centers. In many cases, the execution of a distributed application does not behave as expected, it is thus necessary to understand what causes this strange behavior. Simulation provides the possibility to reproduce experiments under similar conditions. This is a suitable method for experimental validation of a parallel or distributed application.

The tracing instrumentation of a profiling tool is the ability to save all the information about the execution of an application at run-time. Every scientific application executed computes instructions. The originality of our approach is that we measure the completed instructions of the application and not its execution time. This means that if a distributed application is executed on N cores and we execute it again by mapping two processes per core then we need N/2 cores and more time for the execution time of the application. An execution trace of an instrumented application can be transformed into a corresponding list of actions. These actions can then be simulated by SimGrid. Moreover the SimGrid execution traces will contain almost the same data because the only change is the use of half cores but the same number of processes. This does not affect the number of the completed instructions so the simulation time does not get increased because of the overhead. The Grid’5000 platform is used for this work and the NAS Parallel Benchmarks are used to measure the performance of the clusters.

Our main contribution is to propose of a new execution log format that is time-independent. This means that we decouple the acquisition of the traces from the replay. Furthermore we implemented a trace replay tool which relies on top of fast, scalable and validated simulation kernel of SimGrid. We proved that this framework applies for some of the NAS Parallel Benchmarks and we can predict their performance with a good accuracy. Moreover we are working on further improvements for solving some performance issues with the rest benchmarks. We plan to apply some new techniques about the instrumentation of the benchmarks which we have already discussed with people from the performance analysis community and also improve the trace replay tool in order to improve its accuracy. Finally we did a survey on many different tracing tools with regards to the requirements of our methodology which includes all the latest provided tools from the community.
6.2.11. Elastic Scheduling for Functional Workflows

Non-DAG (or functional) workflows are sets of task-graph workflows with non-deterministic transitions between them, that are determined at runtime by special nodes that control the execution flow. In a current work we are focusing on formalizing and evaluating an allocation and scheduling strategy for on-line non-DAG workflows. The goal of this work is to target real-world non-DAG applications and use cloud platforms to perform elastic allocations while keeping cost and stretch fairness constraints.

To address the previous problem we consider each non-DAG workflow as a set of DAG sub-workflows with non-deterministic transitions between them. Whenever an event occurs (a sub-workflow’s execution is completed, a new workflow arrives in the system, a workflow is canceled, etc.) we need to do a rescheduling. The rescheduling strategy considers the currently-running tasks as fixed. Given that the number of events increases proportional to the number of workflows in the system, there is the risk of spending too much time on the scheduling problem and not enough on the workflows themselves. As a result, the scheduling strategy that we will adopt will be a computational inexpensive one, which will give us more room for the number of possible workflows in the system.

This work is currently in the validation step through experimentation with synthetic data. In the near future we will validate against traces of real-world applications that use non-DAG workflows.

6.2.12. Self Adaptive Middleware Deployment

A computer application can be considered as a system of components that exchange information. Each component type has its specific constraints. The application, as a whole, has also its constraints. Deploying an application on a distributed system consist, among other things, to make a mapping between application components and system resources to meet each component constraints, the application constraints, and possibly those set by the the user. Previous work exists on the deployment of middleware, including DIET (with two finished PhD). However, few take into account the issue of redeployment in the event of variation (availability, load, number) of resources. We study this problem of self adaptive deployment of middleware. It consist of achieving an initial deployment, then scrutinizing some changes in the environment, and automatically adjust the deployment (if beneficial) in case of detecting a variation that degrades the performance expected . To do this, we have surveyed the fields of autonomic computing, self adaptive systems and we have defined the different problems that must be solved to achieve this goal. From this, we first define a resource model to represent the physical system, we are to define a model of middleware-based software components, have started the implementation of the resource model to achieve a simulator.

6.2.13. Virtual Machine Placement with Security Requirements

With the number of services using virtualization and clouds growing faster and faster, it is common to mutualize thousands of virtual machines (VMs) within one distributed system. Consequently, the virtualized services, pieces of software and hardware, and infrastructures share the same physical resources. This has given rise to important challenges regarding the security of VMs and the importance of enforcing non-interference between them. Indeed, cross-VM attacks are an important and real world threat. The problem is even worse in the case of adversary users hosted on the same hardware (multi-tenance). Therefore, the isolation facility within clouds needs to be strong. Furthermore, each user has different adversaries and the placement and scheduling processes need to take these adversaries into account.

First, we have worked on resource model to describe distributed system and application model to describe the composition of virtual machine. Then we have formalize isolation requirements between users, between applications and between virtual machines. We also formalized the redundancy requirement. We have created a simulator that can load our resource model and application model. Using it, we have described the Grid’5000 infrastructure and a Virtual Cluster application. We have formalized and implemented an algorithm that takes into account the requirements and place the application. Work in progress considers using Constraint Satisfaction Problems (CSP) and SAT problems to improve the quality of placement. Moreover, we study the trade-off between performance, security requirements and infrastructure consolidation. This works is part of a project on Cloud Security with Alcatel-Lucent Bell Labs and ENSI de Bourges.
6.2.14. Scheduling for MapReduce Based Applications

After a study of the state of the art regarding scheduling, especially scheduling on grid and clouds and MapReduce application scheduling, experiments were performed over the Grid’5000 and Google/IBM Hadoop platforms. We are now working on improving a previous work by Berlinska and Drozdowski which aims at providing a good static schedule of the Map and Reduce phases. A visualization tool has been developed which draws Gantt charts resulting from Berlinska and Drozdowski’s algorithms as well as from our own scheduling heuristics.

A BlobSeer model is also developed in collaboration with the Kerdata research team that will be used for our next developments.

6.3. Parallel Sparse Direct Solvers and Combinatorial Scientific Computing

Participants: Maurice Brémond, Guillaume Joslin, Johannes Langguth, Jean-Yves L’Excellent, Mohamed Sid-Lakhdar, Bora Uçar.

6.3.1. Parallel computation of entries of the inverse of a sparse matrix

Following last year’s work on computing entries of the inverse of a sparse matrix in a serial, in-core or out-of-core environment, and that was implemented in MUMPS, we have pursued work to address this issue in a parallel environment. In such this case, it has been shown that minimizing the number of operations (or the number of accesses to the factors) and balancing the work between the processors are contradictory objectives. Several ideas have been investigated and implemented in order to deal with this issue and to reach high speed-ups. Experimental results are promising and show good speed-ups on relatively small number of processors (up to 16) when dealing with large blocks of sparse right-hand sides, while we used to experience speed-downs before.

6.3.2. Multithreaded parallelism for the MUMPS solver

Apart from using message-passing, we have in the past only exploited multicore parallelism through threaded libraries (e.g. BLAS: Basic Linear Algebra Subroutines), and a few OpenMP directives. We are currently investigating the combination of this fork-join model with threaded parallelism resulting from the task graph, which, in our context, is a tree. To do so, and in order to also target NUMA architectures, we apply ideas from distributed-memory environments to multithreaded environments. Simulations based on benchmarks followed by a first prototype implementation have validated this approach for some classes of matrices on small numbers of cores. We are currently revisiting this implementation and plan to pursue experiments on larger numbers of cores with larger classes of matrices. This starting work was done in the context of a master thesis and is the object of a starting PhD thesis. In a distributed-memory environments, it will be combined with parallelism based on message passing, where the scalability of the existing communication schemes should also be addressed. Both directions will be followed in order to face the multicore (r)evolution.

6.3.3. Low-rank approximations

Low-rank approximations are commonly used to compress the representation of data structures. The loss of information induced is often negligible and can be controlled. Although the dense internal datastructures involved in a multifrontal method, the so-called frontal matrices or fronts, are full-rank, they can be represented by a set of low-rank matrices. Applying to our context the notion of geometric clustering used by Bebendorf to define hierarchical matrices, we have shown that the efficiency of this representation to reduce the complexity of both the factorization and solve phases strongly depends on how variables are grouped. The proposed approach can be used either to accelerate the factorization and solution phases or to build a preconditioner. The ultimate goal of this work is to extend the features of the MUMPS solver to exploit low-rank properties.

This work, and the work described in the two previous paragraphs are in the context of a collaboration with ENSEEIHT-IRIT and with the partners involved in the MUMPS project (see Section 5.2 ).
6.3.4. On partitioning problems with complex objectives

Hypergraph and graph partitioning tools are used to partition work for efficient parallelization of many sparse matrix computations. Most of the time, the objective function that is reduced by these tools relates to reducing the communication requirements, and the balancing constraints satisfied by these tools relate to balancing the work or memory requirements. Sometimes, the objective sought for having balance is a complex function of a partition. We mention some important class of parallel sparse matrix computations that have such balance objectives. For these cases, the current state of the art partitioning tools fall short of being adequate. To the best of our knowledge, there is only a single algorithmic framework in the literature to address such balance objectives. We propose another algorithmic framework to tackle complex objectives and experimentally investigate the proposed framework.

6.3.5. On the Use of Cluster-Based Partial Message Logging to Improve Fault Tolerance for MPI HPC Applications

Fault tolerance is becoming a major concern in HPC systems. The two traditional approaches for message passing applications, coordinated checkpointing and message logging, have severe scalability issues. Coordinated checkpointing protocols make all processes roll back after a failure. Message logging protocols log a huge amount of data and can induce an overhead on communication performance. Hierarchical rollback-recovery protocols based on the combination of coordinated checkpointing and message logging are an alternative. These partial message logging protocols are based on process clustering: only messages between clusters are logged to limit the consequence of a failure to one cluster. These protocols would work efficiently only if one can find clusters of processes in the applications such that the ratio of logged messages is very low. We study the communication patterns of message passing HPC applications to show that partial message logging is suitable in most cases. We propose a partitioning algorithm to find suitable clusters of processes given the communication pattern of an application. Finally, we evaluate the efficiency of partial message logging using two state of the art protocols on a set of representative applications.

6.3.6. Integrated data placement and task assignment for scientific workflows in clouds

We consider the problem of optimizing the execution of data-intensive scientific workflows in the Cloud. We address the problem under the following scenario. The tasks of the workflows communicate through files; the output of a task is used by another task as an input file and if these tasks are assigned on different execution sites, a file transfer is necessary. The output files are to be stored at a site. Each execution site is to be assigned a certain percentage of the files and tasks. These percentages, called target weights, are pre-determined and reflect either user preferences or the storage capacity and computing power of the sites. The aim is to place the data files into and assign the tasks to the execution sites so as to reduce the cost associated with the file transfers, while complying with the target weights. To do this, we model the workflow as a hypergraph and with a hypergraph-partitioning-based formulation, we propose a heuristic which generates data placement and task assignment schemes simultaneously. We report simulation results on a number of real-life and synthetically generated scientific workflows. Our results show that the proposed heuristic is fast, and can find mappings and assignments which reduce file transfers, while respecting the target weights.

6.3.7. UMPa: A Multi-objective, multi-level partitioner for communication minimization

We propose a directed hypergraph model and a refinement heuristic to distribute communicating tasks among the processing units in a distributed memory setting. The aim is to achieve load balance and minimize the maximum data sent by a processing unit. We also take two other communication metrics into account with a tie-breaking scheme. With this approach, task distributions causing an excessive use of network or a bottleneck processor which participates to almost all of the communication are avoided. We show on a large number of problem instances that our model improves the maximum data sent by a processor up to 34% for parallel environments with 4, 16, 64 and 256 processing units compared to the state of the art which only minimizes the total communication volume.
6.3.8. A Divisive clustering technique for maximizing the modularity

We present a new graph clustering algorithm aimed at obtaining clusterings of high modularity. The algorithm pursues a divisive clustering approach and uses established graph partitioning algorithms and techniques to compute recursive bipartitions of the input as well as to refine clusters. Experimental evaluation shows that the modularity scores obtained compare favorably to many previous approaches. In the majority of test cases, the algorithm outperformed the best known alternatives. In particular, among 13 problem instances common in the literature, the proposed algorithm improves the best known modularity in 9 cases.

6.3.9. Constructing elimination trees for sparse unsymmetric matrices

The elimination tree model for sparse unsymmetric matrices and an algorithm for constructing it have been recently proposed [Eisenstat and Liu, SIAM J. Matrix Anal. Appl., 26 (2005) and 29 (2008)]. The construction algorithm has a worst case time complexity $\mathcal{O}(mn)$ for an $n \times n$ unsymmetric matrix having $m$ nonzeros. We propose another algorithm that has a worst case time complexity of $\mathcal{O}(m \log n)$.

6.3.10. Multithreaded clustering for multi-level hypergraph partitioning

Requirements for efficient parallelization of many complex and irregular applications can be cast as a hypergraph partitioning problem. The current-state-of-the art software libraries that provide tool support for the hypergraph partitioning problem are designed and implemented before the game-changing advancements in multi-core computing. Hence, analyzing the structure of those tools for designing multithreaded versions of the algorithms is a crucial task. The most successful partitioning tools are based on the multi-level approach. In this approach, a given hypergraph is coarsened to a much smaller one, a partition is obtained on the the smallest hypergraph, and that partition is projected to the original hypergraph while refining it on the intermediate hypergraphs. The coarsening operation corresponds to clustering the vertices of a hypergraph and is the most time consuming task in a multi-level partitioning tool. We present three efficient multithreaded clustering algorithms which are very suited for multi-level partitioners. We compare their performance with that of the ones currently used in today’s hypergraph partitioners. We show on a large number of real life hypergraphs that our implementations, integrated into a commonly used partitioning library PaToH, achieve good speedups without reducing the clustering quality.

6.3.11. Partitioning, ordering, and load balancing in a hierarchically parallel hybrid linear solver

PDSLin is a general-purpose algebraic parallel hybrid (direct/iterative) linear solver based on the Schur complement method. The most challenging step of the solver is the computation of a preconditioner based on an approximate global Schur complement. We investigate two combinatorial problems to enhance PDSLin’s performance at this step. The first is a multi-constraint partitioning problem to balance the workload while computing the preconditioner in parallel. For this, we describe and evaluate a number of graph and hypergraph partitioning algorithms to satisfy our particular objective and constraints. The second problem is to reorder the sparse right-hand side vectors to improve the data access locality during the parallel solution of a sparse triangular system with multiple right-hand sides. This is needed to eliminate the unknowns associated with the interface in PDSLin. We study two reordering techniques: one based on a postordering of the elimination tree and the other based on a hypergraph partitioning. To demonstrate the effect of these techniques on the performance of PDSLin, we present the numerical results of solving large-scale linear systems arising from numerical simulations of modeling accelerator cavities and of modeling fusion devices.

6.3.12. Experiments on push-relabel-based maximum cardinality matching algorithms for bipartite graphs

We report on careful implementations of several push-relabel-based algorithms for solving the problem of finding a maximum cardinality matching in a bipartite graph and compare them with fast augmenting-path-based algorithms. We analyze the algorithms using a common base for all implementations and compare their relative performance and stability on a wide range of graphs. The effect of a set of known initialization heuristics on the performance of matching algorithms is also investigated. Our results identify a variant of
the push-relabel algorithm and a variant of the augmenting-path-based algorithm as the fastest with proper initialization heuristics, while the push-relabel based one having a better worst case performance.

6.3.13. Towards a scalable hybrid linear solver based on combinatorial algorithms

The availability of large-scale computing platforms comprised of tens of thousands of multicore processors motivates the need for the next generation of highly scalable sparse linear system solvers. These solvers must optimize parallel performance, processor (serial) performance, as well as memory requirements, while being robust across broad classes of applications and systems. In this study, we present a hybrid parallel solver that combines the desirable characteristics of direct methods (robustness) and effective iterative solvers (low computational cost), while alleviating their drawbacks (memory requirements, lack of robustness). We discuss several combinatorial problems that arise in the design of this hybrid solver, present algorithms to solve these combinatorial problems, and demonstrate their impact on a large-scale three-dimensional PDE-constrained optimization problem.
6. New Results

6.1. Communication avoiding algorithms for linear algebra

Participants: Laura Grigori, Simplice Donfack, Amal Khabou, Mathias Jacquelin, Sophie Moufawad.

The focus of this research is on the design of efficient parallel algorithms for solving problems in numerical linear algebra, as solving very large sets of linear equations and large least squares problems, often with millions of rows and columns. These problems arise in many numerical simulations, and solving them is very time consuming.

This research focuses on developing new algorithms for linear algebra problems, that minimize the required communication, in terms of both latency and bandwidth. We have introduced in 2008 two communication avoiding algorithms for computing the LU and QR factorizations, that we refer to as CALU and CAQR (joint work with J. Demmel and M. Hoemmen from U.C. Berkeley, J. Langou from C.U. Denver, and H. Xiang then at INRIA) [6], [9]. Since then, we have also designed a communication avoiding algorithm for rank revealing QR. In addition, we have also extended theoretical lower bounds to sparse Cholesky factorization and identified algorithms that attain these bounds and so minimize communication. The communication avoiding algorithms are now studied by several other groups, including groups at INRIA, and they start being implemented and being available in public libraries as ScaLAPACK.

During 2011, our research has focused on a study of the stability of communication avoiding LU factorization and on its implementation on multicore machines. In [20] we focus on numerical properties of CALU. To decrease the communication required in the LU factorization, CALU uses a new pivoting strategy, referred to as tournament pivoting, that may lead to a different row permutation than the classic LU factorization with partial pivoting. We have further investigated the numerical stability of CALU. The reason to consider CALU is that it does an optimal amount of communication, and asymptotically less than Gaussian elimination with partial pivoting (GEPP), and so will be much faster on platforms where communication is expensive, as shown in previous work. We show that the Schur complement obtained after each step of performing CALU on a matrix A is the same as the Schur complement obtained after performing GEPP on a larger matrix whose entries are the same as the entries of A (sometimes slightly perturbed) and zeros. More generally, the entire CALU process is equivalent to GEPP on a large, but very sparse matrix, formed by entries of A and zeros. Hence we expect that CALU will behave as GEPP and it will be also very stable in practice. In addition, extensive experiments on random matrices and a set of special matrices show that CALU is stable in practice. The upper bound on the growth factor of CALU is worse than of GEPP. However, there are Wilkinson like-matrices for which GEPP has exponential growth factor, but not CALU, and vice-versa.

We present experimental results for random matrices and for a set of special matrices, including sparse matrices, for binary tree based and flat-tree-based CALU. We discuss both the stability of the LU factorization and of the linear solver, in terms of pivot growth and backward errors. The results show that in practice CALU is stable. We present the backward errors measured three ways: by \(\|PA - LU\|/\|A\|\), by the normwise backward error \(\|Ax - b\|/(\|A\|\|x\| + \|b\|)\), and by the componentwise backward error (after iterative refinement in working precision). For random matrices, all CALU’s backward errors were at most 1.9x larger than GEPP’s backward errors. We also tested "special" matrices, including known difficult examples: (1) The ratios of \(\|PA - LU\|/\|A\|\) were at most 1 in over 69% of cases (i.e. CALU was at least as stable as GEPP), and always 1.5 or smaller, except for one ratio of 4.3, in which case both backward errors were much smaller than \(2^{-53} = \) machine epsilon. (2) The ratios of normwise backward errors were at most 1 in over 53% of cases, and always 1.5 or smaller, except for 5 ratios ranging up to 26, in which cases all backward errors were much smaller than 4x machine epsilon. (3) The ratios of componentwise backward errors were at most 1 in over 52% of cases, and always 3.2 or smaller, except for one ratio of 8.3.
In [30] we design a scheduling algorithm for efficiently executing CALU on multicore architectures. We focus on a tunable scheduling strategy that maintains load balance across cores while also maintaining data locality and low dequeue overhead. To achieve this, we use a strategy that combines static and dynamic scheduling. This approach was shown to be successful on regular mesh computations by V. Kale and B. Gropp. This tunable scheduling strategy allows us to flexibly control the percentage of tasks that can be scheduled dynamically; this gives a knob to control load balancing so that it occurs only at the point in computation when the benefits it provides outweighs the costs it induces. On NUMA machines where remote memory access is costly, the percentage of work scheduled dynamically should be small enough to avoid excessive cache misses, but large enough to keep the cores busy during idle times in the static part.

In this work, we show the effectiveness of this method in the context of already highly-optimized dense matrix factorizations. Our prior work on multi-threaded CALU was based on dynamic scheduling. The algorithm performed well on tall and skinny matrices, but became less scalable on square matrices with increasing numbers of processors. We show that the usage of this scheduling in communication avoiding dense factorization leads to significant performance gains. On a 48 core AMD Opteron NUMA machine, our experiments show that we can achieve up to 64% improvement over a version of CALU that uses fully dynamic scheduling, and up to 30% improvement over the version of CALU that uses fully static scheduling. On a 16-core Intel Xeon machine, our hybrid static/dynamic scheduling approach is up to 8% faster than the version of CALU that uses a fully static scheduling or fully dynamic scheduling. Our algorithm leads to speedups over the corresponding routines for computing LU factorization in well known libraries. On the 48 core AMD NUMA machine, our best implementation is up to 110% faster than MKL, while on the 16 core Intel Xeon machine, it is up to 82% faster than MKL. Our approach also shows significant speedups compared with PLASMA on both of these systems.

6.2. Preconditioning techniques for solving large systems of equations

Participants: Laura Grigori, Riadh Fezzanni, Sophie Moufawad.

A different direction of research is related to preconditioning large sparse linear systems of equations. This research is performed in the context of ANR PETALh project (2011-2012), which follows the ANR PETAL project (2008-2009). It is conducted in collaboration with Frederic Nataf from University Paris 6.

Several highly used preconditioners are for example the incomplete LU factorizations and Schwarz based approaches as used in domain decomposition. Most of these preconditioners are known to have scalability problems. The number of iterations can increase significantly when the size of the problem increases or when the number of independent domains is increased. This is often due to the presence of several low frequency modes that hinder the convergence of the iterative method. To address this problem, we study a different class of preconditioners, called direction preserving or filtering preconditioners. These preconditioners have the property of being identical to the input matrix on a given filtering vector. A judicious choice of the vector allows to alleviate the effect of low frequency modes on the convergence.

We consider in particular two classes of preconditioners. The first preconditioner is an incomplete decomposition that satisfies the filtering property [11]. The nested preconditioner has the same property for a specific vector of all ones. However the construction is different and takes advantage of a nested structure of the input matrix. The previous research on these methods considered only matrices arising from the discretization of PDEs on structured grids, where the matrix has a block tridiagonal structure. This structure imposes a sequential computation of the preconditioner and it is not suitable for the more general case of unstructured grids. Hence, while very efficient, the usage of these preconditioners was very limited. At the beginning of this research we have obtained several theoretical results for these methods that demonstrate their numerical behavior and convergence properties for cases arising from the discretization of PDEs on structured grids [11]. But the main result is the development of a generalized method [48], [46] that has two important properties: it allows the filtering property to be satisfied for any input matrix; the matrix can be reordered such that its computation is highly parallel. Experimental results show that the method is very efficient for certain classes of matrices, and shows good scalability results in terms of both problem size and number of processors.
6.3. Microwave Data Analysis for petaScale computers

**Participants:** Laura Grigori, Mikolaj Szydlarski, Meisam Sharify.

In [47] we describe a scalable algorithm for computing an inverse spherical harmonic transform suitable for a cluster of multiple CPU-GPUs. We base our implementation on hybrid programming combining MPI and CUDA. We focus our attention on the two major sequential steps involved in the transforms computation, retaining the efficient parallel framework of the original code. We detail optimization techniques used to enhance the performance of the OpenMP/CUDA-based code and compare them with those implemented in the public domain parallel package, S2HAT.

We also present performance comparisons of the multi GPU version and a hybrid, MPI/OpenMP version of the same transform. We find that one NVIDIA Tesla S1070 can accelerate overall execution time of the SHT by as much as 3 times with respect to the MPI/OpenMP version executed on one quad-core processor (Intel Nehalem 2.93 GHz) and, owing to very good scalability of both versions, 128 Tesla cards perform as good as 256 twelve-core processor (AMD Opteron 2.1 GHz).

The work presented here has been performed in the context of the Cosmic Microwave Background simulations and analysis. However, we expect that the developed software will be of more general interest and applicability.

6.4. Innovative linear system solvers for hybrid multicore/GPU architectures

**Participant:** Marc Baboulin.

The advent of new processor architectures (e.g. multicore, GPUs) requires the rethinking of most of the scientific applications and innovative methods must be proposed in order to take full advantage of current supercomputers [12].

To accelerate linear algebra solvers on current parallel machines, we introduced in public domain libraries a class of solvers based on statistical techniques. A first application concerns the solution of a square linear system $Ax = b$. We study a random transformation of $A$ that enables us to avoid pivoting and then to reduce the amount of communication [54]. Numerical experiments show that this randomization can be performed at a very affordable computational price while providing us with a satisfying accuracy when compared to partial pivoting. This random transformation called Partial Random Butterfly Transformation (PRBT) is optimized in terms of data storage and flops count. In the solver that we developed, PRBT combined with LU factorization with no pivoting take advantage of the latest generation of hybrid multicore/GPU machines and outperform existing factorization routines from current parallel library MAGMA.

A second application is related to solving symmetric indefinite systems via $LDL^T$ factorization for which there was no existing parallel implementation in the dense library ScaLAPACK. We developed an efficient and innovative parallel tiled algorithm for solving symmetric indefinite systems on multicore architectures [59] & [1]. This solver avoids pivoting by using a multiplicative preconditioning based on symmetric randomization. This randomization prevents the communication overhead due to pivoting, is computationally inexpensive and requires very little storage. Following randomization, a tiled LDLT factorization is used that reduces synchronization by using static or dynamic scheduling. We compare Gflop/s performance of our solver with other types of factorizations on a current multicore machine and we provide tests on accuracy using LAPACK test cases.

6.5. MILEPOST GCC: machine learning enabled self-tuning compiler

**Participant:** Grigori Fursin [correspondant].

Tuning compiler optimizations for rapidly evolving hardware makes porting and extending an optimizing compiler for each new platform extremely challenging. Iterative optimization is a popular approach to adapting programs to a new architecture automatically using feedback-directed compilation. However, the large number of evaluations required for each program has prevented iterative compilation from widespread take-up in production compilers. Machine learning has been proposed to tune optimizations across programs systematically but is currently limited to a few transformations, long training phases and critically lacks publicly released, stable tools.
Our approach is to develop a modular, extensible, self-tuning optimization infrastructure to automatically learn the best optimizations across multiple programs and architectures based on the correlation between program features, run-time behavior and optimizations. In this paper we describe MILEPOST GCC, the first publicly-available open-source machine learning-based compiler. It consists of an Interactive Compilation Interface (ICI) and plugins to extract program features and exchange optimization data with the cTuning.org open public repository. It automatically adapts the internal optimization heuristic at function-level granularity to improve execution time, code size and compilation time of a new program on a given architecture. Part of the MILEPOST technology together with low-level ICI-inspired plugin framework is now included in the mainline GCC.

We developed machine learning plugins based on probabilistic and transductive approaches to predict good combinations of optimizations. Our preliminary experimental results show that it is possible to automatically reduce the execution time of individual MiBench programs on various machines from GRID5000, some by more than a factor of 2, while also improving compilation time and code size. We also present a realistic multi-objective optimization scenario for Berkeley DB library using MILEPOST GCC and improve execution time by approximately 17%, while reducing compilation time and code size by 12% and 7% respectively on Intel Xeon processor.

6.6. Loop Transformations: Convexity, Pruning and Optimization

Participant: Cédric Bastoul.

High-level loop transformations are a key instrument in mapping computational kernels to effectively exploit resources in modern processor architectures. However, determining appropriate compositions of loop transformations to achieve this remains a significantly challenging task; current compilers may achieve significantly lower performance than hand-optimized programs. To address this fundamental challenge, we first present a convex characterization of all distinct, semantics-preserving, multidimensional affine transformations. We then bring together algebraic, algorithmic, and performance analysis results to design a tractable optimization algorithm over this highly expressive space. The framework has been implemented and validated experimentally on a representative set of benchmarks run on state-of-the-art multi-core platforms.

6.7. Exact algorithm for the l1-compressive sensing problem using a modified Dantzig- Wolfe method

Participants: Alexandre Borghi, Jerome Darbon, Sylvain Peyronnet.

In this work, we consider the l1-Compressive Sensing problem and presents an efficient algorithm that computes an exact solution. The idea consists in reformulating the problem such that it yields a modified Dantzig-Wolfe decomposition that allows to efficiently apply all standard simplex pivoting rules. Experimental results show the superiority of our approach compared to standard linear programming methods.

6.8. Supple: a flexible probabilistic data dissemination protocol for wireless sensor networks

Participants: Aline Carneiro Viana, Thomas Hérault, Thomas LArgillier, Sylvain Peyronnet, Fatiha Zaidi.

We propose a flexible proactive data dissemination approach for data gathering in self-organized Wireless Sensor Networks (WSN). Our protocol Supple, effectively distributes and stores monitored data in WSNs such that it can be later sent to or retrieved by a sink. Supple empowers sensors with the ability to make on the fly forwarding and data storing decisions and relies on flexible and self-organizing selection criteria, which can follow any predefined distribution law. Using formal analysis and simulation, we show that Supple is effective in selecting storing nodes that respect the predefined distribution criterion with low overhead and limited network knowledge.
6.9. Non-self-stabilizing and self-stabilizing gathering in networks of mobile agents—the notion of speed

Participants: Joffroy Beauquier, Janna Burman, Julien Clément, Shay Kutten.

In the population protocol model, each agent is represented by a finite state machine. Agents are anonymous and supposed to move in an asynchronous way. When two agents come into range of each other (“meet”), they can exchange information. One of the vast variety of motivating examples to the population protocols model is ZebraNet. ZebraNet is a habitat monitoring application where sensors are attached to zebras and collect biometric data (e.g. heart rate, body temperature) and information about their behavior and migration patterns (via GPS). The population protocol model is, in some sense, related to cloud computing and to networks characterized by asynchrony, large scale, the possibility of failures, in the agents as well as in the communications, with the constraint that each agent is resource limited.

In order to extend the computation power and efficiency of the population protocol model, various extensions were suggested. Our contribution is an extension of the population protocol model that introduces the notion of “speed”, in order to capture the fact that the mobile agents move at different speeds and/or have different communication ranges and/or move according to different patterns and/or visit different places with different frequencies. Intuitively, fast agents which carry sensors with big communication ranges communicate with other agents more frequently than other agents do. This notion is formalized by allocating a cover time, cv, to each mobile agent v. cv is the minimum number of events in the whole system that occur before agent v meets every other agent at least once. As a fundamental example, we have considered the basic problem of gathering information that is distributed among anonymous mobile agents and where the number of agents is unknown. Each mobile agent owns a sensed input value and the goal is to communicate the values (as a multi-set, one value per mobile agent) to a fixed non-mobile base station (BS), with no duplicates or losses.

Gathering is a building block for many monitoring applications in networks of mobile agents. For example, a solution to this problem can solve a transaction commit/abort task in MANETs, if the input values of agents are votes (and the number of agents is known to BS). Moreover, the gathering problem can be viewed as a formulation of the routing problem in Disruption Tolerant Networks.

We gave different solutions to the gathering in the model of mobile agents with speed and we proved that one of them is optimal.

6.10. Making Population Protocols Self-stabilizing

Participants: Joffroy Beauquier, Janna Burman, Shay Kutten, Brigitte Rozoy.

As stated in the previous paragraph, the application domains of the population protocol model are asynchronous large scale networks, in which failures are possible and must be taken into account. This work concerns failures and namely the technique of self-stabilization for tolerating them.

Developing self-stabilizing solutions (and proving them) is considered to be more challenging and complicated than developing classical solutions, where a proper initialization of the variables can be assumed. This remark holds for a large variety of models and hence, to ease the task of the developers, some automatic techniques have been proposed to transform programs into self-stabilizing ones.

We have proposed such a transformer for algorithms in the population protocol model introduced for dealing with resource-limited mobile agents. The model we consider is a variation of the original one in that there is a non mobile agent, the base station, and that the communication characteristics (e.g. moving speed, communication radius) of the agents are considered through the notion of cover time.

The automatic transformer takes as an input an algorithm solving a static problem and outputs a self-stabilizing solution for the same problem. To the best of our knowledge, it is the first time that such a transformer for self-stabilization is presented in the framework of population protocols. We prove that the transformer we propose is correct and we make the complexity analysis of the stabilization time.
6.11. Self-stabilizing synchronization in population protocols with cover times

Participants: Joffroy Beauquier, Janna Burman, Shay Kutten, Brigitte Rozoy.

Synchronization is widely considered as an important service in distributed systems which may simplify protocol design. Phase clock is a general synchronization tool that provides a form of a logical time. We have developed a self-stabilizing phase clock algorithm suited to the model of population protocols with cover time. We have shown that a phase clock is impossible in the model with only constant-state agents. Hence, we assumed an existence of resource unlimited agent - the base station. The clock size and duration of each phase of the proposed phase clock tool are adjustable by the user. We provided application examples of this tool and demonstrate how it can simplify the design of protocols. In particular, it yields a solution to Group Mutual Exclusion problem.

6.12. Impossibility of consensus for population protocol with cover times

Participants: Joffroy Beauquier, Janna Burman.

We have extended the impossibility result for asynchronous consensus of Fischer, Lynch and Paterson (FLP) to the asynchronous model of population protocols with cover times. We noted that the proof of FLP does not apply. Indeed, the key lemma stating that two successive factors in an execution, involving disjoint subsets of agents, commute, is no longer true, because of the cover time property. Then we developed a completely different approach and we proved that there is no general solution to consensus for population protocols with cover times, even if there is a single possible crash. We noted that this impossibility result also applies to randomized asynchronous consensus, contrary to what happens in the classical message-passing or shared memory communication models, in which the problem is solvable inside some bounds on the number of faulty processes. Then, for circumventing these impossibility results, we introduced the phase clock oracle and the S oracle, and we shown how they allow to design solutions.

6.13. Routing and synchronization in large scale networks of very cheap mobile sensors

Participants: Joffroy Beauquier, Brigitte Rozoy.

In a next future, large networks of very cheap mobile sensors will be deployed for various applications, going from wild life preserving or environmental monitoring up to medical or industrial system control. Each sensor will cost only a few euros, allowing a large scale deployment. They will have only a few bit of memory, no identifier, weak capacities of computation and communication, no real time clock and will be prone to failures. Moreover such networks will be fundamentally dynamic. The goal of this subject is to develop the basic protocols and algorithms for rudimentary distributed systems for such networks. The studied problems are basic ones, like data collection, synchronization (phase clock, mutual exclusion, group mutual exclusion), fault tolerance (consensus), automatic transformers, always in a context of possible failures. A well known model has already been proposed for such networks, the population protocol model. In this model, each sensor is represented by a finite state machine. Sensors are anonymous and move in an asynchronous way. When two sensors come into range of each other ("meet"), they can exchange information. One of the vast variety of motivating examples for this model is ZebraNet. ZebraNet is a habitat monitoring application in which sensors are attached to zebras in order to collect biometric data (e.g., heart rate, body temperature) and information about their behavior and migration patterns. Each pair of zebras meets from time to time. During such meetings (events), ZebraNet’s agents (zebras’ attached sensors) exchange data. Each agent stores its own sensor data as well as data of other sensors that were in range in the past. They upload data to a base station whenever it is nearby. It was shown that the set of applications that can be solved in the original model of population protocols is rather limited. Other models (such as some models of Delay/Disruption-Tolerant Networks - DTNs), where each node maintains links and connections even to nodes it may interact with only intermittently, do not seem to suit networks with small memory agents and a very large (and unknown) set of anonymous agents. That is why we enhance the model of population protocols by introducing a notion
of "speed". We try to capture the fact that the mobile agents move at different speeds and/or have different communication ranges and/or move according to different patterns and/or visit different places with different frequencies. Intuitively, fast agents which carry sensors with large communication ranges communicate with other agents more frequently than other agents do. This notion is formalized by the notion of cover time for each agent. The cover time of an agent is the unknown number of events (pairwise meetings) in the whole system that occur (during any execution interval) before agent \( v \) meets every other agent at least once. The model we propose is somehow validated by some recent statistical results, obtained from empirical data sets regarding human or animal mobility. An important consequence of our approach is that the analytic complexity of the protocols designed in this model is possible, independently of any simulation or experimentation. For instance, we consider the fundamental problem of gathering different pieces of information, each sensed by a different anonymous mobile agent, and where the number of agents is unknown. The goal is to communicate the sensed values (as a multi-set, one value per mobile agent) to a base station, with no duplicates or losses. Gathering is a building block for many monitoring applications in networks of mobile agents. Moreover, the gathering problem can be viewed as a special case of the routing problem in DTNs, in which there is only one destination, the base station. Then we are able to compute the complexity of solutions we propose, as well as those of solutions used in experimental projects (like ZebraNet), and to compare them. The algorithms we present are self-stabilizing. Such algorithms have the important property of operating correctly regardless of their initial state (except for some bounded period). In practice, self-stabilizing algorithms adjust themselves automatically to any changes or corruptions of the network components (excluding the algorithm’s code). These changes are assumed to cease for some sufficiently long period. Self-stabilization is considered for two reasons. First, mobile agents are generally fragile, subject to failures and hard to initialize. Second, systems of mobile agents are by essence dynamic, some agents leave the system while new ones are introduced. Self-stabilization is a well adapted framework for dealing with such situations.

6.14. Self-Stabilizing Control Infrastructure for HPC

Participants: Thomas Hérault, Camille Coti.

High performance computing platforms are becoming larger, leading to scalability and fault-tolerance issues for both applications and runtime environments (RTE) dedicated to run on such machines. After being deployed, usually following a spanning tree, a RTE needs to build its own communication infrastructure to manage and monitor the tasks of parallel applications. Previous works have demonstrated that the Binomial Graph topology (BMG) is a good candidate as a communication infrastructure for supporting scalable and fault-tolerant RTE.

In this work, we presented and analyzed a self-stabilizing algorithm to transform the underlying communication infrastructure provided by the launching service (usually a tree, due to its scalability during launch time) into a BMG, and maintain it in spite of failures. We demonstrated that this algorithm is scalable, tolerates transient failures, and adapts itself to topology changes.

The algorithms are scalable, in the sense that all process memory, number of established communication links, and size of messages are logarithmic with the number of elements in the system. The number of synchronous rounds to build the system is also logarithmic, and the number of asynchronous rounds in the worst case is square logarithmic with the number of elements in the system. Moreover, the self-stabilizing property of the algorithms presented induce fault-tolerance and self-adaptivity. Performance evaluation based on simulations predicts a fast convergence time (1/33s for 64K nodes), exhibiting the promising properties of such self-stabilizing approach.

We pursue this work by implementing and evaluating the algorithms in the STCI runtime environment to validate the theoretical results.

6.15. Large Scale Peer to Peer Performance Evaluations

Participant: Serge Petiton.
6.15.1. Large Scale Grid Computing

Recent progress has made possible to construct high performance distributed computing environments, such as computational grids and cluster of clusters, which provide access to large scale heterogeneous computational resources. Exploration of novel algorithms and evaluation of performance is a strategic research for the future of computational grid scientific computing for many important applications [88]. We adapted [68] an explicit restarted Lanczos algorithm on a world-wide heterogeneous grid platform. This method computes one or few eigenpairs of a large sparse real symmetric matrix. We take the specificities of computational resources into account and deal with communications over the Internet by means of techniques such as out-of-core and data persistence. We also show that a restarted algorithm and the combination of several paradigms of parallelism are interesting in this context. We perform many experimentations using several parameters related to the Lanczos method and the configuration of the platform. Depending on the number of computed Ritz eigenpairs, the results underline how critical the choice of the dimension of the working subspace is. Moreover, the size of platform has to be scaled to the order of the eigenproblem because of communications over the Internet.

6.15.2. High Performance Cluster Computing

Grid computing focuses on making use of a very large amount of resources from a large-scale computing environment. It intends to deliver high-performance computing over distributed platforms for computation and data-intensive applications. We propose [99] an effective parallel hybrid asynchronous method to solve large sparse linear systems by the use of a Grid Computing platform Grid5000. This hybrid method combines a parallel GMRES(m) (Generalized Minimum RESidual) algorithm with the Least Square method that needs some eigenvalues obtained from a parallel Arnoldi algorithm. All of these algorithms run on the different processors of the platform Grid5000. Grid5000, a 5000 CPUs nation-wide infrastructure for research in Grid computing, is designed to provide a scientific tool for computing. We discuss the performances of this hybrid method deployed on Grid5000, and compare these performances with those on the IBM SP series supercomputers.

6.15.3. Large Scale Power aware Computing

Energy conservation is a dynamic topic of research in High Performance Computing and Cluster Computing. Power-aware computing for heterogeneous world-wide Grid is a new track of research. We have studied and evaluated the impact of the heterogeneity of the computing nodes of a Grid platform on the energy consumption. We propose to take advantage of the slack-time caused by the heterogeneity in order to save energy with no significant loss of performance by using Dynamic Voltage Scaling (DVS) in a distributed eigensolver [69]. We show that using DVS only during the slack-time does not penalize the performances but it does not provide significant energy savings. If DVS is applied to all the execution, we get important global and local energy savings (respectively up to 9% and 20%) without a significant rise of the wall-clock times.

6.16. High Performance Linear Algebra on the Grid

Participants: Thomas Hérault, Camille Coti.

Previous studies have reported that common dense linear algebra operations do not achieve speed up by using multiple geographical sites of a computational grid. Because such operations are the building blocks of most scientific applications, conventional supercomputers are still strongly predominant in high-performance computing and the use of grids for speeding up large-scale scientific problems is limited to applications exhibiting parallelism at a higher level.

In this work, we have identified two performance bottlenecks in the distributed memory algorithms implemented in ScaLAPACK, a state-of-the-art dense linear algebra library. First, because ScaLAPACK assumes a homogeneous communication network, the implementations of ScaLAPACK algorithms lack locality in their communication pattern. Second, the number of messages sent in the ScaLAPACK algorithms is significantly greater than other algorithms that trade flops for communication.
This year, we presented a new approach for computing a QR factorization one of the main dense linear algebra kernels of tall and skinny matrices in a grid computing environment that overcomes these two bottlenecks. Our contribution is to articulate a recently proposed algorithm (Communication Avoiding QR) with a topology-aware middleware (QCG-OMPI) in order to confine intensive communications (ScalAPACK calls) within the different geographical sites.

An experimental study conducted on the Grid5000 platform shows that the resulting performance increases linearly with the number of geographical sites on large-scale problems (and is in particular consistently higher than ScalAPACKs).

### 6.17. Emulation of Volatile Systems

**Participants:** Thomas Largillier, Benjamin Quetier, Sylvain Peyronnet, Thomas Hérault, Franck Cappello.

In the process of developing grid applications, people need to often evaluate the robustness of their work. Two common approaches are simulation, where one can evaluate his software and predict behaviors under conditions usually unachievable in a laboratory experiment, and experimentation, where the actual application is launched on an actual grid. However, simulation could ignore unpredictable behaviors due to the abstraction done and experimentation does not guarantee a controlled and reproducible environment, and simulation often introduces a high level of abstraction that makes the discovery and study of unexpected, but real, behaviors a rare event.

In this work, we proposed an emulation platform for parallel and distributed systems including grids where both the machines and the network are virtualized at a low level. The use of virtual machines allows us to test highly accurate failure injection since we can destroy virtual machines, and network virtualization provides low-level network emulation. Failure accuracy is a criteria that evaluates how realistic a fault is. The accuracy of our framework has been evaluated through a set of micro benchmarks and a very stable P2P system called Pastry.

We are in the process of developing a fault injection tool to work with the platform. It will be an extension of the work started in the tool Fail. The interest of this work is that using Xen virtual machines will allow to model strong adversaries since it is possible to have virtual machines with shared memory. These adversaries will be stronger since they will be able to use global fault injection strategies.

### 6.18. Exascale Systems

**Participant:** Franck Cappello.

Over the last 20 years, the open-source community has provided more and more software on which the world’s high-performance computing systems depend for performance and productivity. The community has invested millions of dollars and years of effort to build key components. Although the investments in these separate software elements have been tremendously valuable, a great deal of productivity has also been lost because of the lack of planning, coordination, and key integration of technologies necessary to make them work together smoothly and efficiently, both within individual petascale systems and between different systems. A repository gatekeeper and an email discussion list can coordinate open-source development within a single project, but there is no global mechanism working across the community to identify critical holes in the overall software environment, spot opportunities for beneficial integration, or specify requirements for more careful coordination. It seems clear that this completely uncoordinated development model will not provide the software needed to support the unprecedented parallelism required for peta/exascale computation on millions of cores, or the flexibility required to exploit new hardware models and features, such as transactional memory, speculative execution, and GPUs. We presented a rational promoting that the community must work together to prepare for the challenges of exascale computing, ultimately combing their efforts in a coordinated International Exascale Software Project.
Over the past few years resilience has became a major issue for high-performance computing (HPC) systems, in particular in the perspective of large petascale systems and future exascale systems. These systems will typically gather from half a million to several millions of central processing unit (CPU) cores running up to a billion threads. From the current knowledge and observations of existing large systems, it is anticipated that exascale systems will experience various kind of faults many times per day. It is also anticipated that the current approach for resilience, which relies on automatic or application level checkpoint/restart, will not work because the time for checkpointing and restarting will exceed the mean time to failure of a full system. This set of projections leaves the community of fault tolerance for HPC systems with a difficult challenge: finding new approaches, which are possibly radically disruptive, to run applications until their normal termination, despite the essentially unstable nature of exascale systems. Yet, the community has only five to six years to solve the problem. In order to start addressing this challenge, we synthesized the motivations, observations and research issues considered as determinant of several complimentary experts of HPC in applications, programming models, distributed systems and system management.

As a first step to adress the resilience challenge, we conducted a comprehensive study of the state of the art. The emergence of petascale systems and the promise of future exascale systems have reinvigorated the community interest in how to manage failures in such systems and ensure that large applications, lasting several hours or tens of hours, are completed successfully. Most of the existing results for several key mechanisms associated with fault tolerance in high-performance computing (HPC) platforms follow the rollback-recovery approach. Over the last decade, these mechanisms have received a lot of attention from the community with different levels of success. Unfortunately, despite their high degree of optimization, existing approaches do not fit well with the challenging evolutions of large-scale systems. There is room and even a need for new approaches. Opportunities may come from different origins: diskless checkpointing, algorithmic-based fault tolerance, proactive operation, speculative execution, software transactional memory, forward recovery, etc.

We provided the following contributions: (1) we summarize and analyze the existing results concerning the failures in large-scale computers and point out the urgent need for drastic improvements or disruptive approaches for fault tolerance in these systems; (2) we sketch most of the known opportunities and analyze their associated limitations; (3) we extract and express the challenges that the HPC community will have to face for addressing the stringent issue of failures in HPC systems.
HIEPACS Project-Team

6. New Results

6.1. Algorithms and high-performance solvers

6.1.1. Dense linear algebra solvers for multicore processors accelerated with multiple GPUs

In collaboration with the Inria RUNTIME team and the University of Tennessee, we have designed dense linear algebra solvers that can fully exploit a node composed of a multicore processor accelerated with multiple GPUs. This work has been integrated in the latest release of the MAGMA package (http://icl.cs.utk.edu/magma/

6.1.2. Hybrid direct/iterative solvers based on algebraic domain decomposition techniques

A first release of the MaPHyS package should be made available early in 2012 thanks to the developments conducted in the last year of the ADT. An approximation of the local Schur complement has been studied that is based on approximated inverse technique. This work is a natural extension of part of the PhD research of Mikko Byckling. Furthermore, during his master internship, Stojce Nakov has investigated the design of

6.1.3. Resilience in numerical simulations

In his master internship work, Mawussi Zounon investigated recovery strategies for core faults in the framework of parallel preconditioned Krylov solvers. The underlying idea is to recover fault entries of the iterate via interpolation from existing values available on neighbor cores. He will continue this work in the framework of his PhD funded by the ANR-RESCUE. Notice that theses activities are also part of our contribution to the G8-ECS (Enabling Climate Simulation at extreme scale).

6.1.4. Full geometric multigrid method for 3D Maxwell equations

In the context of a collaboration with the CEA/CESTA center, Mathieu Chanaud continued his PhD work on a tight combination between multigrid methods and direct methods for the efficient solution of challenging 3D irregular finite element problems arising from the discretization of Maxwell equations. A parallel solver dedicated to the ODYSSEE challenge (electromagnetism) of CEA/CESTA has been implemented and integrated. The novel parallel solver was able to solve a 1.3 billion system given a 20 million unknown problem at the coarsest level. The input mesh defines the coarsest level. This mesh is further refined to defined the grid hierarchy, where matrix free smoothers are considered to reduce the memory consumption.

6.1.5. Scalable numerical schemes for scientific applications

A work is currently carried on with TOTAL (Rached Abdelkhalek PhD). The extraordinary challenge that the oil and gas industry must face for hydrocarbon exploration requires the development of leading edge technologies to recover an accurate representation of the subsurface. Seismic modeling and Reverse Time Migration (RTM) based on the full wave equation discretization, are tools of major importance since they give an accurate representation of complex wave propagation areas. Unfortunately, they are highly compute intensive. The recent development in GPU technologies with unified architecture and general-purpose languages coupled with the high and rapidly increasing performance throughput of these components made General Purpose Processing on Graphics Processing Units an attractive solution to speed up diverse applications. We have designed a fast parallel simulator that solves the acoustic wave equation on a GPU cluster. Solving the acoustic wave equation in an oil exploration industrial context aims at speeding up seismic modeling and Reverse Time Migration. We consider a finite difference approach on a regular mesh, in both
2D and 3D cases. The acoustic wave equation is solved in a constant density or a variable density domain. All the computations are done in single precision, since double precision is not required in our context. We use nvidia CUDA to take advantage of the GPU computational power. We study different implementations and their impact on the application performance. We obtain a speed up of 16 for Reverse Time Migration and up to 43 for the modeling application over a sequential code running on general purpose CPU. The defense of this thesis is planned early 2012.

For the solution of the elastodynamic equation on meshes with local refinments, we are currently collaborating with Total to design a parallel implementation of a local time refinement technique on top of a discontinuous Galerkin space discretization. This latter technique enables to manage non-conforming meshes suited to deal with multiblock approaches that capture the locally refined regions. This work is developed in the framework of Yohann Dudouit PhD thesis. A software prototype is currently developed to address these simulations.

The calculation of acoustic modes in combustion chambers is a challenging calculation for large 3D geometries. It requires the calculation of a few of the smallest eigenpairs of large unsymmetric matrices in a parallel environment. A new block Arnoldi approach is currently developed to best benefit from the continuation scheme used in this application context. This is part of the PhD research activity of Pablo Salas.

6.2. Efficient algorithmics for code coupling in complex simulations

The performance of the coupled codes depends on how the data are well distributed on the processors. Generally, the data distributions of each code are built independently from each other to obtain the best load-balancing. But once the codes are coupled, the naive use of these decompositions can lead to important imbalance in the coupling area. Therefore, the modeling of the whole coupling is crucial to improve the performance and to ensure a good scalability. The goal is to find the best data distribution for the whole coupled codes and not only for each standalone code. The key idea is to use a graph/hypergraph model that will incorporate information about the coupling itself. Then, we propose new algorithms to perform a coupling-aware partitioning in order to improve the load-balancing of the whole coupled simulation.

Let us consider two coupled codes, modeled by two graphs (or hypergraphs) $A$ and $B$, connected by inter-edges $I(A, B)$ that represents the coupling communications between codes. Formally, the problem consists in partitioning $A$ in $M$ and $B$ in $N$ with accounting for $I(A, B)$. This algorithm should optimize both the edge cut for each graph and the coupling communications while maintaining each graph balance. Our general strategy is divided in three main steps:

1. first, we freely partition $A$ in $M$ to obtain the partition $A/M$;
2. then, we projects this partition to $B$ according to $I(A, B)$, that provides the partition $B/M$;
3. finally, we compute the partition $B/N$ by repartitioning $B$ from $M$ existing parts into $N$.

The final repartitioning step is particularly tedious, because it must handle a variable number of processes. However, as far as we know, the state-of-the-art graph/hypergraph repartitioning tools are limited to a fixed number of processes (i.e. $M = N$). To overcome this issue, we have proposed a new repartitioning algorithm – assuming the load is constant – based on hypergraph partitioning technics with fixed vertices. Our algorithm uses an *optimal* communication pattern, that we have proved to minimize the total number of messages between the former and newer parts. Experimental results validate our work comparing it with other approaches [ 20 ]. We currently investigate how to extend our algorithm for the dynamic load-balancing of parallel adaptive codes ($A = B$), whose load evolution is variable and difficult to predict. In this case, it would be convenient to dynamically adapt the number of processes used at runtime ($M \neq N$), while minimizing migration cost during the repartitioning step. This work is currently conducted in the framework of Clément Vuchener PhD thesis.

6.3. Distributed Shared Memory approach for the steering of parallel simulations
As a different approach of EPSN, we conceived and developed an in-transit visualization framework for interfacing an arbitrary HPC simulation code with an interactive ParaView session using the HDF5 parallel IO library as the API. The library called H5FDdsm is coupled with a ParaView plugin ICARUS (Initialize Compute Analyze Render Update Steer).

Because our interface is based on files, stored in a distributed shared memory (DSM), we sought during this year different redistribution strategies to optimize the bandwidth and the transfers between the simulation and the ParaView servers hosting the DSM. This work showed real benefits, particularly on one of our Cray XE6 testing machines using a block cyclic redistribution. On these large HPC machines that do not support the dynamic MPI process management set of functions, we improved our connection system so that simulation and post-processing can be coupled within an MPMD job. Taking also advantage of one-sided communication models and of the Cray Gemini interconnect communication performance, our framework has been sensibly improved and should be optimal in the coming months.

The interface has also been enhanced with a steering interface that allows us to control the simulation workflow and send back not only parameters, but also complete meshes in parallel, which can then be read by the simulation using either our steering interface or HDF5 calls. This has been demonstrated with SPH-flow, a CFD code developed by Ecole Centrale de Nantes and HydrOcean, replacing dynamically and in parallel a falling wedge with a deforming sphere.

This work has been realized and is currently carried on at CSCS - Swiss National Supercomputing Centre in the framework of Jérôme Soumagne PhD thesis (under the co-supervision of Mr. John Biddiscombe) and within the NextMuSE European project 7th FWP/ICT-2007.8.0 ([17], [18], [19]).

6.4. Material physics

6.4.1. Hybrid materials

The study of hybrid materials based on a coupling between molecular dynamics (MD) and quantum mechanism (QM) simulation has been conducted in collaboration with IPREM (Pau) within the ANR CIS 2007 NOSSI (ended December 2011). These simulations are complex and costly and may involve several length scales, quantum effects, components of different kinds (mineral-organic, hydro-philic and -phobic parts). Our goal was to compute dynamical properties of hybrid materials like optical spectra. The computation of optical spectra of molecules and solids is the most consuming time in such coupling. This requires new methods designed for predicting excited states and new algorithms for implementing them. Several tracks have been investigated in the project and new results obtained as described below.

Optical spectra.

Some new improvements in our TD-DFT code have been introduced. Our method is based on the LCAO method for densities and excited states that computes electronic excitation spectra. We have worked in two directions:

- As the method introduces a regularization parameter to obtain regularized spectra we have used it to build better algorithms. In particular, we have developed a new hierarchical algorithm that builds a well adapted frequency distribution to better capture the biggest peaks (strongest oscillator strengths) in the spectrum. Moreover, a nonlinear fit method was added and used to compute the transitions and the oscillator strengths of the spectrum.

- In our algorithm, we used a coarse grain paradigm to parallelize the spectrum computation. This approach leads to a memory bottleneck for large systems. In that respect, we have explored a new parallel approach based on a fine grain paradigm (matrix-vector parallelization) to better exploit the manycore architecture of the emerging computers.

Finally, we have improved the packaging of the code to prepare a public release of the code. Our TD-DFT code will be soon available on request.
**QM/MM algorithm.** For structure studies or dynamical properties, we have coupled QM model based on pseudo-potentials (SIESTA code) with dynamic molecular (DL-POLY code). Therefore we have developed a new algorithm to avoid accounting twice for the forces and the quantum electric field in the molecular model. All algorithms involved in the coupling have been introduced both in SIESTA and in DL-POLY codes. The following new developments needed by the coupling have been introduced in the SIESTA code:

- We have implemented a fast evaluation of the molecular electrostatic field on the quantum grid.
- We have introduced a non periodic Poisson solver based on the parallel linear Hypre solver. This solver allows us to use computation domains as small as possible.
- We have implemented the ElectroStatic Potential (ESP) fit method to obtain more physical point charges than those given by SIESTA with the Mulliken method. These point charges are used by the MM codes to compute electrostatic forces.

Thanks to all our developments introduced in SIESTA a collaboration with the SIESTA research team has started. This enables us to have access to their private svn like repository. Preliminary results on a water dimer and a water box systems show good agreement with other methods developed in SIESTA and DL-POLY teams. All these results were presented in the final international NOSSI workshop in Biarritz on December.

### 6.4.2. Material failures

We have started in the context of the OPTIDIS ANR to work on dislocation simulations. The main characteristic of these simulations is that they are highly dynamical. This year, we have started the study of the state of the art on this topic in two directions. The first direction concerns the study of the algorithms used in such simulations and how we can efficiently parallelize them on manycore clusters. In the second one for isotropic materials, we are investigating how to adapt our fast multipole method to compute constraints and then forces in this kind of simulations.
6. New Results

6.1. Massive mobile dense wireless networks

**Participants:** Cédric Adjih, Aline Carneiro Viana, Emmanuel Baccelli, Philippe Jacquet, Pascale Minet, Paul Mühlethaler, Yasser Toor.

6.1.1. Executive summary

Scaling properties of mobile ad hoc network lead to an increase of global capacity when the network density increases or when the packets can be stored for a while in mobile nodes instead of being immediately retransmitted.

Gupta and Kumar have shown in 2000 that the transport capacity per node in a multihop ad hoc network decreases in $1 / \sqrt{N \log N}$, $N$ being the number of nodes in the network. Therefore the global capacity of the network increases in $\sqrt{N} / \sqrt{\log N}$. This is a surprising result since in wired network a collection of nodes connected to a single communication resource has a transport capacity that just remains constant (i.e. the average per node capacity decreases in $1/N$).

Therefore adding space to a multihop wireless network increases the capacity: this is the space capacity paradox.

When nodes randomly move, it turns to be more advantageous to store packets for a while on mobile routers instead of forwarding them immediately like hot potatoes. When the mobile router moves closer to the destination, then it can delivers packets on a much smaller number of hops. Of course the delivery delay is much longer, but the network capacity also increases by slowing non urgent packets. This is the time capacity paradox: by slowing packets, nodes mobility increases network capacity. This was hinted the first time by Grossglauser and Tse in 2002.

The great challenge is to find the good protocol and tunings that allow to adjust the delivery delay from zero to infinity in order to get a continuous increase in capacity. The challenge is two-sided: one has to keep the delivery delay between reasonable bounds and one has to consider realistic mobility models.

Existing protocols for Mobile Ad Hoc Networks (MANET) are highly efficient in routing data between mobile nodes that belong to the same connected component (cf. the protocols which have received the RFC status by the manet group of IETF). What about a disconnected network where source and destination may be located in two different connected components? In this case usual routing protocols drop packet due to host unreachable as no end-to-end route exists at that time.

A simple idea is to allow the router that has no available route to the destination to keep the packet in buffer until the conditions become more appropriate for forwarding. The forwarding conditions will change because of mobility: the router can move closer to the destination so that they belong to the same connected component and the packet can be delivered.

Indeed, the network may be continuously partitionned due to high mobility, and the traditional approach to allow a mobile node to wait for the network to be fully connected (i.e. form a unique component) or to wait to be in range of the destination may lead to unacceptable delays. Furthermore, concrete applications, such as Defence and Disaster-Relief, cannot always rely on such assumptions.

Nevertheless, even if the communicating nodes may never be within the same connected component, it is important to observe that a “communication path” may be available through time using intermediate nodes that are temporarily within reach of each other while moving, hence making such networks viable for critical applications. Depending on the nature of the environment, these networks are now commonly referred as Intermittently Connected MANET and Delay Tolerant Networks.
In between stands the problem of the fully connected network that forms a single connected component, but for which maintaining full knowledge of the topology would simply make the network collapse under its huge control traffic. In fact this is the main problem that wireless network engineering has to face, in most experiments the generation of control traffic is the main source of disruption.

6.1.2. Scientific achievements

6.1.2.1. Scaling and spatial capacity in non uniform wireless networks

We found a more precise instance of Gupta-Kumar result by using a simple but realistic network model based on slotted ALOHA with Poisson traffic. It turns out that when the traffic density increases then the average node neighborhood area shrinks so that the average encircled traffic load remains constant with an analytical expression.

In their original model Gupta and Kumar assume that the traffic density is constant, which is far from realistic. However we have derived similar generalized results when the traffic density is not uniform. In this case, the heavier is the local traffic, the smaller are the local neighborhood and the larger is the number of hops needed to cross the congested region. Therefore the shortest paths (in hop number as computed by OLSR) will have a natural tendency to avoid congested areas. The path tend to follow trajectory that have analogy in non linear optic with variable indices.

6.1.2.2. Time capacity and node mobility

We have defined a protocol that takes advantage of node mobility in a general way. In short the packet stay with its host router as long as the latter does not evade too fast from its next hop (computed via a shortest path protocol such as OLSR). In the way we understand “too fast” stands the tuning parameters we discussed above. There is no need to have node geographical location and to physically measure motion vector, since everything can be done via the analysis of the dynamic of neighborhood intersections. We analytically derived performance evaluation under random walk mobility models. We plan to simulate the protocol in a real mobility scenario. This algorithm has application in Intelligent Transport System.

6.1.2.3. Overhead reduction in large networks

The first limitation of multihop wireless network is the size of the overhead per node that increases linearly with the size of the network. This is a huge improvement compared to classic internet protocols which have quadratic overhead increases. Nevertheless this till limit the network size to some thousands. We have analyzed the performance of OLSR with Fisheye feature that significantly reduce the overhead with respect to distance. In theory the overhead reduction allows to network size of several order of magnitude. Anyhow the tuning of the overhead attenuation with distance must be carefully done when the network is mobile, in order to avoid tracking failure. We showed that an overhead reduction within square root of the network size achieve this goal.

An alternative way to overhead reduction is ad hoc hierarchical routing and Distributed Hashing Table. Work has just begun in this area.

6.1.2.4. Coloring in wireless networks

Coloring is used in wireless networks to improve communication efficiency, mainly in terms of bandwidth, energy and possibly end-to-end delays. Nodes access the medium according to their color. It is the responsibility of the coloring algorithm to ensure that interfering nodes do not have the same color. First, we established complexity results about the h-hop coloring problem. Second, we focused on wireless sensor networks with grid topologies. We proposed the Vector-Based Coloring Method, denoted VCM, a new method that is able to provide an optimal periodic coloring for any radio transmission range and for any h-hop coloring, h>=1. Third, we also designed OSEVEN "Optimized SchEduling RoutEr Node Activity", a distributed coloring algorithm optimized for dense wireless networks.

6.1.2.5. Complexity results about the h-hop coloring problem

In the paper we published at the WMNC 2011 conference, we define the h-hop node coloring problem, with h any positive integer, adapted to two types of applications in wireless networks. We specify both general mode for general applications and strategic mode for data gathering applications. We prove that the associated decision problem is NP-complete.
6.1.2.6. Grid coloring and the Vector-Based Coloring Method

In 2011, we also focused on wireless sensor networks with grid topologies. How does a coloring algorithm take advantage of the regularity of grid topology to provide an optimal periodic coloring, that is a coloring with the minimum number of colors? We propose the Vector-Based Coloring Method, denoted VCM, a new method that is able to provide an optimal periodic coloring for any radio transmission range and for any h-hop coloring, h>=1. This method consists in determining at which grid nodes a color can be reproduced without creating interferences between these nodes while minimizing the number of colors used. We compare the number of colors provided by VCM with the number of colors obtained by a distributed coloring algorithm with line and column priority assignments. We also provide bounds on the number of colors of optimal general colorings of the infinite grid, and show that periodic colorings (and thus VCM) are asymptotically optimal. Finally, we discuss the applicability of this method to a real wireless network.

6.1.2.7. Opportunistic routing

The model of wireless networks based on dynamic graph does not well assess the real processes in a wireless network. In particular the range of transmission can greatly vary between packets, the graph keeping only the average range. Opportunistic routing consists into taking advantage of temporary extension of the transmission range in order to gain several hops.

We have strong established theoretical performance limits in opportunistic routing. The limits are based on realistic interference scenarios in slotted Aloha. We have also investigated the impact of mobility on this theoretical limits.

We have designed an opportunistic routing protocol whose performance are within a small margin of the theoretical limits.

We have also conducted studies to support intelligent and adaptive forwarding, which allows a good trade-off between reliability and resource-efficiency. We then design a new protocol, called GrAnt, a new prediction-based forwarding protocol for complex and dynamic delay tolerant networks (DTNs). The proposed protocol uses the Ant Colony Optimization (ACO) metaheuristic with a greedy transition rule. This allows GrAnt to select the most promising forwarder nodes or allow for the exploitation of previously found good paths. The main motivation for using ACO is to take advantage of its population-based search and the rapid adaptation of its learning framework. Considering data from heuristic functions and pheromone concentration, the GrAnt protocol includes three modules: routing, scheduling, and buffer management.

6.1.2.8. Intermittent and delay tolerant networks

We consider the problem of routing in these networks, with the sole assumption that the speed of the node mobility is less than the speed of transmitting a packet to a neighbour. We compare this problem with sound propagation in liquid. We show that various pattern of mobility and network clustering can be described by a single parameter such as the information speed propagation.

We introduce new algorithms that route a packet toward a remote destination. The different algorithms vary depending on the buffering and the capacity capabilities of the network (i.e. if one or more copies of a packet can be sent and/or be kept). All algorithms are based on link aging rumors across connected components. The packet bounces from connected components to connected components, thanks to node mobility. We establish several analytical properties using an analogy with the sound propagation in liquid where molecules creates temporary connected components where sounds travel very fast.

Previous models assumed that the propagation of information path evolves like in a dynamic Erdos-Renyi graph leading to an epidemic flooding in O(log n) or O(1), n being the number of nodes in the network. We disprove the Erdos-Renyi model by showing via space-time considerations that the set of information path from a source to a destination is in fact much smaller than the path set in the Erdos-Renyi model. This lead to a much larger minimal delay in square root of n instead of log n. This correspond to a bounded maximal information propagation speed, whose estimate depends on the mobility model and the node density, and is root of multivariate explicit Bessel formulas.
Additionally, we have also considered the problem of data collection in global sensing and intermittently connected systems while avoiding the use of costly infrastructures (e.g., 3G). Motivated by the observation that node encounters are sufficient to build a connected relationship graph, we propose to take advantage of such inherent interactions to transform some mobile devices into delegates. We use then opportunistic delegation as a data traffic offload solution by investigating two main questions: (i) How to gain insights into social mobile networking scenarios?, (ii) How to utilize such insights to design solutions to alleviate overloaded 3G networks?. Our solution leverages usage of mobile applications requiring large data transfers by channeling the traffic to a few, socially important users in the network called VIP delegates. Mobile collectors need then only to meet delegates that, in turn, are responsible for gathering data from a subset of standard producers. We first investigate several delegation strategies based on the relative importance of nodes in their social interactions. Second, by considering a prediction strategy that estimates the likelihood of two nodes meeting each other, we investigate how the delegation strategies perform on predicted traces.

6.1.2.9. Network Coding
We study network coding for multi-hop wireless networks. We focus on the case of broadcasting where one source transmits information to all nodes in the network. Our goal is energy-efficient broadcast, that is, minimizing the total number of transmissions for broadcasting to the entire network. Note that this is a different problem for the classical problem of capacity maximization; and assuming we are far from the network capacity limit, hence in fact, we could assume interference-free transmissions.

Our previous results, they had shown that network coding (and a simple coding strategy) was able to reach optimality for asymptotically large and dense networks, with asymptotically 100 % of the received transmissions being useful (innovative). We extended the results with the combined use of connected dominating sets and network coding: we were able to quantify (and bound) the benefits of network coding in networks where the area of the network stays fixed, and only the density increases.

We have proved that the performance of wireless random network coding are optimal in the following network model: the Erdos-Renyi random graph model and the unit disk random graph model. In particular we show in the Erdos-Renyi the network coding capacity rate outperform any Connected Dominating Set strategy by a factor of order log n. In the unit disk model we gain is larger than 60%. The result is based on the analysis of the connectivity stretch ratio of the random graphs. The connectivity stretch ratio is the ratio of the smallest degree over the connectivity number, and the connectivity stretch ratio tends to one in the two graph models.

6.1.3. Collaborations
- Professor Bernard Mans, Macquarrie University, Sydney, Australia,
- TREC INRIA team,
- Professors Anelise Munaretto and Myriam Regattieri Delgado from Federal Technological University of Parana (UTFPR), Brazil,
- CNRS researcher Marcelo Dias de Amorim, LIP6/UPMC, France,
- Mathias Boc, CEA LIST, France,
- Computer Science Department, Sapienza University of Rome, Italy,
- University of St. Andrews, UK,
- Professor Leila Saidane, ENSI, Tunisia.

6.2. New generation of OLSR, new services and protocols
Participants: Cédric Adjih, Aline Carneiro Viana, Emmanuel Baccelli, Thomas Clausen, Philippe Jacquet, Pascale Minet, Ichrak Amdouni, Ridha Soua, Erwan Livolant, Paul Mühlethaler, Yasser Toor.

6.2.1. Executive summary
The user of a mobile network very quickly experience problems with quality of service: links fade, connectivity disrupts, delays accumulate.
In a wireless network, the set of neighbors which with one node can communicate depends on transmission range, and numerous factors, and in addition the transmission range is often lower than the interference range (the range within which a node prevents correct transmissions of other nodes). Thus bandwidth reservation, a crucial step of quality of service, is an important and difficult problem.

The services and protocols that need careful adaptation are

- Connectivity continuity
- Bandwidth reservation
- Delay routing
- Connectivity control
- Autoconfiguration
- Security
- Energy efficiency
- Localization

The connectivity continuity is the most important problem. Trivial in the wired world where a link failure is a rare event, it becomes problematic in the mobile world where link failure caused by mobility are frequent and normal. The first experiments of mobile ad hoc networks with regular internet protocols miserably failed simply because either the protocol was to slow to recover link failure, or when tuned appropriately was generating such a huge overhead that the network collapsed under its own weight. A new generation of routing protocols has arised that allow a suitable control of connectivity in mobile networks. Among them the **Optimized Link State Routing** combines the optimization of overhead for mobile networks and the full internet legacy. It naturally provides path redundancy which accelerate link failure recovery.

The most important lesson that must be retained is that most of these optimization become NP complete, which is a significant complication compared to their counterpart in the classical wired world. The reason for the NP-completeness is two-sided: on one side the co-interferences make impossible an optimization link by link, on the other side, the large dispersion of performance measurement makes simple heuristic ineffective. As an example, routing with respect to shortest delay average does not guarantee smallest probability of high delay.

Since the bandwidth is scarce, any multimedia application such as video streaming is resource demanding. For example a TV broadcast that uses a mesh network will rapidly exhaust the bandwidth if all connections are point to point. In this case multicast protocols that allows to gather all these point to point connections in a single flow is a need.

There are two classes of multicast protocols: the tree based protocols and the network coding protocols. In the first class the protocols take advantage of the relatively small size of the recipient node set. One can show equivalent results of Gupta and Kumar scaling properties but in the multicast plan when the ratio of recipient versus network size is a fundamental parameter. When this ratio tends to one the performance naturally worsen. When the recipient set is the whole network, one can apply the network coding scheme with random packet combination. In network coding the packets are no longer isolated: relay nodes makes linear combination of packets and transmitted mixed packets. In theory the performance of network coding is better than isolated packet multicast. In practice network coding is simpler to operate does not need topology management such as spanning trees or Connectedc Dominating Set. The reason for this is highly non intuitive, as if packet superposition was acting like state superposition in quantum mechanic, leading to non expected results.

Quality of service has become the central requirement that users expect from a network. High throughput, service continuity are critical issue for multimedia application over the wireless internet where the bandwidth is more scarce than in the wired world. A significant issue in the ad-hoc domain is that of the integrity of the network itself. Routing protocols allow, according to their specifications, any node to participate in the network - the assumption being that all nodes are behaving well and welcome. If that assumption fails - then the network may be subject to malicious nodes, and the integrity of the network fails. An important security service over mobile networks is to ensure that the integrity of the network is preserved even when attacks are launched against the integrity of the network.
6.2.2. Scientific achievements

6.2.2.1. Optimized Link State Routing (OLSR)

The routing protocol OLSR is universally known in the mobile wireless community (more than 475,000 hits on Google). It has numerous implementations and is used in many wireless networks. It is a proactive protocol with full internet legacy which is based on partial topology information exchange, that non the less provide optimal path with additive metrics (such as BGP/OSPF). It is an experimental RFC within IETF and soon will become a full standard under the name OLSRv2.

6.2.2.2. OSPF extension for wireless mesh networking

Long a near-future myth, ad hoc networks are now becoming a reality as a variety of wireless mesh networks are being deployed. Wireless mesh networks are a specific kind of ad hoc network, where terminals are essentially fixed. Even in such cases, which somewhat resembles usual networks, specific routing protocols have nevertheless to be employed, to cope with the characteristics of wireless, multi-hop communications. Such characteristics include scarce bandwidth over inherently unreliable, versatile, semi-broadcast links, and absence of a central authority in general. One of the main difficulties in this context is to cope with contradictory requirements such as, on one hand, dealing with bandwidth scarcity, which typically requires decreasing control traffic, while on the other hand, dealing with unreliable, versatile links which typically requires increasing control traffic. The two prominent routing protocols that have been developed for ad hoc networks and studied over the past decade, are the IETF standards AODV and OLSR. AODV is based on a reactive scheme (i.e. on-demand flooding to discover a path to a new destination), while OLSR is based on a proactive scheme, which is essentially an optimization of link state routing (i.e. pre-provisioning of paths to all possible destinations). OLSR is to date the most deployed such protocol, as it powers numerous wireless mesh community networks that currently flourish in various cities throughout Europe and North America. Based on this experience, the integration of ad hoc networking in the "standard" networking body is going further in several directions. One direction is the IEEE 802.11s standardization effort, which uses AODV and OLSR-derived algorithms to provide wireless mesh routing capabilities below IP. Another direction, spearheaded by the IETF, is the extension of IP routing standards such as OSPF to support ad hoc routing: in this realm we recently spun RFC 5449, as well as a series of academic publications on the subject. The idea behind extending OSPF to support ad hoc networks comes from a simple observation: OSPF is algorithmically quite similar to OLSR, as both are based on a proactive, link state approach. As on the other hand OSPF is a well-understood, widely deployed, industry-standard protocol, employing it to integrate ad hoc networks with existing infrastructure is considered by users as an easy migration path.

6.2.2.3. Multi-metric routing

Quality of service involves finding routes between two nodes in the network that satisfies a number of constraints. These constraints could be the requested bandwidth, the maximum delay, the minimum loss probability, the reliability of links, etc. This problem is NP-Complete because it combines additive metrics in the optimization problem. Hipercom proposed heuristics for finding routes that respect up to four metrics when calculating routes between source and destination. Another QoS issue is the creation of models that estimate the actual value of a metric. For example, computing the available bandwidth or the transfer delay on a link, etc. is very complex in a non-deterministic medium access such as WiFi. To resolve this problem, we developed a model for estimating the available bandwidth in a wireless network. This model is based on considering interfering links in cliques, after which we provide the maximum capacity that could be deployed in a clique. We may still enhance the model by adding a scaling factor to the clique equations in order to become more accurate when compared to real measurements.

In particular we have investigated the metric based on packet delay distribution. Since propagation delays between routers are negligible, most delays occur in queueing and medium access control processing. Contrary to previous common belief there is no need of network synchronization. The objective is to proactively determine the delay in absence of packet data traffic. The estimate of delay distribution is done via analytical method. In order to keep control on quality of service flows we use source routing forwarding options.
6.2.3. Collaborations

- Many contractual collaborations:
  - MoD (QoS, security, interconnection between the OLSR and OSPF routing domains),
  - Hitachi (Vehicular applications, OLSRv2),

- Non contractual:
  - BAE (OLSRv2),
  - Deutsche Telekom Labs/TU-Berlin, Germany,

6.3. Wireless Sensor Networks

Participants: Cédric Adjih, Aline Carneiro Viana, Emmanuel Baccelli, Thomas Clausen, Philippe Jacquet, Pascale Minet, Ichrak Amdouni, Ridha Soua, Erwan Livolant, Paul Mühlenthaler, Yasser Toor.

6.3.1. Executive summary

In wireless sensor networks, we focus more particularly on:

- Spatial reuse of the bandwidth,
- Routing according to a cross-layering approach,
- Security,
- Energy efficiency,
- IPv6 support.

6.3.2. Scientific achievements

6.3.2.1. Cryptographic Protocols to Fight Sinkhole Attacks on Tree-based Routing in Wireless Sensor Networks

Wireless Sensor Networks (WSN) are penetrating more and more in our daily life. As a consequence, security has become an important matter for these networks. We introduce two new cryptographic protocols of different complexity and strength in limiting network degradation caused by sinkhole attacks on tree-based routing topologies in Wireless Sensor Networks (WSNs). The main goal of both protocols is to provide continuous operation by improving resilience against, rather than detection of, these attacks. The main benefit of providing resilience is that it allows operating (or graceful degradation) in the presence of attacks. Furthermore, while resilience mechanisms do not dismiss detection mechanisms, detection mechanisms often introduce more complexity and so, more weaknesses to the system, which might not justify their benefits. More specifically our two RESIlient and Simple Topology-based reconfiguration protocols are: RESIST-1 and RESIST-0. RESIST-1 prevents a malicious node from modifying its advertised distance to the sink by more than one hop, while RESIST-0 does not allow such lying at the cost of additional complexity.

6.3.2.2. IPv6 Protocol suite for Sensor Networks

Wireless sensor networking is a key element of the Internet of Things (IoT), a substantial part of the billions of smart objects that are soon to blend into the global IP network, from actuators to home appliances, from smart meters, to smart dust. Sensor nodes are devices used for distributed and automated monitoring of various parameters such as temperature, movement, noise or radioactivity levels etc. Sensors are scattered with minimum planning with respect to their precise physical position (including the central role of the sink, if any), and the set of peers with which a sensor can directly communicate through its wireless interface may change rapidly over time due to asynchronous sleep mode strategies, fluctuations in the radio environment, device failure or mobility. Through its wireless interface, a sensor thus connects to a communication link with undetermined connectivity properties. Sensor networks are a challenge to current IP standards, since on the one hand these protocols were designed to work on wired links and on the other hand these protocols were designed to work on machines that do not have drastic constraints in terms of CPU, power capacities, and memory, as sensor nodes do. In consequence, several key standard protocols (including TCP, UDP, DHCP,
NDP, SLAAC, and OSPF) do not function correctly in this environment. Nevertheless, IPv6-based sensor networking is a viable long term goal because it would enable generic, large scale, seamless integration of millions of sensing devices using heterogeneous radio technologies, at a low cost, and in a future-proof manner. The Internet Engineering Task Force (IETF) is currently engaged into multiple efforts addressing the limitations of existing standards concerning wireless sensor IP networking. Some of the standards under construction aim at fitting IP formats, especially IPv6 formats, to direct wireless communications using low power radio technologies such as IEEE 802.15.4, which require IP format compression. Other standards in development aim at providing multi-hop wireless sensor communication with IPv6, which requires specific routing protocols, efforts in which we actively participate, prompting numerous joint publications with both industrial and academic partners.

6.3.2.3. Coloring in wireless sensor networks
Graph coloring is used in wireless networks to optimize network resources: bandwidth and energy. We focus on grid topologies that constitute regular topologies for large or dense wireless networks. We consider various transmission ranges and identify a color pattern that can be reproduced to color the whole grid with the optimal number of colors. We obtain an optimal periodic coloring of the grid for the considered transmission range. We then evaluate the performance of a 3-hop distributed coloring algorithm, called SERENA. Through simulation results, we highlight the impact of node priority assignment on the number of colors obtained for any network and grids in particular. We then compare these optimal results on grids with those obtained by SERENA and identify directions to improve SERENA.

6.3.2.4. Coloring algorithm optimized for dense wireless networks
In 2011, we also designed OSERENA "Optimized SchEduling RoutEr Node Activity", a distributed coloring algorithm optimized for dense wireless networks. Network density has an extremely reduced impact on the size of the messages exchanged to color the network. Furthermore, the number of colors used to color the network is not impacted by this optimization. We describe the properties of the algorithm and prove its correctness and termination. Simulation results point out the considerable gains in bandwidth.

6.3.2.5. Multichannel access in wireless sensor networks
In 2011 we started a research activity on multichannel access in wireless sensor networks. A state of the art has been published at the IFIP Wireless Days 2011 Conference.

6.3.3. Collaborations
- Many contractual collaborations:
  - Hitachi (Vehicular applications, OLSRv2),
  - OCARI2 project (industrial wireless sensor network, QoS, cross layer, energy efficiency, routing, node activity scheduling),
  - SAHARA project (wireless sensor network embedded in aircrafts),
  - STIC INRIA-Tunisian Universities: the team of Prof. Leila Saidane at ENSI (Performance improvement in a wireless sensor network),

- Non contractual:
  - BAE (OLSRv2),
  - Freie Universitaet (sensor networks, DHT),
  - Deutsche Telekom Labs/TU-Berlin, Germany,
  - University of Athens, Greece.

6.4. Vehicular and mobile applications
Participants: Cédric Adjih, Emmanuel Baccelli, Thomas Clausen, Philippe Jacquet, Pascale Minet, Paul Mühlethaler, Yasser Toor.
6.4.1. Executive summary

We have the following vision: in the future mobile internet and static internet will have their core deeply intricated. This means that mobile ad hoc networks will be attached to the core network, form extension and even be part of it. For example in disaster area, a wireless network could replace the destroyed infrastructure and help to the emergency operations.

With this perspective items such as Autoconfiguration, Security are of crucial importance. However there is a potential conflict between a large population of fixed nodes based on ancient protocol and a smaller but more dynamic population based on new protocols. In the integration both population must cooperate in an hybrid protocol.

The difficulty is to build protocols that are as dynamic and efficient as MANET protocols but can support the legacy of the old and heavy internet protocols. The challenge is nevertheless achievable, because the dynamic part of the network needs less frequent updates from the fixed part of the network. Moreover the natural abundance of resource in the fixed part of the network allows it to support the more frequent updates from the mobile part.

OLSR has been found to be the natural best candidate for this challenge since it gathers dynamic and optimization with internet legacy.

6.4.2. Scientific achievements

6.4.2.1. Military tactical networks

During year 2011, we conducted several expertises about industrial propoals dealing with OLSR use in military tactical networks.

6.4.2.2. Protocols for vehicular networks

We have achieved numerous studies and design of protocols for vehicular networks and more specificaly for V2V (Vehicle-to-Vehicle) network.

First we have studied the channel occupancy induced by the OLSR proactive routing protocol used in a linear Vehicular Ad hoc Network (VANET). Unlike previous studies, which usually use simulations to evaluate the overhead, we have proposed a simple analytical model to carry out this evaluation. Moreover, we did not evaluate the total overhead induced by the routing protocol as is usually proposed, but, for a given node, the channel occupation induced by the routing protocol.

We have studied flooding techniques for safety applications in VANETs. The typical scenario is the diffusion of an alert message after a car crash in a platoon of vehicles. The packet is diffused with the pure flooding, the multipoint relay (MPR) diffusion of OLSR and a geographic aware protocol. For OLSR we have introduced a variant (Robust-MPR) to improve the reliability. Different realistic scenarios were considered and various parameters such as vehicle density, and background traffic load were scrutinized. We have shown that the Robust-MPR and the geographic aware protocol satisfy the requirements of the safety applications while using considerably less overhead than pure flooding.

We have shown that the geographic aware protocols can be improved for the diffusion of an alert message by using opportunistic routing. We have designed OB-VAN (Opportunistic Broadcast for VANets ) a new protocol that uses this idea. One of the novelty of this protocol is the use of an active signalling technique in the acknowledgement procedure to select the best relay taking advantage of the reception pattern of each message. We have studied OB-VAN in a linear VANET and have shown that it outperforms the flooding for the delay and the amount of overhead. However the delivery ratio of OB-VAN may be insufficient for safety applications. This remark has led to the design of R-OB-VAN which is a reliable variant of OB-VAN. With extensive simulations, we have shown that R-OB-VAN maintains a high delivery ratio even in the presence of packet loss due to shadowing.
We have studied the performance of the Aloha scheme in linear VANETs. This analysis assumes a SINR (Signal over Interference plus Noise Ratio) based model. In this model, we have derived the probability of a successful transmission between two vehicles at a distance of $R$ meters. We have also computed the mean throughput according to Shannon’s law. In these two models, we have optimized the two quantities directly linked to the achievable network throughput i.e., the mean packet progress and the density of transport.

Finally, we have studied the utilization of opportunistic routing and shown that this technique is also beneficial for point to point traffic. It decreases the delay and increases the throughput compared with shortest path first routing. Moreover, we have also shown that opportunistic routing for point to point traffic eases considerably the optimization of the MAC scheme e.g. the transmission probability for Aloha and the carrier sense threshold for CSMA.

6.4.3. Collaboration

We received support from MoD for this activity.
6. New Results

6.1. Security

Participants: Ilaria Castellani, Zhengqin Luo, Tamara Rezk [correspondant], José Santos, Manuel Serrano.

6.1.1. Secure session calculi

We have pursued our work on controlling information flow in session calculi, started in previous years in collaboration with colleagues from the university of Torino. We also started investigating a notion of (objective) reputation for principals participating in sessions. The reputation of a principal is based on her previous behaviour as a user of a service. A principal's reputation can be checked both by the service itself, before admitting again the principal as a user, or by other principals to evaluate the reputation of the current users before they join a service (we consider multi-user services). We plan to apply this idea to refine our previous work on information flow control in multiparty sessions, by considering reputations built on the “security behaviour” of principals.

In the work “Information flow safety in multiparty sessions” [11], we consider a calculus for multiparty sessions enriched with security levels for messages. We propose a monitored semantics for this calculus, which blocks the execution of processes as soon as they attempt to leak information. We illustrate the use of our monitored semantics with various examples, and show that the induced safety property implies the security property studied previously for the same calculus. This work was presented at the 18th International Workshop on Expressiveness in Concurrency (EXPRESS’11).

In the work “A Reputation System for Multirole Sessions” [10], we extend role-based multiparty sessions with reputations and policies associated with principals. The reputation associated with a principal in a service is built by collecting her relevant behaviour as a participant in sessions of the service. The service checks the reputation of principals before allowing them to take part in a session, and decides whether to accept them or not depending on their reputation and on the role they want to play. Furthermore, principals can declare policies that must be fulfilled by the other participants of the same service. These policies are used by principals to check the reputation of the current participants and to decide accordingly whether or not to join the service. Our approach is illustrated by an example describing a real-world protocol. This work was presented at the 6th International Symposium on Trustworthy Global Computing (TGC’11).

Both [11] and [10] were partially funded by the ANR-08-EMER-010 grant PARTOUT.

6.1.2. Automatic Code Injection Prevention for Web Applications

We propose a new technique based on multitier compilation for preventing code injection in web applications. It consists in adding an extra stage to the client code generator which compares the dynamically generated code with the specification obtained from the syntax of the source program. No intervention from the programmer is needed. No plugin or modification of the web browser is required. The soundness and validity of the approach are proved formally by showing that the client compiler can be fully abstract. The practical interest of the approach is proved by showing the actual implementation in the HOPE environment.

This work was presented in TOSCA’11 and appeared in the LNCS series [13]. See also software section.

6.1.3. A Certified Lightweight Non-Interference Java Bytecode Verifier

We propose a type system to verify the non-interference property in the Java Virtual Machine. We verify the system in the Coq theorem prover.

This work will appear in the journal of Mathematical Structures in Computer Science [8].
6.1.4. Information-flow types for homomorphic encryptions

We develop a flexible information-flow type system for a range of encryption primitives, precisely reflecting their diverse functional and security features. Our rules enable encryption, blinding, homomorphic computation, and decryption, with selective key re-use for different types of payloads. We show that, under standard cryptographic assumptions, any well-typed probabilistic program using encryptions is secure (that is, computationally non-interferent) against active adversaries, both for confidentiality and integrity. We illustrate our approach using ElGamal and Paillier encryption.

We present two applications of cryptographic verification by typing: (1) private search on data streams; and (2) the bootstrapping part of Gentry’s fully homomorphic encryption.

We provide a prototype typechecker for our system.

This work appeared in CCS’11 [12]. See also software section.

6.1.5. The Mashic compiler

Mashups are a prevailing kind of web applications integrating external gadget APIs often written in the Javascript programming language. Writing secure mashups is a challenging task due to the heterogeneity of existing gadget APIs, the privileges granted to gadgets during mashup executions, and Javascript’s highly dynamic environment.

We propose a new compiler, called Mashic, for the automatic generation of secure Javascript-based mashups from existing mashup code. The Mashic compiler can effortlessly be applied to existing mashups based on a wide-range of gadget APIs. It offers security and correctness guarantees. Security is achieved by using the Same Origin Policy. Correctness is ensured in the presence of benign gadgets, that satisfy confidentiality and integrity constrains with regard to the integrator code. The compiler has been successfully applied to real world mashups based on Google maps, Bing maps, YouTube, and Zwibbler APIs.

See also software section.

6.1.6. Secure Information Flow by Self-Composition

Information flow policies are confidentiality policies that control information leakage through program execution. A common means to enforce secure information flow is through information flow type systems. Although type systems are compositional and usually enjoy decidable type checking or inference, their extensibility is very poor: type systems need to be redefined and proven sound for each new single variation of security policy and programming language for which secure information flow verification is desired. In contrast, program logics offer a general mechanism to enforce a variety of safety policies, and for this reason are favored in Proof Carrying Code, a promising security architecture for mobile code. However, the encoding of information flow policies in program logics is not straightforward, because they refer to a relation between two program executions. The purpose of this work is to investigate logical formulations of secure information flow based on the idea of self-composition, that reduces the problem of secure information flow of a program P to a safety property for program P composed with itself.

This work appeared in the special issue of MSCS of PLID [7].

6.1.7. Secure Information flow enforcement techniques for dynamic security policies

We performed a comprehensive investigation of alternative static mechanisms to enforce information flow policies considering a setting in which programs run under an authority that is only known at runtime and that yields a relaxation of the base security policy. The devised method aims at eliminating the need to reanalyse a program each time its authority changes. The soundness of the proposed approach was established for a concurrent higher-order imperative lambda calculus with reference creation. This work resulted in the report “Typing Illegal Information Flows as Program Effects” available at http://www-sop.inria.fr/members/Jose.Santos/reportInfFlow.pdf.
6.2. Models, semantics, and languages

Participants: Pejman Attar, Gérard Berry [correspondant], Gérard Boudol, Frédéric Boussinot, Johan Grande, Cyprien Nicolas, Manuel Serrano.

6.2.1. HipHop

HipHop is a new Domain Specific Language for HOP dedicated to request and event orchestration. HipHop follows the synchronous reactive model of the Esterel and ReactiveC languages, originally developed for embedded systems programming. It is based on synchronous concurrency and preemption primitives, which are known to be key components for the modular design of complex temporal behaviors. Although the language is concurrent, the generated code is purely sequential and thread-free. HipHop is translated to Hop code to be interpreted by the Hop runtime, either on server or client sides. HipHop has been described in a paper [9] accepted at the new International Workshop on Programming Language And Systems Technologies for Internet Clients (Plastic 2011).

6.2.2. Tesard

Tesard is a programming language of the Caml family, designed to offer simple constructs for shared memory concurrency and a deadlock-free semantics.

It features in particular the 2 following constructs:

- thread e which launches a thread that will execute expression e, and returns immediately;
- lock x in e which takes a lock on mutable value x (possibly waiting to be able to do so), executes expression e, releases the lock on x and returns the result of the execution of e.

A type and effect system is used at compile-time to:

- associate a mutex with every mutable value
- make the lock ... in... construct implement deadlock avoidance.

The language is implemented as an interpreter and a bytecode compiler. It is a fork of Llama Light, which is itself derived from Caml. Llama Light is functionnaly roughly equivalent to Caml Light, but a large part of its code comes from OCaml, along with the threads library and runtime that we ported ourselves.

Several key parts of the language have been implemented: the runtime and most of the type and effect system, including inference of the creation of mutexes (let region r in ...) but not region polymorphism yet. The project is not in a usable state (no release has been made yet). Development versions are publicly available via GitHub at https://github.com/nahoj/llama.

6.2.3. Synchronous orchestration and beyond

We studied DSL, an orchestration language based on the synchronous/reactive model. In DSL, systems are composed of several sites executed asynchronously. Within each site, scripts are run in a synchronous parallel way. Scripts may call functions that are treated in an abstract way: their effect on the memory is not considered, but only their “orchestration”, i.e., the organisation of their calls in time and in place (the site where they are called). The mapping of sites onto cores allows one to benefit from multicore architectures. Two properties are required by DSL: reactivity of sites and absence of interferences between scripts run by distinct sites. We also introduced DSLM, which adds a memory level to DSL and a way to automatically adapt the execution to get a maximal use of the available cores. This work, presented respectively in [18] and [15], was funded by the ANR-08-EMER-010 grant PARTOUT.

6.3. Web programming

Participants: Zhengqin Luo, Cyprien Nicolas, Tamara Rezk, Bernard Serpette, Manuel Serrano [correspondant].
6.3.1. A new evaluator for Hop

At the time where Hop programs were basic scripts, the performance of the server-side interpreter was not a concern. An inefficient interpreter was acceptable. As Hop expanded, Hop programs got larger and more complex. A more efficient interpreter was necessary. Therefore, this year we have designed and implemented a new interpreter for Hop. It is compact, its whole implementation counting no more than 2.5 KLOC. It is more than twice faster than the old interpreter and consumes less than a third of its memory. The architecture and the performance of the new interpreter are described in [14].

6.3.2. Abstraction of Hop with a bicolored lambda-calculus

We have studied an extension of the $\lambda$-calculus where each expression has a color. These colors designate the sites where expressions are evaluated, i.e., in the server or in the client. Colors are similar to the $\$ and $\sim$ annotations of Hop. With this, we have defined a transformation, using $\beta$-expansion, which groups together expressions with the same color. Correction, confluence and termination of the transformation was verified using the Coq system and its description was described in a paper to appear in 2012 [16].

Following Hop’s syntax, $\$ mentions that the followed expression is evaluated on the server while the $\sim$ character introduce client code. Inside the client code it is allowed to reintroduce some server expressions by reusing a $\$. We can imagine, for example, that the action associated to a client’s button is dependent on some server’s data (order number, proxies, data bases, ...). This kind of example is depicted with the abstract syntax tree in the upper left part of figure 1.

![Example of Hop colored tree](image)

Figure 1. Example of Hop colored tree

From the server’s point of view, the client’s code following a $\sim$ is ignored, but the environment in which this client code is activated must be preserved for all $\$ inside this code. Therefore, there exists a relation between a node $\sim$ and its including $\$. These relations are depicted by the dashed arrows in the upper right tree of Figure 1. The proposed transformation highlights this relation and is shown in the bottom of the figure where it can be observed that the transformation groups together expressions of the same color.

6.3.3. HopTeX, an Hop application for authoring documents

HOPTEX is a Hop application for authoring HTML and LaTeX documents. The content of the document is either expressed in HTML or in a blending of HTML and a dedicated wiki syntax, for the sake of conciseness and readability. The rendering of the document is expressed by a set of CSS rules. The main originality of HOPTEX is to consider LaTeX as a new media type for HTML and to express the compilation from HTML to LaTeX by the means of dedicated style sheet rules. HOPTEX can then be used to generate high quality documents for both paper printed version and electronic version.
HOPTEX is implemented in HOP, a multi-tier programming language for the Web 2.0. This implementation extensively relies on two facilities generally only available on the client-side that HOP also supports on the server-side of the application: DOM manipulations and CSS server-side resolutions.

The online version of the paper describing HOPTEX [17] is available at the HOPTEX web page (http://hop.inria.fr/hop/weblets/homepage?weblet=hoptex). Contrary to the PDF version published in the proceeding of the workshop, the online version is convenient for both desktop computers and Smartphones.
6. New Results

6.1. BlobSeer and Map-Reduce programming

6.1.1. BlobSeer-based cloud storage

**Participants:** Alexandra Carpen-Amarie, Alexandru Costan, Gabriel Antoniu, Luc Bougé.

As data volumes generated and processed by such applications increase, a key requirement that directly impacts the adoption rate of the Cloud paradigm is efficient and reliable data management. In this context, we investigate the requirements of Cloud data services in terms of data-transfer performance and access patterns and we explore the ways to leverage and adapt existing data-management solutions for Cloud workloads. We aim at building a Cloud data service both compatible with state-of-the-art Cloud interfaces and able to deliver high-throughput data storage.

To achieve this goal, we developed a file system layer on top of BlobSeer, which exposes a hierarchical file namespace enhanced with the concurrency-optimized BlobSeer primitives. Furthermore, we integrated the BlobSeer file system as a backend for Cumulus, an efficient open-source Cloud storage service. We validated our approach through extensive evaluations performed on Grid’5000. We devised a set of synthetic benchmarks to measure the performance and scalability of the Cumulus system backed by BlobSeer, showing it can sustain high-throughput data transfers for up to 200 concurrent clients.

Next, we explored the advantages and drawbacks of employing Cloud storage services for distributed applications that manage massive amounts of data. We investigated two types of applications. We relied on an atmospheric phenomena modeling application to conduct a set of evaluations in a Nimbus Cloud environment. This application is representative for a large class of simulators that compute the evolution in time set of parameters corresponding to specific points in a spatial domain. As a consequence, such applications generate important amounts of output data. We evaluated an S3-compliant Cloud storage service as a storage solution for the generated data. To this end, we employed distributed Cumulus services backed by various storage systems. The reason for targeting this approach is that storing output data directly into the Cloud as the application progresses can benefit higher-level applications that further process such simulation data. As an example, visualization tools need to have real-time access to output data for analysis and filtering purposes.

We built an interfacing module to enable the application to run unmodified in a Cloud environment and to send output data to an S3-based Cloud service. Our experiments show that distributed Cumulus backends, such as BlobSeer or PVFS, sustain a constant throughput even when the number of application processes that concurrently generate data becomes 3 times higher than the number of storage nodes.

6.1.2. Optimizing Intermediate Data Management in MapReduce Computations

**Participants:** Diana Moise, Gabriel Antoniu, Luc Bougé.

MapReduce applications, as well as other cloud data flows, consist of multiple stages of computations that process the input data and output the result. At each stage, the computation produces intermediate data that is to be processed by the next computing stage. We studied the characteristics of intermediate data in general, and we focused on the way it is handled in MapReduce frameworks. Our work addressed intermediate data at two levels: inside the same MapReduce job, and during the execution of pipeline applications.
We focused first on efficiently managing intermediate data generated between the “map” and “reduce” phases of MapReduce computations. In this context, we proposed to store the intermediate data in the distributed file system used as underlying backend. In this direction, we investigated the features of intermediate data in MapReduce computations and we proposed a new approach consisting in storing this kind of data in a DFS. The major benefit of this approach is better illustrated when considering failures. Existing MapReduce frameworks store intermediate data on nodes local disk. In case of failures, intermediate data produced by mappers can no longer be retrieved and processed further by reducers, the solution of most frameworks is to reschedule the failed tasks and to re-generate all the intermediate data that was lost because of failures. This solution is costly in terms of additional execution time. With our approach of storing intermediate data in a DFS, we avoid the re-execution of tasks in case of failures that lead to data loss. As storage for intermediate data, we considered BSFS as being a suitable candidate for providing for the requirements of intermediate data: availability and high I/O access. The tests we performed in this context, measured the impact of using a DFS as storage for intermediate data instead of the local-disk approach. We then assessed the performance of BSFS and HDFS when serving as storage for intermediate data produced by several MapReduce applications.

We then considered another type of intermediate data that appears in the context of pipeline MapReduce applications. In order to speed-up the execution of pipeline MapReduce applications (applications that consist of multiple jobs executed in a pipeline) and also, to improve cluster utilization, we proposed an optimized Hadoop MapReduce framework, in which the scheduling is done in a dynamic manner. We introduced several optimizations in the Hadoop MapReduce framework in order to improve its performance when executing pipelines. Our proposal consisted mainly in a new mechanism for creating tasks along the pipeline, as soon as the tasks’ input data becomes available. As our evaluation showed, this dynamic task scheduling leads to an improved performance of the framework, in terms of job completion time. In addition, our approach ensures a more efficient cluster utilizations, with respect to the amount of resources that are involved in the computation.

We evaluated both approaches for intermediate data through a set of experiments on the Grid’5000 [56] testbed. Preliminary results [17] show the scalability and efficiency of our proposals, as well as additional benefits brought forward by our approach.

6.1.3. A-Brain: Perform genetic and neuroimaging data analysis in Azure clouds

Participants: Radu Tudoran, Alexandru Costan, Gabriel Antoniu, Louis-Claude Canon.

Joint genetic and neuroimaging data analysis on large cohorts of subjects is a new approach used to assess and understand the variability that exists between individuals. This approach has remained poorly understood so far and brings forward very significant challenges, as progress in this field can open pioneering directions in biology and medicine. As both neuroimaging- and genetic-domain observations represent a huge amount of variables (of the order of $10^{6}$), performing statistically rigorous analyses on such amounts of data represents a computational challenge that cannot be addressed with conventional computational techniques.

In order to perform an accurate analysis we need to provide a programming platform and a high throughput storage. The target infrastructure is the Azure clouds. Hence we have adapted the BlobSeer storage approach for Azure, thus providing a new way to store data in clouds, that federates the local storage space from computational nodes into a uniform shared storage, called TomusBlobs. Using this storage system as a storage backend, we have built a MapReduce prototype for Azure clouds. This MapReduce system, called TomusMapReduce -TMR, is used to perform the simulation of the joint genetic and neuroimaging application. For validating the framework, a toy application that simulate the data access and computation patterns of the real application, was used. The next step, after the evaluation of the framework, that has just began, consists in replacing the toy application with the real one and the scaling of the framework in the limit allowed by the cloud provider. In addition a demo for this project is in progress, that will consists in providing a visualization tool for the framework. This will be used to intuitively represent the results for the simulation of the scientific application, this being useful both for better presenting the project to interested parties and for the researchers from bioinformatics.

6.2. Efficient VM management in clouds

Participants: Alexandru Costan, Alexandra Carpen-Amarie, Gabriel Antoniu.
Infrastructure as a Service (IaaS) cloud computing allows users to lease computational resources from the cloud provider’s datacenter for a short time by deploying virtual machines (VMs) on these resources. This model raises new challenges in the design and development of IaaS middleware. One of those challenges is the need to deploy a large number (hundreds or even thousands) of VM instances simultaneously. Once the VM instances are deployed, another challenge is to simultaneously take a snapshot of many images and transfer them to persistent storage to support fault tolerance and management tasks, such as suspend-resume and migration. With datacenters growing rapidly and configurations becoming heterogeneous, it is important to enable efficient concurrent deployment and snapshotting that are at the same time hypervisor independent and ensure a maximum compatibility with different configurations.

We addressed these challenges by proposing a virtual file system specifically optimized for virtual machine image storage [19]. It is based on a lazy transfer scheme coupled with object versioning that handles snapshotting transparently in a hypervisor-independent fashion, ensuring high portability for different configurations. Large-scale experiments on hundreds of nodes demonstrate excellent performance results: speedup for concurrent VM deployments ranges from a factor of 2 up to 25, with a reduction in bandwidth utilization of as much as 90% [18]. We implemented this deployment scheme in the Nimbus cloud and presented a demo illustrating it at the Grid’5000 School [26].

Given the dynamic nature of IaaS clouds and the long runtime and resource utilization of scientific applications, an interesting use-case for the multi-snapshotting techniques is for efficient checkpoint-restart. We introduced an approach that leverages VM disk-image multi-snapshotting and multi-deployment inside checkpoint-restart protocols running at guest level in order to efficiently capture and potentially roll back the complete state of the application, including file system modifications. This framework is specifically optimized for tightly-coupled scientific applications that were written using a message passing system (in particular MPI) and need to be ported to IaaS clouds. Our solution introduces a dedicated checkpoint repository that is able to efficiently take incremental snapshots of the whole disk attached to the virtual machine instances, thus offering support to use any checkpointing protocol that can save the state of processes into files, including application level mechanisms, where the process state is managed by the application itself, and process-level mechanisms, where the process state is managed transparently at the level of the message passing library. Experiments on the G5K testbed show substantial improvement for MPI applications over existing approaches, both for the case when customized checkpointing is available at application level and the case when it needs to be handled at process level.

We integrated this checkpointing scheme inside the Nimbus cloud with some promising preliminary results. We plan to complement the existing solution with live incremental snapshotting using asynchronous background transfers for high checkpointing efficiency and with adaptive prefetching to achieve high restart efficiency.

6.3. Cloud data storage management

6.3.1. Autonomic storage for cloud services

Participants: Alexandru Costan, Alexandra Carpen-Amarie, Gabriel Antoniu, Florin Pop, Ciprian Dobre, Elena Apostol.

A means to achieve performance improvement and resource-usage optimization in cloud storage systems consists in enabling an autonomic behavior based on introspection. Self-adaptation incurs a high degree of complexity in the configuration and tuning of the system, with possible repercussions on its availability and reliability. To address these challenges we introduced in BlobSeer in [11] a three-layered architecture designed to identify and generate relevant information related to the state and the behavior of the system, based on the MonALISA monitoring framework. Such information is then expected to serve as an input to a higher-level self adaptation engine. These data are yielded by an (1) introspection layer, which processes the raw data collected by a (2) monitoring layer. The lowest layer is represented by the (3) instrumentation code that enables BlobSeer to send monitoring data to the upper layers.
A first approach to leverage the introspection framework aims at enhancing BlobSeer with self-configuration capabilities, as a means to support storage elasticity through dynamic deployment of data providers. This solution enables the data providers to scale up and down depending on the detected system’s needs. The component we designed adapts the storage system to the environment by contracting and expanding the pool of storage providers based on the system’s load. The key idea of this component is the automatic decision that has to be made on how many resources the system needs to operate normally while keeping the resources utilization down to a minimum. This problem is addressed by using a test-decided heuristic based on the monitoring data. The introspective architecture has been evaluated on the Grid’5000 testbed, with experiments that prove the feasibility of generating relevant information related to the state and the behavior of the system.

We plan to use the introspective BlobSeer to develop a distributed data aggregation service. Its primary goal will be to serve as a repository backend for complex analysis and automatic mining of scientific data. Another direction that will be explored is to use the introspective BlobSeer as a cloud-based storage layer for sensitive context data, collected from a vast amount of sources: from smartphones to sensors located in the environment. Clouds are perfect candidates to handle the storage and aggregation of such data for even larger context-aware applications. Such solutions rely on more relaxed storage capabilities than traditional relational databases (eventual consistency suffices for example). This, combined with the high concurrency support and the flexible storage schema make BlobSeer a suitable candidate for the storage layer. We plan to develop a new layer on top of BlobSeer targeting context aware applications. At the logical level, this layer will provide transparency, mobility, real-time guarantees and access based on meta-information. At the physical layer, the most important capability will rely on BlobSeer’s elasticity to scale up and down according to real-time usage, in order to reduce the costs within the Cloud.

6.3.2. Managing data access on Clouds through security policies

Participants: Alexandru Costan, Alexandra Carpen-Amarie, Gabriel Antoniu.

With the emergence of Cloud computing, there has been a great need to provide an adequate security level in such environments, as they are vulnerable to various attacks. Malicious behaviors such as Denial of Service attacks, especially when targeting large-scale data management systems, cannot be detected by typical authentication mechanisms and are responsible for drastically degrading the overall performance of such systems.

In [14] we proposed a generic security management framework allowing providers of Cloud data management systems to define and enforce complex security policies. The generality of this approach comes from the flexibility both in terms of supporting custom security scenarios and interfacing with different Cloud storage systems. This security framework is designed to detect and stop a large array of attacks defined through an expressive policy description language and to be easily interfaced with various data management systems. We introduced a modular architecture consisting of three components. The Policy Management module represents the core of the framework, where security policies definition and enforcement takes place. This module is completely independent of the Cloud system, as its input only consists in user activity events monitored from the system. The User Activity History module is a container for monitoring information describing users’ actions. It collects data by employing monitoring mechanisms specific to each storage system and makes them available for the Policy Management module. The Trust Management module incorporates data about the state of the Cloud system and provides a trust value for each user based on his past actions. The trust value identifies a user as a fair or a malicious one. Furthermore, the trust values enable the system to take custom actions for each detected policy violation, by taking into account the history of each user.

As a case study, we applied the proposed framework to BlobSeer. We defined a specific policy to detect DoS attacks in BlobSeer and we evaluated the performance of our framework through large scale experiments on the Grid’5000 testbed. The results show that the Policy Management module meets the requirements of a data storage system in a large-scale deployment: it was able to deal with a large number of simultaneous attacks and to restore and preserve the performance of the target system.

As a next step we will focus on more in-depth experiments involving the detection of various types of attacks in the same time. Moreover, we will investigate the limitations of our Security Management framework, with
respect to the accuracy of the detection in the case of more complex policies, as well as the probability and the impact of obtaining false positive or false negative results. Another research direction is to further develop the Trust Management component of the security management framework and study the impact it has on the Policy Enforcement decisions for complex scenarios.

6.4. Storage architecture and adaptive consistency for clouds

Participants: Houssem-Eddine Chihoub, Gabriel Antoniu.

As more and more applications are becoming data-intensive, the design of a scalable storage architecture providing a huge file sharing and fine grain access with high throughput under heavy concurrency is a timely and relevant challenge.

In [24], we introduce a storage architecture for Cloud computing. This architecture proposes efficient and scalable storage support for both VM images and application data. Our architecture relies on BlobSeer [12], a data sharing platform optimized for concurrent accesses, as a basic storage backend enhanced in term of quality of service and efficiency by GloBeM tool [61] that rely on behavior modeling and monitoring to avoid bad case scenarios. The architecture uses this approach to have a better and efficient storage VM images, that allow a faster image deployment and efficient versioning and snapshotting. Furthermore, the architecture provides a platform to store, manage, and share cloud application data allowing several key features to clouds such as storage elasticity.

The main aim is to provide high availability and good scalable performance at low cost. This is justified by the growing need of data-intensive applications for managing huge sets of data replicated over several data centers. In order to provide good performance and high availability, data replication is mandatory. But this generates the issue of replicas consistency as shown by the CAP theorem [52]. To achieve the aforementioned goals, relaxing consistency rules is unavoidable. On the other hand, opting for weaker consistency, all the time, can be too costly.

In current work we leverage the trade-off between consistency and availability and performance. We are investigating an approach that changes the consistency level at runtime considering system and application needs. In order to choose the most suitable consistency level, our approach monitor the storage system and collect useful information, such as network load and applications access patterns, that enable the system to estimate the amount of expected stale reads.

6.5. Modelling cloud storage performance

Participants: Daniel Higuero, Louis-Claude Canon, Alexandru Costan, Gabriel Antoniu.

The objective of this research direction is to provide comprehensive performance models for storage systems. Their role is to capture how the system components interact for different usage patterns (number of reads or writes). The objective is to determine the incurred costs in terms of storage space and efficiency for a given workload.

One application of this model consists in dynamically adjusting the parameters of the storage system as required in an autonomic approach. For this purpose, it is necessary to identify the characteristics of the storage system for meeting a given level of requirements. Progress has been made on this part during the 3-month visit of Daniel Higuero (University Carlos III, Madrid). A preliminary performance model currently predicts the available bandwidth when multiple concurrent transfers occur. This model serves as a basis for a dimensioning strategy that is formulated through a linear program.

This approach has further been complemented with an offline analysis of several traces of the BlobSeer storage system when it is used as a backend for MapReduce applications. Mining this information in an automated fashion allowed to detect the different trade-offs that influence a BlobSeer deployment: time required to execute the application vs. number of machines used by the storage system, communication costs vs. space usage. The final goal is to tune BlobSeer for specific applications. The proposed strategy is currently being evaluated.
Future directions are directed towards refining the proposed model. Several parameters significantly impact the performance of storage systems such as the redundancy mechanism, the data placement strategy or disk-related effects. As a first step, experiments for assessing the quality of finer models will be designed. Ultimately, we aim to capture the I/O variability of storage systems, in particular in the context of the cloud.

This work will enable new collaborations. It is planned to work on the models mentioned above with the Mescal INRIA team in the context of a collaboration between the MapReduce ANR project and the Songs ANR project. Moreover, in the framework of the MapReduce project, we expect to work on a performance model for designing decision algorithms that are required by the component-based MapReduce framework that is developed in the GRAAL/Avalon INRIA team. Finally, the GRAAL/AVALON team works on scheduling algorithms that could beneficially profit from a storage performance model.

6.6. Scalable I/O and visualization for post-petascale HPC simulations

**Participants:** Matthieu Dorier, Gabriel Antoniu.

In the context of the Joint INRIA/UIUC Laboratory for Petascale computing (JLCP), we are addressing the new challenges related to I/O, data analysis and visualization for extreme-scale simulations. As HPC resources approaching millions of cores become a reality, a growing challenge in maintaining high performance is the presence of high variability in the effective throughput of codes performing I/O operations. Since I/O is mainly performed for the purpose of subsequent data analysis and visualization, another way to limit the impact of I/O performance on scientific discovery consists in enabling in-situ visualization. This brings again new challenges such as how to efficiently couple large-scale simulations with visualization software.

We started the development of Damaris, a middleware targeting multicore SMP nodes to efficiently address the problems mentioned above. Damaris has been evaluated with the CM1 atmospheric simulation [43], one of the targeted application for the BlueWaters project. To show the capability of our approach to efficiently hide I/O jitter and related costs, experiments have been carried on the French Grid’5000, on a Power5 cluster at NCSA and on the Kraken Cray XT5 supercomputer (currently 11th in the Top500) with up to 9K cores. By gathering data into large files while avoiding synchronization between processes, our solution brings several benefits:

1. it increases the sustained write throughput of the simulation by a factor of almost 15;
2. it provides almost 70% overall application speedup on 9,000 cores;
3. it fully hides I/O-related costs;
4. it enables a 600% compression ratio without any additional overhead, leading to a major reduction of storage requirements.

A poster [16] presenting some of these results has been accepted at ICS’11 and awarded the second price at the ACM Student Research Competition. All these results are presented in a research report [25] pending for publication.

Current work addresses the efficient coupling of large-scale simulations and visualization tools through Damaris. We have been able to get access to the Jaguar supercomputer hosted at ORNL and we are planning very-large scale experiments on up to 100,000 cores to show the benefits of our Damaris approach.

6.7. Scalable array-oriented active storage

**Participants:** Viet-Trung Tran, Gabriel Antoniu, Luc Bougé.
The recent explosion in data sizes manipulated by distributed scientific applications has prompted the need to develop specialized storage systems capable to deal with specific access patterns in a scalable fashion. In this context, a large class of applications focuses on parallel array processing: small parts of huge multi-dimensional arrays are concurrently accessed by a large number of clients, both for reading and writing. However, many established storage solutions such as parallel file systems and database management systems expose data access models (e.g., file systems, structured databases) that are too general and do not exactly match the nature requirements of the application. This forces the application developer to either adapt to the exposed data access model or to use an intermediate layer that performs a translation. In either case, the mismatch leads to suboptimal data management: the one-storage-solution-fits-all-needs has reached its limits.

Thus, there is an increasing need to specialize the I/O stack to match the requirements of the application. The objective of this research is to design Pyramid: an array-oriented active storage system optimizing for applications that represent and manipulate data as huge multi-dimensional arrays. However, a specialized storage system that deals with such an access pattern faces several challenges at the level of data/metadata management, we carefully design the system with the following principles: (1) we introduce a dedicated array-oriented data access model that offers support for active storage and versioning; (2) we enrich striping techniques specifically optimized for multi-dimensional arrays with a distributed metadata management scheme that avoids potential I/O bottlenecks observed with centralized approaches.

We evaluated Pyramid through a set of experiments on the Grid’5000 [56] testbed that aims to evaluate both the performance and the scalability of our approach under concurrent accesses. Preliminary evaluation in our recent papers [22], [13] shows promising results: our prototype demonstrates good performance and scalability under concurrency, both for read and write workloads.
6. New Results

6.1. Behavioral Fingerprinting

Participant: Olivier Festor [contact].

Device fingerprinting aims to automatically determine the types (name and version of software, brand name and series of hardware) of remote devices for a given protocol. Hence, keeping an up-to-date inventory database of devices in use on a network is possible and helpful as for example to check remotely if unauthorized applications have been installed. Some types of devices for which vulnerabilities are known can be easily detected in order to patch them or at least send alerts to the owners. From a security point of view, attackers use specific tools to perform their attack which may also be detected rapidly thanks to fingerprinting. Most current systems rely only on signatures of differences in implementation of a given protocol stack and signatures are often outdated.

We have designed a new fingerprinting scheme that is accurate even on protocol stacks that are completely identical, but which run on hardware having different capabilities (CPU power, memory resources, etc). Our fingerprinting scheme can learn distinctive patterns in the state machine of a particular implementation. We see such a pattern as a restricted tree finite state machine that provides additional time-related information about the transitions performed [15]. The captured identification models were then used to automatically build attack prevention rules [19].

This work was done in cooperation with Jérôme Francois, Radu State and Thomas Engel from the University of Luxembourg.

6.2. Management and monitoring of P2P networks

Participants: Isabelle Chrisment [contact], Olivier Festor, Juan Pablo Timpanaro.

Content pollution is one of the major issues affecting P2P file sharing networks. However, since early studies on FastTrack and Overnet, no recent investigation has reported its impact on current P2P networks. In [21], we presented a method and the supporting architecture to quantify the pollution of contents in the KAD network. We first collected information on many popular files shared in this network. Then, we proposed a new way to detect content pollution by analyzing all filenames linked to a content with a metric based on the Tversky index and which gives very low error rates. By analyzing a large number of popular files, we showed that 2/3 of the contents are polluted, one part by index poisoning but the majority by a new, more dangerous, form of pollution that we call index falsification. This work was done, in collaboration with the University of Technology of Troyes, within the context of the ACDA-P2P Project funded by GIS-3SGS.

BitTorrent is a widely deployed P2P file sharing protocol, extensively used to distribute digital content and software updates, among others. Recent actions against torrent and tracker repositories have fostered the move towards a fully distributed solution based on a distributed hash table to support both torrent search and tracker implementation. We conducted an analysis on one of the BitTorrent’s DHT (Mainline DHT) and developed a monitoring architecture, so as to measure and discover security flaws on the network. In [23] we compared KAD DHT against BitTorrent DHT in terms of security by deploying different attacks on the network. We showed that the lack of security in Mainline DHT allows very efficient attacks that can easily impact the operation of the whole network. We also provided a peer-ID distribution analysis of the network, so as to adapt previous protection schemes to the Mainline DHT. The mechanisms are assessed through large-scale experiments on the real DHT-based BitTorrent tracker.

3Approche Collaborative pour la Détection d’Atttaques dans les réseaux Pair à Pair
4Groupement d’Intérêt Scientifique - Surveillance, Sureté et Sécurité des grands Systèmes
If BitTorrent’s Mainline DHT is exposed to several identified security issues, in parallel, the KAD DHT has been the core of intense research and was improved over years. We presented a study that motivates the integration of both worlds. We provided a performance comparison of both DHTs in terms of publishing efficiency. We investigated the security threats and showed that the current BitTorrent’s Mainline DHT is more vulnerable to attacks than KAD while the download service of BitTorrent has much better performance. Given the strengths and weaknesses of both DHTs, we designed a hybrid architecture [24], which is based on KAD’s indexation mechanism and BitTorrent download protocol. On the one hand, the client is able to index its files in the well-known KAD DHT, taking advantage of KAD’s security mechanism and its double-indexation scheme. On the other hand, the client uses the BitTorrent download protocol so as to download a given file, which has been proven to surpass KAD’s. We implemented this hybrid architecture, that we called hMule, as a unified KAD-BitTorrent file-sharing application, which is compatible with both P2P file sharing networks and provides the KAD advantages on indexation and the BitTorrent speed for transfer without losing backward compatibility.

We started our research about being anonymous when downloading from BitTorrent. We conducted a set of measurements from High Security Lab aiming to characterize the usage of the I2P network, a low-latency anonymous network based on garlic routing [35]. Our goal was to answer the following questions: what is the network used for? when is it used the most? which kind of applications the network designers should pay more attention to? We designed a distributed monitoring architecture for the I2P network and we showed that, through three one-week long experiments, we were able to identify 32% of all running applications, among web servers and file-sharing clients. Additionally, we identified 37% of published I2P applications, which turned out to be unreachable after their publication on the I2P distributed database.

We built-up a model of I2P encryption/decryption approach and using the Avispa tool, we able to find a possible attack on the network. Further work will be focused on probing right and on developing a proof-of-concept of this.

6.3. Configuration security automation

Participants: Rémi Badonnel [contact], Martin Barrere, Olivier Festor.

The main research challenge addressed in this work has focused on enabling configuration security automation in autonomic networks and services. In particular our objective has been to increase vulnerability awareness in the autonomic management plane in order to prevent configuration vulnerabilities. The continuous growth of networking significantly increases the complexity of management. It requires autonomic networks and services, which are capable of taking in charge their own management by optimizing their parameters, adapting their configurations and ensuring their protection against security attacks. However, the operations and changes they execute during these management activities may generate vulnerable configurations. A first part of our work has therefore consisted in consolidating a security automation strategy for preventing vulnerabilities and maintaining safe configurations in autonomic infrastructures [7]. This solution relies on the integration of configuration vulnerability descriptions into the management plane [8]. The OVAL language, part of the SCAP protocol, has become the de-facto standard for specifying configuration vulnerabilities in a technical viewpoint. We have refined a mathematical modeling for mapping OVAL descriptions into policy rules which can be interpreted by the autonomic Cfengine configuration system. These policies enable the Cfengine system to assess and detect vulnerabilities. We have designed a functional architecture and formalized a translation algorithm for supporting this security automation. We have also prototyped an OVAL-to-Cfengine translation module, called Ovalyzer, and analyzed its interactions with the components of the Cfengine system. Based on vulnerability descriptions extracted from the official OVAL repository, we have performed an extensive set of experiments to quantify the performance and coverage of the Ovalyzer module. A second part of our work has consisted in investigating how our security automation solution can be extended to distributed configuration vulnerabilities. In SCAP-based traditional approaches, a distributed vulnerability is typically understood as the aggregation of individual configuration vulnerabilities which are spread in the network and might allow a multi-step attack. We have shown through the analysis of a case study that this definition does not offer a complete outlook of the problem. In particular, each network device can individually present a secure
configuration, but when combined across the network, a global vulnerable configuration may be produced. In that context, we have introduced in [27] a mathematical definition for distributed vulnerabilities and have specified the DOVAL language (Distributed OVAL), on top of OVAL, as a means for describing these vulnerabilities in a machine readable manner. A case study in the area of VoIP networks and services has been considered for demonstrating the instantiation of DOVAL main constructs. The DOVAL descriptions constitute useful security definitions that in turn can be exploited for security automation. We have built a framework for supporting these distributed configuration vulnerabilities based on the Cfengine system. In particular, we have proposed and evaluated collaborative strategies and optimized algorithms for performing the assessment of DOVAL descriptions.

6.4. Online Risk Management

Participants: Rémi Badonnel [contact], Oussema Dabbebi, Olivier Festor.

Telephony over IP has known a large scale deployment and has been supported by the standardization of dedicated signaling protocols. This service is however exposed to multiple attacks due to a lower confinement in comparison to traditional PSTN networks. While a large variety of methods and techniques has been proposed for protecting VoIP networks, their activation may seriously impact on the quality of such a critical service. Risk management provides new opportunities for addressing this challenge. In particular, our work aims at performing online risk management for VoIP networks and services. The purpose is to adapt the service exposure with respect to the threat potentiality, while maintaining a low security overhead. Based on the classification of VoIP attacks and the analysis of their properties, we have refined in [11] an extended risk modeling for IP telephony infrastructures. This modeling permits to cover a large spectrum of security attacks. It supports our online risk management strategy which is capable of dynamically activating or deactivating security safeguards in the VoIP infrastructure. The mitigation is based on the control of the service exposure using these safeguards. We have compared our solution to other traditional strategies, and have quantified the benefits and limits according to multiple performance criteria. We have also analyzed the impact of the risk model parameters on our mitigation, and showed to what extent the parameterization can be partially automated in [12]. An important part of our efforts has focused in the year 2011 on extending our online risk management strategy to more distributed configurations [32]. While our initial work was centered around Asterisk-based enterprise networks, we have taken a particular interest in P2PSIP networks. They constitute an open decentralized solution where the registration and location servers are implemented by a distributed hash table responsible for storing the bindings between the address-of-record SIP-URI and the contact SIP-URI. We have identified different attack sources and attack scenarios in these P2PSIP networks, considering the functional roles that are played by the SIP peers. The security threats are specific to the P2PSIP protocol or are the result of inheritance from the SIP layer and the peer-to-peer area. In that context, we have analyzed the instantiation of our online risk modeling by taking into account the properties and components of the P2PSIP architecture, and have established a portfolio of dedicated countermeasures, including replication-based and certification-based techniques. We have evaluated the strategy performance and scalability through an extensive set of experiments performed with the OMNET++ simulator. We also have quantified the complementarity of our solution with the RELOAD security framework which relies on a central certificate enrolment server.

6.5. VoIP Security

Participants: Laurent Andrey, Olivier Festor, Abdelkader Lahmadi [contact].

In previous work, we have proposed the prevention system SecSIP [5] for SIP-based networks which uses a rule-based approach to build prevention specifications on SIP protocol activities that stop attacks exploiting an existing vulnerability before reaching their targets. We have pursued our efforts in VoIP security which led to two new contributions:

- Building and maintaining prevention rules using the VeTo language can become a time consuming and error prone task, especially when addressing an important number of vulnerabilities discovered
using a fuzzing tool. The discovered vulnerabilities using such process are usually based on a single exploit message with a malformed field or sequence of vulnerable messages. To reduce this effort, we have designed a generation method to produce VeTo specifications targeting those vulnerable messages. The method mainly characterizes a malformed field within an exploit message or the vulnerable sequence of messages and generates a set of VeTo rules specifications to prevent their exploit. The generated VeTo rules are then deployed and maintained on the SecSIP engine to be applied against the SIP traffic. The solution [19] relies on generating rules using genetic algorithms operating on a set of candidate regular expressions to match a malformed pattern within a SIP message, and evaluate their quality using a well defined fitness function to ensure that their are specific enough to only match exploit messages.

- SecSIP uses a plain text configuration file in which VeTo specifications are authored and managed manually. While extending the deployment of the framework beyond our own lab, support for remote configuration was required. Given the promise of Netconf, we naturally turned our investigations towards this protocol and embraced the YANG data-modeling framework. In [20] we have presented the Yang model built for VeTo policies and the Netconf framework put in place.

We have developed a flexible SIP honeypot. It is flexible in the sense that a behavior can be externally and easily defined. The goal of such a honeypot is to be able to be quickly customized in response to an observation made on a more generic and large scale honeypot. If the initial observation is likely to be an attack the customized honeypot would eventually get deeper and more informative interactions with the attacker. The realization is a module of the Dionaea general framework for honeypot (successor of the well-known nepenthes framework) and we use the SIPP test tool as an engine to animate SIP interactions provided as automata in some XML file. More detail on the implementation can be found in [26].

6.6. VoIP Fraud

Participants: Olivier Festor [Contact], Mohamed Nassar.

In the context of a cooperation with the University of Liege, we have addressed the problem of SPIT from a new perspective [22]. Based on end-user feedback, we have proposed a scheme for generating SPIT signatures from the SIP INVITE messages. Hence it is possible to filter the next SPIT calls before ringing their destinations. The generated SPIT signatures are adaptive to the benign signaling traffic in the sense that they do not conflict with it. The generation of signatures is based on supervised machine learning techniques. We namely investigated decision trees with categorical attributes obtained by parsing the SIP messages.

Our system works in two modes: a batch and an online mode. The batch mode consists on training the decision tree over a labeled (spit, normal) data-set and then transforming the tree into an if-else rule-set. In online mode, the successive learnt signatures are aggregated and the possible conflicts are resolved. Experimentation on off-the-shelf SPIT tools showed the efficiency of our approach to find the good signatures. However, experiments show that the J48 decision tree is easily defeated using some obfuscation techniques. We therefore proposed a generalisation approach to translate the tree into an if-else rule-set shows instead good robustness against such attacks. The overall framework provides suitable performance for operational deployment in terms of learning time, required memory, size of the rule-set and the call setup delay. The different parameters of the system (i.e. size of the different buffers and windows) are easily configurable. Different SPIT signatures may imply different SPIT capabilities. For example, a spitter may break a Captcha test by brute-forcing a DTMF guess. Another spitter may start talking by a human-like congratulation in order to bypass a Turing test. One of the goals of our approach is to provide a framework for applying reinforcement learning techniques and hence increasing the efficiency of the filtering process. The reinforcement learning aims at selecting the best challenge to be used when a given SPIT signature is detected. Basically the re-inforcement learning maintains a table matching each signature with the best challenge response discovered so far. The table is continuously updated using a trial and error scheme.

We did validate the approach on multiple data-sets obtained from Voice over IP operators members of the SCAMSTOP project.
6.7. Pervasive computing

Participants: Laurent Ciarletta [contact], Tom Leclerc, Julien Siebert, Olivier Festor, André Schaff.

Vincent Chevrier (MAIA Team)

In Pervasive or Ubiquitous Computing, a growing number of communicating/computing devices are collaborating to provide users with enhanced and ubiquitous services in a seamless way. Madynes is focusing on the networking aspects of ubiquitous systems. We cooperate with the Maia (and Trio) team(s) to be able to encompass issues and research questions that combine both networking and cognitive aspects.

Pervasive Computing is about interconnected and situated computing resources providing users with contextual services. These systems, embedded in the fabric of our daily lives, are complex: numerous interconnected and heterogeneous entities are exhibiting a global behavior impossible to forecast by merely observing individual properties. Firstly, users physical interactions and behaviors have to be considered. They are influenced and influence the environment. Secondly, the potential multiplicity and heterogeneity of devices, services, communication protocols, and the constant mobility and reorganization also need to be addressed. Our research on this field as detailed in [10] is going towards both closing the loop between humans and systems and taming the complexity, using multi-modeling (to combine the best of each domain specific model) and co-simulation (to design, develop and evaluate) as part of a global conceptual and practical toolbox.

In 2011 we worked on the following research topics:

• Multi-models of these Pervasive Computing environments (including the users in the modeling and the simulations). We have been focusing on the collaborative simulations of dynamic networks/elements, namely P2P and adhoc networks using agents to drive those simulations. This work is done in collaboration with the MAIA team. The results have been extensively described in the PhD thesis of Julien Siebert [3].

• Study of service discovery protocols, contextual metrics in adhoc networks, and Service Discovery in adhoc networks using an hybrid model between cluster-like (WCPD) and MPR-based (OLSR) broadcasting. The results have been extensively described in the PhD thesis (Contributions for Advanced Service Discovery in Ad hoc Networks) of Tom Leclerc [2]. In this thesis, we consider service discovery in MANETs, that are a collection of devices that communicate with each other over a wireless medium. Such networks are formed spontaneously whenever devices are in transmission range without any preexisting infrastructure. The main characteristic of MANETs is the high dynamics of nodes (induced by the users moving around), the volatile wireless transmissions, the user behavior, the services and their usage. We’ve proposed a complete solution for service discovery in ad hoc networks, from the underlying network up to the service discovery itself. A first contribution, is the Stable Linked Structure Flooding (SLSF) protocol that creates stable based cluster structure and thereby provides scalable and efficient message dissemination. The second contribution is the Stable Linked Structure Routing (SLSR) protocol that uses the SLSF dissemination structure to enable routing capabilities. Using those protocols as basis, we propose to improve service discovery by additionally considering context awareness and adaptation.

• Context awareness and mobility/usage models

We contributed on improving simulations by coupling simulators and models that, together, can model and simulate the variety and richness of ad hoc related usage scenarios and their human characteristic. A guideline for all of our contributions was to be able to integrate and/or consider context and context awareness in both the proposed protocols and the related research tools and models. On one hand, The proposed protocols all have the capacity to adapt their efforts according to certain metrics, that represent the context. The simulator coupling architecture, on the other hand, permits to model and design scenarios in which the context, such as the service usages or the human behavior, has an impact and matters.

• Energy-constraint geolocalization, addressing, routing and management of wireless devices: a research collaboration with Fireflies RTLS was started in March 2009 and is ongoing. The initial work
has been extended in a joint work with the TRIO Team and leads towards finding a global energy-cost function, and life expectancy of the wireless sensor system.

In the future work, we plan to apply those results to Cyper Physical Systems, within the Aetournos (Airborne Embedded auTonomOUs Robust Network of Objects and Sensors) platform at Loria. We aim at developing cross-layer solutions to robust routing between flying drones.

We are also working inside a CPER project towards management solutions of wireless network sensors (project ECOSUR) used to control Smart Spaces.

6.8. Co-Simulation and multi-modeling

Participants: Laurent Ciarletta [contact], Julien Siebert, Tom Leclerc.

Vincent Chevrier (MAIA team, LORIA) and Tomas Navarette are external collaborators.


Participants: Laurent Ciarletta [contact], Julien Siebert.

Vincent Chevrier (MAIA team, LORIA) is an external collaborator.

this work has been extensively detailed in Julien Siebert’s PhD thesis [3] and partially in Tom Leclerc’s, with an application to ubiquitous adhoc networks and services.

This work has been done between the fields of ubiquitous networks and multi-agent based simulation. The main context is to study mutual influences existing between ubiquitous network performances and their users behaviours. We have highlighted the need for reusing and coupling modelling and simulation softwares together in order to simultaneously integrate several abstraction levels in the study. We target those needs by a multiagent approach and we propose a metamodel: AA4MM. The core idea in AA4MM is to build a society of models, simulators and simulation softwares that solves the core challenges of multimodelling and simulation coupling in an homogeneous perspective. AA4MM major contributions are the possibility to easily reuse, to make interoperable and modular existing heterogeneous models and softwares, to manage scale changes and a simulation algorithm fully decentralized. We apply this metamodel to the field of ubiquitous networks in order to target the question of mutual influences between networks performances and users behaviours.

6.8.2. Adaptive control of a complex system based on its multi-agent model

Participants: Laurent Ciarletta [contact], Julien Siebert.

Vincent Chevrier (MAIA team, LORIA) and Tomas Navarette are external collaborators and main investigators of this theme.

As a starting point, we are exploring how the behavior and other factors such as spatial and temporal dimensions are mutually influencing and the impact of parameters variability of our models in environment where collective behaviors can emerge [6]. We did comparison of five different models. These models are built upon the same individual behavior hypothesis of a collective phenomenon present in peer-to-peer file exchange networks: "free-riding". We studied a global analytical model and four multi-agent models. Multi-agent models include the space and time dimensions rarely seen in the literature discussing aggregated models of the collective phenomenon in question. We have demonstrated that one individual decision algorithm can lead to contradictory information.

Using these results, we want to build a control mechanism for a complex/dynamic system. Specifically, we want to evaluate the effectiveness of creating a control mechanism based on a multi-agent model of the system.

Multi-agent models can be adapted to that purpose since usual approaches using analytical models as a basis can be intractable when dealing with such systems; and if we consider that the available control actions are meant to be applied locally, a multi-agent model is necessary. We are currently working on a case study within the dynamic networks domain, namely the free-riding phenomenon present in peer-to-peer networks.
We propose an architecture that gathers information from the system and uses it to parametrize and tune a set of multi-agent models. The outcome of simulations is used to decide which control actions have to be applied to the system, in order to achieve a predefined control objective. We consider that we do not have complete information to characterize the state of the system.

### 6.9. Sensor networks management

**Participants:** Cyril Auburtin, Alexandre Boeglin, Olivier Festor, Abdelkader Lahmadi, Emmanuel Nataf [contact].

6LowPAN networks denotes many embedded devices interconnected by a variety of links ranging from wireless technologies such as 802.15.4, bluetooth, Low Power Wifi to wired technologies such as low power PLC. The common property of such networks is the limited resources of their nodes in terms of power, computing, memory and communication. The network could be described with thousands of devices with very limited internal and external resources and their communication channels are low-bandwith, high loss rate and volatile links subject to failure over time. These networks rely on the 6LowPAN protocol defined by the IETF as an adaptation layer for the IPv6 protocol to address their low power and lossy properties.

During the year 2011, we have started a research activity around the monitoring and security assessment of 6LowPAN networks. Our contributions are mainly as follows:

- We are developing a novel approach to assign monitoring roles in 6LowPAN networks using available local information provided by the routing layer. The resulting monitoring architecture is adaptive taking benefit from the reactivity of the routing protocol when dynamic changes occur due to connectivity or nodes mobility. Our first simulations results reveal that our assignment approach is more efficient, less aggressive and less resources consuming than its competitors.

- We have also designed and implemented a piggybacking technique to deliver monitoring report into existing packets traveling through 6LowPAN networks. In our solution, we have extended the IPv6 Hop-by-Hop extension header with a new option which contains status data of monitored nodes. This technique can reduce the number of packets and bytes sent across the network since there is no specific monitoring packets competing with existing traffic. Monitoring data shares the routing path of application data packets until it reaches a management node. We have applied our piggybacking technique to discover coap-enabled management agents. Each agent in the deployed wireless sensor network piggybacks its identifier into the RPL routing protocol messages until it reaches a manager node.

- Regarding security management of these networks, we have developed a stateless fuzzing tool for the 6LowPAN protocol [28]. The tool is build upon the Scapy packets manipulation library. It provides different mutation algorithms to be applied on 6LowPAN messages. These messages are defined by interaction scenarios described in an XML format.

- Related also to security, we have modelled an ontology for intrusion detection system in sensor networks [17]. The model exposes family of intrusions depending on their objectives. The service provided by the network, the communication channels and the security mechanisms are the main classes of the model.

### 6.10. High Security Lab

**Participants:** Alexandre Boeglin [contact], Olivier Festor, Mohamed Nassar.

The objective of the High Security Lab at INRIA Nancy Grant Est is to provide both the infrastructure and the legal envelope to researchers to perform sensitive security oriented experimentations. We do contribute to this laboratory by (1) designing and operating a large network telescope and (2) performing vulnerability assessment research, network data and malware collection and analysis.
During the year 2011, some maintenance tasks have been carried out on the High Security lab:

- the SDSL line, which previously had a capacity of 1Mbps, has been upgraded to a 2Mbps line, and traffic shaping rules have been added to the router, that allow honeypots to run alongside experiments, without impacting them,
- the storage capacity of our database server, which was starting to get full, has been multiplied by four, and existing data has been migrated to the new equipment.

A set of new experiments have also been deployed:

- a server has been dedicated to a new variant of SGNet, for the VAMPIRE project. This one specifically targets attacks on SIP services, which the other one cannot do,
- in collaboration with the INRIA Nancy Grant Est IT service, we started to log public (thus anonymous) DNS queries and responses made by the research center’s recursive DNS servers, to use the collected data as input set for experiments.

In 2011 we worked also on the automated analysis of malware traces to extract flow-level signatures of malware. We obtained early results regarding network flow-graphs and tested several clustering techniques to separate malware traffic.

6.11. Sensas

Participants: Cyril Auburtin, Alexandre Boeglin [contact], Olivier Festor.

The goal of the SensAS ADT, which started in 2011, is to propose applications based on wireless sensor networks, building upon work that has been done through the SensLab and SensTools projects.

The Madynes team is responsible of the SensMGT part of the SensAS project, which focuses on sensor network management and configuration applications.

First, we adapted the existing contiki-snmp implementation to the SensLab WSN430 nodes. We did so by (1) reducing the memory footprint of the code and (2) by implementing several SNMP MIBs.

To reduce the memory footprint, we had to disable some optional features and unused drivers of the Contiki OS.

The MIBs that we chose to implement were:

- the SNMPv2-MIB that provides generic system information,
- the IF-MIB that provides information and statistics about the network interface of the sensor,
- and the ENTITY-SENSOR-MIB, that provides access to the actual sensors data.

Then, we were facing a problem, as the Contiki versions provided by the SensTools project were only stable releases, and we found it difficult to track the development version of Contiki with them. We then decided to create our own WSN430 drivers and platform definition for Contiki, well integrated with the development repository, and reusing as much as possible of already existing code. Our next step in this direction will be to have our contribution officially integrated in the Contiki OS.

And finally, we devised and implemented a COAP server discovery protocol using the piggybacking technique, which allows every node that offers COAP resources to announce itself to the grounded root of the sensor network, without requiring the transmission of additional packets.
5. New Results

5.1. IP networks

Participants: Eitan Altman, Konstantin Avrachenkov.

5.1.1. Interdisciplinary study of the Internet access and of network neutrality

In our previous research we have identified large inefficiencies that occur when one allows one type of provider (e.g., access provider) to impose costs on another type of provider (e.g., content provider). This part in which E. Altman collaborated with P. Bernhard (INRIA project-team BIOCORE), S. Caron and G. Kesidis (both from Pennsylvania State Univ., USA), J. Rojas-Mora (Univ. of Barcelona, Spain), and S. Wong (Univ. of A Coruña, Spain) has now appeared in [96].

This investigation has been pursued in various directions. In [42], E. Altman, A. Legout (INRIA project-team PLANETE) and Y. Xu (Univ. Avignon/LIA) have studied a hierarchical structure of ISPs, the economic impact of some cashing placement policies, and more complex demand functions (the demands of users for content). In [33], E. Altman and the law specialist S. Wong (Univ. of A Coruña, Spain) analyze in cooperation with the economist J. Rojas-Mora (Univ. of Barcelona, Spain) the impact of legislation related to network neutrality on the quality of service for the end users.

5.1.2. Adaptive monitoring system for IP networks

The remarkable growth of the Internet infrastructure and the increasing heterogeneity of applications and users’ behavior make more complex the manageability and monitoring of ISP networks and raises the cost of any new deployment. The main consequence of this trend is an inherent disagreement between existing monitoring solutions and the increasing needs of management applications. In this context, in [62] K. Avrachenkov, I. Lassoued, A. Krifa and C. Barakat (all three from INRIA project-team PLANETE) present the design of an adaptive centralized architecture that provides visibility over the entire network through a network-wide cognitive monitoring system. Practically, given a measurement task and a constraint on the volume of collected information, the proposed architecture drives the sampling rates on the interface of network routers to achieve the maximum possible accuracy, while adapting itself to any change in network traffic conditions. The authors tune the system parameters with the help of FAST sensitivity test.

5.1.3. Size based scheduling

Size-based scheduling is a promising solution to improve the response time of small flows (mice) that have to share bandwidth with large flows (elephants). To do this, one important task is to track the size of the ongoing flows at the router. However, most of the proposed size-based schedulers either employ the trivial way of tracking the size information of all flows, or require changes at end-hosts. Hence, either they are not scalable or they increase complexity. In [55], E. Altman, D. Mon Divakaran (IIT Mandi, India) and P. Vicat-Blanc Primet (Lyatiss, France) have proposed a new way of performing size-based scheduling in a practical and scalable fashion, by identifying and ‘de-prioritizing’ elephants only at times of high load. They exploit TCP’s behavior by using a mechanism that detects a window of packets – called spikes - when the buffer length exceeds a certain threshold. This spike-detection is used to identify elephant flows and thereafter de-prioritize them. Two-level processor-sharing (TLPS) scheduling is employed to schedule flows in two queues, one with the high-priority flows, and the other with the de-prioritized flows. They show that the proposed mechanism not only improves the response time of mice flows in a scalable way, but also gives better response times to other flows by treating them preferentially as long as they do not overload the high-priority queue.
5.1.4. Accuracy of fluid models for bandwidth-sharing networks

Optimal control of stochastic bandwidth-sharing networks is typically difficult. In order to facilitate the analysis, deterministic analogues of stochastic bandwidth-sharing networks, the so-called fluid models, are often chosen for analysis, as their optimal control can be found more easily. The tracking policy translates the fluid optimal control policy back into a control policy for the stochastic model, so that the fluid optimality can be achieved asymptotically when the stochastic model is scaled properly. In [20], K. Avrachenkov, A. Piunovsky and Y. Zhang (both from the University of Liverpool, UK) study the efficiency of the tracking policy, that is, how fast the fluid optimality can be achieved in the stochastic model with respect to the scaling parameter. In particular, the result of [20] shows that, under certain conditions, the tracking policy can be as efficient as feedback policies.

5.1.5. Bootstrap method for simulating bandwidth sharing

In [71], E. Altman, T. Jimenez and J. Rojas-Mora (both from Univ. Avignon/LIA) identify difficulties in evaluating through simulations the expected transfer time of a file when several TCP connections share a common bottleneck buffer. The main difficulties are due to the fact that the file size over the Internet has been reported to have a Pareto distribution with parameter smaller than 1.5. This implies that the number of ongoing connections as well as the sojourn times have infinite variance. This has two implications: one cannot estimate the confidence intervals for simulation based on the CLT (central Limit Theory approach), and the duration of the simulations needed to get to steady state is very long. The authors show how to solve both problems by the use of the bootstrap approach.

5.2. Wireless communications

Participants: Sara Alouf, Eitan Altman, Konstantin Avrachenkov, Nicaise Choungmo Fofack, Mahmoud El Chamie, Majed Haddad, Manjesh Kumar Hanawal, Philippe Nain, Giovanni Neglia, Manoj Panda, Sreenath Ramanath.

5.2.1. Green networking

Green networking is a new trend in network design that is more aware of the impact of technology on the environment and on humans. Reducing energy has been so far the main concern in that approach, and much of the research has been devoted to understanding the tradeoffs between reducing energy and other performance measures such as coverage and delay.

For several years we have been contributing to this research effort, many of which have been summarized in the survey [67] by E. Altman in collaboration with G. S. Paschos (Center for Research and Technology, Hellas, Greece), P. Mannensalo (VTT, Finland), S. Stanczak (HH-Fraunchofer, Berlin, Germany) and L. Tassiulas (Univ. of Thessaly, Greece). In particular, much of the work involving members of MAESTRO that had appeared in previous years in conferences concerning energy saving in WiMax has now appeared in a journal publication [21] by A. P. Azad, S. Alouf, E. Altman, in cooperation with V. Borkar (TATA Inst. of Fundamental Research, Mumbai, India) and G. S. Paschos (Center for Research and Technology, Hellas, Greece).

In 2011 we started investigating policies for switching off base stations using two new tools: multimodularity and stochastic geometry. The latter has been used by E. Altman and M. K. Hanawal (also with Univ. Avignon/LIA), in cooperation with R. El-Azouzi (Univ. Avignon/LIA) and S. Shamai (Technion, Israel) in [98] to study the tradeoffs related to the uplink, and by E. Altman in cooperation with C. Hasan and J. M. Gorce (both from INSA-Lyon and INRIA project-team SWING) in [89] for the downlink. Optimal policies were obtained within the class of policies that switch off base stations with some fixed probability but independently of each other. To relax this restriction of independence and thus obtain even better policies, S. Ramanath and E. Altman, in collaboration with V. Kavitha (Univ. Avignon/LIA), have used in [69] the theory of multimodularity, which is the discrete counterpart of convexity. Among the most striking points in this research has been the observation in [98] that the conventional energy saving approach can have the opposite effect on the humans in the uplink: when the base station closest to a mobile phone is switched off (for energy
saving) then the mobile phone has to transmit with a larger power so as to reach a more remote base station. It turns out that the main source of radiation to the human brain is indeed the uplink transmission, which implies that switching off base stations could cause more exposure to radiation. This is of particular concern in view of the announcement by the World Health Organization (May 31, 2011) that cell phones cause cancer.

5.2.2. Cellular networks with continuous connectivity
In [65], S. Alouf and V. Mancuso (Institute IMDEA Networks, Madrid, Spain) analyze the power save and its impact on web traffic performance when customers adopt the continuous connectivity paradigm. Considering realistic http traffic, they evaluate the user access delay, the download time and the expected economy of energy in the cell. The model, validated through packet-level simulations, shows that dramatic energy save can be achieved by both mobile users and base stations. In case of Poisson arrivals, the aggregate behavior of a base station’s users is studied by means of a processor-shared queueing system [105]. The model can be used to maximize the base station energy savings under a given set of QoS performance constraints. With the participation of N. Choungmo Fofack, the work in [65] has been complemented with a sensitivity analysis [95]. The impact of model parameters on the performance and cost metrics is thoroughly assessed.

5.2.3. Power allocation in multicell networks
Power allocation to satisfy user demands in the presence of large number of interferers in a multicellular network is a challenging task. Further, power to be allocated depends upon the system architecture, for example, upon components like coding, modulation, transmit precoder, rate allocation algorithms, available knowledge of the interfering channels, etc. This calls for an algorithm via which each base station in the network can simultaneously allocate power to their respective users so as to meet their demands (when they are within the achievable limits), using whatever information regarding the other users is available. In [70], S. Ramanath, V. Kavitha (Univ. Avignon/LIA) and M. Debbah (SUPELEC) devise such an algorithm which is in fact universal: the proposed algorithm works from a fully cooperative setting to almost no cooperation and or for any configuration of modulation, rate allocation, etc. schemes. The algorithm asymptotically satisfies the user demands, running simultaneously and independently within a given total power budget at each base station. Further, it requires minimal information to achieve this: every base station needs to know its own users demands, its total power constraint and the transmission rates allocated to its users in every time slot. The authors formulate the power allocation problem in a system specific game theoretic setting, define system specific capacity region and analyze the proposed algorithm using ordinary differential equation (ODE) framework. Simulations confirm the effectiveness of the proposed algorithm.

5.2.4. Small cell networks
In [28], S. Ramanath and E. Altman, in collaboration with V. Kavitha (Univ. Avignon/LIA), characterize the performance of Picocell networks in the presence of moving users. They model various traffic types between base-stations and mobiles as different types of queues. They derive explicit expressions for the expected waiting time, service time and drop/block probabilities for both fixed and random velocity of mobiles. They obtain (approximate) closed-form expressions for optimal cell size when the velocity variations of the mobiles is small for both non-elastic and elastic traffic. They conclude from the study that, if the expected call duration is long enough, the optimal cell size depends mainly on the velocity profile of the mobiles, its mean and variance. It is independent of the traffic type or duration of the calls. Further, for any fixed power of transmission, there exists a maximum velocity beyond which successful communication is not possible. This maximum possible velocity increases with the power of transmission. Also, for any given power, the optimal cell size increases when either the mean or the variance of the mobile velocity increases.

5.2.5. New concepts in fair resource allocation
Fair resource allocation is usually studied in a static context, in which a fixed amount of resources is to be shared. In dynamic resource allocation one usually tries to assign resources instantaneously so that the average share of each user is split fairly. The exact definition of the average share may depend on the application, as different applications may require averaging over different time periods or time scales. Our main contribution is to introduce new refined definitions of fairness that take into account the time over which one averages the
performance measures. In [39] E. Altman, K. Avrachenkov and S. Ramanath examine how the constraints on the averaging durations impact the amount of resources that each user gets. The authors apply this new concept in [68] to spectrum allocation and indoor-outdoor femtocells.

5.2.6. Self organization in cellular networks

Time-slots and frequencies are contended in cellular networks and their allocation is determined by base stations. For scalability purposes the resource allocation is decentralized, so that base stations do not share their information with each other. Actions are often taken based on partial information on the system. In particular, the statistics of the channels are often not available. Scheduling decisions of a base station concerning mobiles in its cell cause interference in other cells and there is thus a need to dynamically adjust to interference and to converge to a satisfactory operation point. This has motivated a large amount of work on self-organization in cellular networks based on OFDMA. E. Altman, Z. Altman, R. Combes (both from Orange Labs, Issy les Moulineaux) and M. Haddad have written a series of papers on self-organization. In [53] and [51] self-organization in interference coordination is studied. In [50], R. Combes, Z. Altman and E. Altman, further propose and analyze a self-optimization method for coverage-capacity optimization in OFDMA networks with MIMO. Moreover, they study in [52] self-organization when adding relays so as to increase coverage. Static and dynamic resource sharing mechanisms are investigated. In the static case they use a queuing model to calculate the optimal resource sharing strategy and the maximal capacity of the network analytically. The influence of relay planning and number of deployed relays is investigated, and gains resulting from good planning are evaluated analytically. Self-optimizing dynamic resource allocation is tackled using a Markov Decision Process (MDP) model.


Self-organization has also been used in the past to obtain opportunistic scheduling in a way that achieves proportional fair resource sharing. In [23], R. Combes, Z. Altman (both from Orange Labs, Issy les Moulineaux) and E. Altman, have extended this to the general $\alpha$-fair concept. A dynamic choice of the factor $\alpha$ is proposed, which has the interpretation of trading optimality with fairness in a dynamic way.

5.2.7. Dynamic networks

In source routing, a complete path is chosen for a packet to travel from source to destination. While computing the time to traverse such a path may be straightforward in a fixed, static graph, doing so becomes much more challenging in dynamic graphs, in which the state of an edge in one timeslot (i.e., its presence or absence) is random, and may depend on its state in the previous time step. The traversal time is due to both time spent waiting for edges to appear and time spent crossing them once they become available. In [99], P. Nain in collaboration with A. Bar-Noy (City University of New York), P. Basu (Raytheon BBN Technologies), M. P. Johnson (Pennsylvania State University), F. Yu (City University of New York) and D. Towsley (University of Massachusetts at Amherst) computes the expected traversal time (ETT) for a routing path in a number of special cases of stochastic edge dynamics models, and for three edge failure models, culminating in a surprisingly challenging yet realistic setting in which the initial configuration of edge states for the entire path is known. We show that the ETT for this “initial configuration” setting can be computed in quadratic time, by an algorithm based on probability generating functions. The authors also give several linear-time upper and lower bounds on the ETT.

5.2.8. Sensor networks

In many application scenarios sensors need to calculate the average of some local values, e.g. of local measurements. A possible solution is to rely on consensus algorithms. In this case each sensor maintains a local estimate of the global average, and keeps improving it by performing a weighted sum of the estimates of all its neighbors. The number of iterations needed to reach an accurate estimate depends on the weights used at each sensor. K. Avrachenkov, G. Neglia and M. El Chamie have proposed a new average consensus algorithm, where each sensor selects its own weights on the basis of some local information about its neighborhood [45]. In realistic sensor network topologies, the algorithm shows faster convergence than other existing consensus protocols.
5.2.9. Delay and disruption-tolerant networks (DTNs)

5.2.9.1. Applying risk sensitive control to delay tolerant networks

When controlling the propagation of a message in DTNs, the objective is often to maximize the successful delivery probability of a message within a given deadline. It takes often the form of the expectation of the exponent of some integral cost. So far, models involving such costs have been solved by interchanging the order of expectation and the exponential function. While reducing the problem to a standard optimal control problem, this interchange is only tight in the mean-field limit obtained as the population tends to infinity. In [41] E. Altman, V. Kavitha (Univ. Avignon/LIA), F. De Pellegrini (Create-Net, Trento, Italy), V. Kamble (UC Berkeley, CA, USA) and V. Borkar (TATA Inst., Mumbai, India), identify a general framework from optimal control in finance, known as risk sensitive control, which allows handling the original (multiplicative) cost and obtaining solutions to several novel control problems in DTNs. New optimal control problems which consider the effect of wireless propagation path loss factor and the power constraints at the source and or the destination are proposed for DTNs within this framework. Optimal policies of non-threshold type are found.

5.2.9.2. Multiple destinations

In [73], C. Singh, A. Kumar and R. Sundaresan (all three from IISc Bangalore, India) in collaboration with E. Altman, use Markov Decision Processes to study optimal policies for propagation of contents in DTNs to multiple destinations. They obtain structural properties for a discretized system which allows them to derive the structure of optimal policies to the original problem.

5.2.9.3. Reliable unicast and multicast

In case the DTN does not deliver a packet within some time $T$, it has to be retransmitted. In [36] E. Altman and M. Panda, in collaboration with T. Chahed and A. Ali, (both from Telecom SudParis) and L. Sassatelli (Univ. Nice Sophia Antipolis/I3S), propose protocols for unicast and for multicast that render the connection reliable. These protocols include ACKs and retransmissions. The authors compute the value of $T$ that optimizes the throughput and address energy consumptions aspects.

5.2.9.4. Network coding

In [18], E. Altman studies, in cooperation with F. De Pellegrini (Create-Net, Trento, Italy), how to improve the performance of DTNs by adding network coding. The latter has the effect of efficiently adding spatial redundancy to the network. They identify the structure of optimal policies, which are shown not always to be of a threshold type.

G. Neglia, in collaboration with X. Zhang (Fordham University, New York) and J. Kurose (University of Massachusetts at Amherst), has published a survey on the application of network coding to DTNs [78].

5.2.9.5. Ferry based local area networks

Polling systems are used to model the Ferry assisted Wireless LANs and thereby to obtain the stationary workload performance. Not much theory is available for calculating the stationary workload of polling systems with arrivals in a continuum. In [103], V. Kavitha (Univ. Avignon/LIA) and E. Altman propose a discretization approach, by which the so-called “pseudo conservation law” of the discrete polling systems is utilized to derive the stationary performance of continuous polling systems. The continuous polling results are used in deriving optimal ferry routes.

5.2.9.6. Adaptive epidemic routing in DTNs

G. Neglia and R. Masiero (University of Padua, Italy) have explored a recently proposed optimization framework that relies on local sub-gradient methods and consensus algorithms. The research is described in MAESTRO 2010 activity report and has appeared in [66].

5.2.9.7. Routing in quasi-deterministic networks

G. Neglia, U. Acer (Bell labs Antwerp), P. Giaccone and S. Tarapiah (both from Politecnico di Torino, Italy) and D. Hay (Hebrew University of Jerusalem), have investigated routing in DTNs where the underlying node mobility is known in advance but can be modified by random effects. The research is described in MAESTRO 2010 activity report and has appeared in [35] and [94].
5.3. Information systems

Participants: Eitan Altman, Konstantin Avrachenkov, Nicaise Choungmo Fofack, Majed Haddad, Alain Jean-Marie, Dorian Mazauric, Philippe Nain, Marina Sokol.

5.3.1. Web crawler optimization

A typical web search engine consists of three principal parts: crawling engine, indexing engine, and searching engine. The work [19] by K. Avrachenkov and P. Nain, together with A. Dudin, V. Klimenok, and O. Semenova (all three from Belarusian State University, Belarus), aims to optimize the performance of the crawling engine. The crawling engine finds new web pages and updates existing web pages in the database of the web search engine. The crawling engine has several robots collecting information from the Internet. The authors first calculate various performance measures of the system (e.g., probability of arbitrary page loss due to the buffer overflow, probability of starvation of the system, average time waiting in the buffer). Intuitively, one would like to avoid system starvation and at the same time to minimize the information loss. The authors formulate the problem as a multi-criteria optimization problem and solve it in the class of threshold policies. The authors consider a very general web page arrival process modeled by Batch Marked Markov Arrival Process and a very general service time modeled by Phase-type distribution. The model has been applied to the performance evaluation and optimization of the crawler designed by INRIA MAESTRO team in the framework of the RIAM INRIA-Canon research project (see MAESTRO 2006 and 2007 activity reports).

5.3.2. PageRank node centrality

In [48] K. Avrachenkov and M. Sokol, together with D. Nemirovsky (former MAESTRO team member), E. Smirnova (INRIA project-team AXIS) and N. Litvak (University of Twente, The Netherlands), study a problem of quick detection of top-k Personalized PageRank (PPR) lists. This problem has a number of important applications such as finding local cuts in large graphs, estimation of similarity distance and person name disambiguation. The authors suggest that two observations are important when finding top-k PPR lists. Firstly, it is crucial that one detects fast the top-k most important neighbors of a node, while the exact order in the top-k list and the exact values of PPR are by far not so crucial. Secondly, by allowing a small number of “wrong” elements in top-k lists, one achieves great computational savings, in fact, without degrading the quality of the results. Based on these ideas, the authors propose Monte Carlo methods for quick detection of top-k PPR lists. We demonstrate the effectiveness of these methods on the Web and Wikipedia graphs, provide performance evaluation and supply stopping criteria.

5.3.3. Analysis of YouTube

E. Altman and M. Haddad, in collaboration with S.-E. Elayoubi (Orange Labs, Issy les Moulineaux), R. El-Azouzi, T. Jimenez and Y. Xu (all three from Univ. Avignon/LIA) have been investigating streaming protocols similar to the one used by YouTube. After preparing a survey on the state-of-the-art in [57], they used Ballot theorems in [106] in order to compute the starvation probabilities (these are the probability that the queue empties before completing to send a streaming application).

This work is carried out in the framework of the Grant with Orange Labs (see Section 6.3) on “Quality of Service and Quality of Experience”.

5.3.4. Peer-to-peer networks

5.3.4.1. Real-time control of contents download

In the course of the VOODDO project, the question of assessing the theoretical limits of prefetching information in real-time arose. Given a network bandwidth and a graph of documents, is it possible to download documents in advance, so that the document surfer is never blocked because of missing information? The problem is modeled using a “cops-and-robbers” game and some of its algorithmic properties are derived. This work of A. Jean-Marie and D. Mazauric is joint with F. Fomin (Univ. Bergen) and F. Giroire and N. Nisse (both from INRIA project-team MASCOTTE) [86].
5.3.4.2. P2P traffic classification

P2P downloads still represent a large portion of today’s Internet traffic. More than 100 million users operate BitTorrent and generate more than 30% of the total Internet traffic. According to the Wikipedia article about BitTorrent, the traffic generated by BitTorrent is greater than the traffic generated by Netflix and Hulu combined. Recently, a significant research effort has been done to develop tools for automatic classification of Internet traffic by application. The purpose of the work [47] by K. Avrachenkov and M. Sokol, together with A. Legout (INRIA project-team PLANETE) and P. Gonçalves (INRIA project-team RESO), is to provide a framework for subclassification of P2P traffic generated by the BitTorrent protocol. The general intuition is that users with similar interests download similar contents. This intuition can be rigorously formalized with the help of graph based semi-supervised learning approach. In particular, the authors propose to work with PageRank based semi-supervised learning method, which scales well with very large volumes of data.

5.3.4.3. BitTyrant

The success of BitTorrent has fostered the development of variants to its basic components. Some of the variants adopt greedy approaches aiming at exploiting the intrinsic altruism of the original version of BitTorrent in order to maximize the benefit of participating to a torrent. G. Neglia, D. Carra (University of Verona, Italy), P. Michiardi and F. Albanese (both from INSTITUT EURECOM) have studied BitTyrant, a recently proposed strategic client. The research is described in MAESTRO 2008 activity report. Results have been extended and supported by PlanetLab experiments in [22].

5.3.5. Content-centric networks

In [100] N. Choungmo Fofack, P. Nain and G. Neglia, together with D. Towsley (University of Massachusetts at Amherst), provide building blocks for the performance evaluation of Content Centric-like Networks (CCNs). In CCNs if a cache receives a request for a content it does not store, it forwards the request to a higher-level cache, if any, or to the server. When located, the document is routed on the reverse-path and a copy is placed in each cache along the path. In this work the authors consider a cache replacement policy based on Time-to-Lives (TTLs) like in a DNS network. A local TTL is set when the content is first stored at the cache and is renewed every time the cache can satisfy a request for this content (at each hit). The content is removed when the TTL expires. Under the assumption that requests follow a renewal process and the TTLs are exponential random variables, we determine exact formulas for the performance metrics of interest (average cache occupancy, hit and miss probabilities/rates) for some specific architectures (a linear network and a tree network with one root node and \(N\) leaf nodes). For more general topologies and general TTL distributions, an approximate solution is proposed. Numerical results show the approximations to be accurate, with relative errors smaller than \(10^{-3}\) and \(10^{-2}\) respectively for exponentially distributed and constant TTLs.

This work is carried out in the framework of the Grant with Orange Labs on “Content-centric networks” (Section 6.2).

5.4. Game theory applied to networking

Participants: Eitan Altman, Konstantin Avrachenkov, Majed Haddad, Manoj Panda, Giovanni Neglia.

5.4.1. Resource allocation in wireless networks

5.4.1.1. Power control

In [14] E. Altman, K. Avrachenkov and A. Garnaev (St. Petersburg State University, Russia), study power control for Gaussian interference channel in optimization and game frameworks. In the optimization framework there is a single decision maker who assigns network resources and in the game framework users share the network resources according to Nash equilibrium. The authors enhance the water-filling technique with explicit analytic solutions. The authors also provide an alternative simple proof of the convergence of the Iterative Water Filling Algorithm. Finally, the authors compare the non-cooperative approach with the cooperative approach and show that the non-cooperative approach results in a more fair resource distribution.
There has been a debate between those proposing protocols based on a centralized controller and those favoring decentralized protocols based on non-cooperative game theory. In [26] E. Altman and M. Haddad, together with S.-E. Elayoubi and Z. Altman (both from Orange Labs, Issy les Moulineaux), consider a situation where a base station lets mobiles take power control decisions in some system states and imposes actions in other states. The authors study how best to choose what information to make available and how mobiles should react.

5.4.1.2. Joint power and rate allocation

In [63] X. Lei and L. Cottatellucci (both from INSTITUT EURECOM) and K. Avrachenkov consider a block fading interference channels with partial channel state information and address the issue of joint power and rate allocation in a game theoretic framework. Resource allocation algorithms based on Bayesian games are proposed. The existence, uniqueness, and some stability properties of Nash equilibria are analyzed. For some asymptotic setting, closed-form expressions of Nash equilibria are also provided.

5.4.1.3. Jamming in wireless networks

Jamming is a form of a denial of service attack in which an adversary can degrade the quality of the reception by creating interference. One can study jamming both in the purpose of protecting a wireless network against such attack or, on the contrary, in order to efficiently disrupt the communications of some adversary. In both cases jamming is part of a conflict for which game theory is an appropriate tool. In [15] E. Altman, K. Avrachenkov and A. Garnaev (St. Petersburg State University), investigate the effect of partially available information in which the user does not even know whether or not the jammer is indeed present. The problem is formulated as a zero-sum game. The authors find the equilibrium strategies in closed-form and specify the range of sub-carriers where the user can expect the jamming attack.

5.4.1.4. Channel access

In WiFi networks, mobile nodes compete for accessing the shared channel by means of a random access protocol called Distributed Coordination Function (DCF), which is long term fair. Selfish nodes could benefit from violating the protocol and increasing their transmission probability. G. Neglia, I. Tinnirello and L. Giarré (University of Palermo, Italy) have been studying the interaction of selfish nodes in the last two years (the research activity is described in MAESTRO 2009 and MAESTRO 2010 activity reports). [31], [74] further extend the results to a heterogeneous scenario, where nodes have different requirements in terms of uplink/downlink ratios.

5.4.2. Network formation games

The continued growth of computer networks such as the Internet has raised the interest in understanding how networks get formed. The design of such networks is generally carried out by a large number of self-interested actors (users, Internet Service Providers ...), all of whom seek to optimize the quality and cost of their own operation. Previous works have addressed the “Network Formation” problem considering almost exclusively networks designed by selfish users, which can be consistently suboptimal. In [46] K. Avrachenkov and G. Neglia, together with J. Elias (University Paris Descartes), F. Martignon (University Paris-Sud 11), and L. Petrosyana (St. Petersburg State University, Russia), address the network formation issue using cooperative game theory, which permits to study ways to enforce and sustain cooperation among agents. Both the Nash bargaining solution and the Shapley value are investigated. After the comparison of these two approaches, the authors conclude that the Nash bargaining solution is more suitable to enforce cooperation in the network formation game in terms of cost allocation to users and computation time to get the solution.

5.4.2.1. Network design with socially-aware users

In many scenarios network design is not enforced by a central authority, but arises from the interactions of several self-interested agents. This is the case of the Internet itself. K. Avrachenkov and G. Neglia, in collaboration with J. Elias (University Paris Descartes) and F. Martignon (University Paris-Sud 11), have proposed two novel socially-aware network design games. The research has been described in MAESTRO 2010 activity reports. [24] extends the results for the case when users’ utility functions incorporate a socially-aware component.
5.4.2.2. Stochastic games for cooperative network routing

In [64] K. Avrachenkov, L. Maggi and L. Cottatellucci (both from INSTITUT EURECOM) consider a system where several providers share the same network and control the routing in disjoint sets of nodes. They provide connection toward a unique server (destination) to their customers. The objective is to facilitate the design of the available network links and their costs such that all network providers are interested in cooperating and none of them withdraw from the coalition. More specifically, the authors establish the framework of a coalition game by providing an algorithm to compute the transferable coalition values. As by-product, the authors apply the proposed algorithm to two-player games both in networks subject to hacker attacks and in epidemic networks.

5.4.2.3. Association games

Using tools from coalition game theory, E. Altman, in cooperation with C. Singh (IISc Bangalore, India), considers in [72] a wireless framework in which several mobile terminals can receive and decode the same signal of the base station, and where the cost for broadcasting is taken to be the transmission power. They begin by proposing various schemes to share the cost and study their properties. Then, they study the association with partial information: an arriving user knowing its location has to decide without knowledge of the location of the other users and their number whether to join the multicast tree and pay according to a given cost sharing scheme, or to have a unicast connection at a given cost. The unicast alternative that each mobile has, results in a limitation on the coverage (area covered by the multicast session) and on the capacity (number of mobiles connected to the multicast session). The authors derive the expected capacity and coverage as a function of the cost sharing policy. This work is extended in [58] to the case of several base stations by E. Altman in collaboration with C. Hasan and J. M. Gorce (both from INSA Lyon and INRIA project-team SWING).

5.4.3. Routing games

In [40], E. Altman, M. Panda and A. Estanisla (Master student at UPMC) study ring networks extensively used in both road traffic and telecommunications (in local area networks) in which each source with a given origin and destination on the ring, can split its traffic and send some part in one direction of the ring and some other part in the other direction. They compute the equilibria and find out that due to non-cooperation, much traffic is sent at equilibrium along long paths.

In [16], E. Altman, O. Pourtallier (INRIA project-team COPRIN), T. Jimenez (Univ. Avignon/LIA) and H. Kameda (Univ. Tsukuba, Japan) study a load balancing processor sharing problem. The classical framework of routing games turned out not to apply here. Indeed, it had been used to model situations where the flow from each class of users is split among paths without any information on the realization of the sizes of each packet. In contrast, in this paper, each individual knows its size. The authors have succeeded in computing the equilibrium within the new setting.

Collusion is the situation where several players decide to cooperate and to choose their actions as if they were a single player - each player maximizes the sum of utilities of that group instead of only its own utility. In [90], E. Altman in collaboration with Y. Hayel (Univ. Avignon/LIA) and H. Kameda (Univ. Tsukuba, Japan), has proposed various concepts that evaluate the impact of collisions. The authors have further studied collusions in routing games and identified situations where collusions are bad for all players: both those that collide loose in performance as well as those who remain independent.

5.5. Stochastic processes, queueing, control theory and game theory

Participants: Eitan Altman, Julien Gaillard, Majed Haddad, Alain Jean-Marie.

5.5.1. Convergence of rolling horizon control

In collaboration with E. Della Vecchia and S. Di Marco (both from National Univ. Rosario, Argentina), A. Jean-Marie has investigated the performance (convergence and error bounds) of the Rolling Horizon heuristic for optimal stochastic control and stochastic games in different modeling situations.
In the case of the long-term average expected gain, they have shown\cite{85} that convergence occurs whenever the value iteration algorithm converges. They have then considered zero-sum semi-Markov games with discounted payoff\cite{54,76}, for which they have proved geometric convergence under the usual assumptions of the literature.

5.5.2. Impulse control versus continuous control

Impulse control is a modeling framework of optimal control theory, in which the control actions can provoke instantaneous changes in the value of the state. For modelers, it has the features of both continuous-time and discrete-time models, and it can help understand which one to choose in a given optimization situation. A. Jean-Marie has studied the question in conjunction with K. Erdlenbruch (CEMAGREF), M. Tidball (INRA) and M. Moreaux (Univ. Toulouse 1). In a quite generic single-dimensional model, they show that the optimality of impulse policies with respect to “smooth” control policies is strongly related to a submodularity property of the instantaneous cost function\cite{101}.

5.5.3. Routing games

Several fundamental results have been obtained in routing games that model finite number of sources of traffic (players) who decide how to split the traffic among various paths. When the number of players is large, the Wardrop equilibrium concept is often used, where the problem is modeled as one with a continuum of decision makers where each has a negligible impact (non atomic game) on other's performance. E. Altman and his co-workers have studied the question of whether Wardrop equilibrium is a good approximation for a problem with finitely many players for which the Nash equilibrium is the solution concept. In\cite{38}, E. Altman, in collaboration with Z. Altman, R. Combes (both from Orange Labs, Issy les Moulineaux) and S. Sorin (Univ. Pierre and Marie Curie (UPMC)) establishes the convergence under mild convexity assumptions on the link costs (or delays). The proof is based on yet another fundamental result derived in that reference and that was later extended in\cite{44} by E. Altman in collaboration with O. Pourtallier (INRIA project-team COPRIN), T. Jimenez (Univ. Avignon/LIA) and H. Kameda (Univ. Tsukuba, Japan), that states that if there is some symmetry in a network then any Nash equilibrium will inherit the symmetric properties (for example, if two users have the same source and destination and the same demand then at equilibrium, they will send the same amount of traffic over each link).

In all the above work there is an assumption that the link cost (or delay) per packet is class independent (it depends on the flows through the link only through their sum). In the case of Wardrop equilibrium this assumption implies that the game has an equivalent global optimization problem whose solution coincides with the equilibrium. The link cost evaluated at some \(x\) in the equivalent problem is the integral of the original link cost (from 0 to \(x\)) and is in fact a potential. In the case of class dependent cost, that is, when the cost depends in other ways on the traffic of each class then the result of the integration may depend on the path and one cannot transform the problem to an equivalent optimization one. H. Kameda and J. Li (both from Univ. Tsukuba, Japan) in collaboration with E. Altman, identify in\cite{27} other class-dependent cost that have the property of a field, that is, it can be expressed as the gradient of a potential. They obtain the Wardrop equilibrium and study its properties.

Another difficulty occurs in rouging games when the paths available are not the same for all users. This is the case, in particular, when there are priorities. This problem is addressed in\cite{25} by J. Elias (University Paris Descartes), F. Martignon (University Paris-Sud 11), A. Capone (Politecnico di Milano, Italy) and E. Altman within an application to non-cooperative spectrum access in cognitive radio networks.

5.5.4. Bio-inspired paradigms

5.5.4.1. Epidemiology

For several years now, E. Altman has been developing techniques for dynamic optimal control and games in cooperation with S. Sarkar’s group from the University of Pennsylvania (which used to be part of the DAWN associated team with MAESTRO). This year this collaboration has resulted in three additional publications co-authored by M.H.R. Khouzani and S. Sarkar (both from Univ. of Pennsylvania, PA, USA) and E. Altman\cite{61,60,104}. All three papers use the Pontriagin maximal principle to derive the structure.
of optimal policies applied to a mean-field approximation of the problem. The first two papers do that in a context of optimal control theory while the third one does it in the context of a dynamic game.

5.5.4.2. Sequential anonymous games (SAG)

Sequential Anonymous Games (SAG) can be viewed as an extension of Markov Decision evolutionary games. In both formalisms there are many players modeled as a continuum number of players. A Markov chain is associated with each player. There are several types of players. The fraction of players in each class is called a global state and the state of the Markov chain of an individual is called the individual state. An individual chooses at some sequential decision opportunities actions. It earns some immediate reward (fitness) at each slot and moves with some probability to another individual state. In SAG, both the transition probabilities and the immediate fitness of an individual depend on its current state and action as well as on the current global system state. The latter evolves according to some function averaged over the fitness of the individuals in each class (the fraction of individuals in a class grows if they do better than those in the other classes). In [75] E. Altman investigates, in collaboration with P. Wiecek (Wroclaw Univ. of Technology, Poland), the case where the objective of an individual is to maximize either its total expected fitness during its life time or its expected average fitness. The authors establish the existence of equilibria and study its properties. Applications to power control have appeared in [32] by E. Altman, in collaboration with P. Wiecek (Wroclaw Univ. of Technology, Poland) and Y. Hayel (Univ. Avignon/LIA).

5.5.4.3. Markov decision evolutionary games (MDEGs)

Since his 2004 Infocom paper, E. Altman has been working on this novel paradigm. The model is similar to that of the previous paragraph (SAG) except that in MDEG both immediate fitness and transition probabilities depend linearly on the global state. This reflects a scenario where the interaction between a player and the rest of the population occurs through pairwise interactions: each player encounters from time to time a randomly chosen other player and it finds itself playing a matrix game with that player. The entries of the matrix corresponding to each player as well as the transition probabilities for each player depend on the individual states of the players. E. Altman has applied this model to the dynamic Hawk and Dove game, in which individuals have to choose the degree of aggressiveness in their behavior as a function of their energy state.

Below are three publications both with biological applications and applications to wireless communications (where depending on one’s remaining battery energy, one has to decide at what power to transmit). The first publication in [30], by E. Altman, H. Tembine (SUPELEC), R. El-Azouzi and Y. Hayel (both from Univ. Avignon/LIA), lays the foundations of MDEGs and presents the application to power control in which the individual state is the battery level of energy. The second publication in [59], by Y. Hayel (Univ. Avignon/LIA), E. V. Belmega (SUPELEC) and E. Altman, studies theoretical aspects that arise in case that the global state cannot represent the fractions of different populations but rather their actual size. The third publication in [97] by E. Altman, J. Gaillard, M. Haddad and P. Wiecek (Wroclaw Univ. of Technology, Poland) again studies MDEGs (as in the first paper), but restricts to policies that use static policies: the same mixed strategy is taken by a player at each state. The authors manage to compute explicitly the equilibrium in this game within this class of policies.

5.5.4.4. Delayed evolutionary games

Evolutionary game theory includes much theory on the description of the global system state as a function of the fitness of individuals. The models are often described through differential equations (e.g. the “replicator dynamics”). In many scenarios it is realistic to consider delays between the moment that one receives a given fitness till this is translated to a change in the population size. For example, if the lifetime of a computer is three years then an application that performs better with one computer may take more than a year till it is adopted by other users who do not have the same computer. In [30], H. Tembine (SUPELEC), E. Altman, R. El-Azouzi and Y. Hayel (both from Univ. Avignon/LIA) investigate instability phenomena that are introduced by the delay and derive necessary stability conditions.
6. New Results

6.1. Network Design and Optimization

Participants: Jean-Claude Bermond, Nathann Cohen, David Coudert, Frédéric Giroire, Dorian Mazauric, Joanna Moulierac, Nicolas Nisse, Ronan Pardo Soares, Issam Tahiri.

6.1.1. Backbone and Broadband Networks

Network design is a very wide subject that concerns all kinds of networks. We mainly study telecommunications networks which can be either physical networks (backbone, access, wireless, ...) or virtual (logical) ones. The objective is to design a network able to route a (given, estimated, dynamic, ...) traffic under some constraints (e.g. capacity) and with some quality of service (QoS) requirements. Usually the traffic is expressed as a family of requests with parameters attached to them. In order to satisfy these requests, we need to find one (or many) path(s) between their end nodes. The set of paths is chosen according to the technology, the protocol or the QoS constraints. For instance, optical backbones use the WDM technology to take better advantage of the capacity of the optical fibers often already installed. This is achieved through the multiplexing of several wavelength channels onto the same fiber. In that case a resource allocation is an optical channel, also called lightpath, which includes a path and wavelengths assigned to its links, one per link. If wavelength translation is performed in optical switching, then each channel may be assigned different wavelengths on the links of its path; otherwise the wavelength continuity imposes all the links to have the same wavelength. Of course, two lightpaths sharing a link must use different wavelengths on that link. The design can be done at the conception of the network (i.e. when conceiving a virtual network in MPLS where we have to establish virtual paths) or to adapt the network to changes (failures, new link, updates of routers, variation of traffic, ...). Finally there are various optimization criteria which differ according to the point of view: for a network user they are related to his/her satisfaction (minimizing delays, increasing available bandwidth, ...), while for a network operator, economics criteria like minimizing deployment and operating costs are more important.

This very wide topic is addressed by a lot of academic and industrial teams in the world. Our approach is to attack these problems with tools from Discrete Mathematics.

6.1.1.1. Traffic Grooming

In a WDM network, routing a connection request consists in assigning to this request a route in the physical network and a wavelength. When each request uses at most $1/C$ of the bandwidth of the wavelength, we say that the grooming factor is $C$. It means that on a given link of the network we can groom at most $C$ requests on the same wavelength. Under this constraint the objective can be either to minimize the number of wavelengths (related to the transmission cost) or to minimize the number of Add/Drop Multiplexers (ADM) used in the network (related to the cost of the nodes). During the last years, we have addressed this problem in various WDM network topologies with the goal of minimizing the total number of required ADMs.

This year, we considered the minimization of the number of ADMs in optical WDM bidirectional rings, considering symmetric shortest path routing and all-to-all unitary requests [24]. We formulate the problem in terms of graph decompositions, and state a general lower bound for all the values of the grooming factor $C$ and $N$, the size of the ring. We have studied exhaustively the cases $C = 1$, $C = 2$, and $C = 3$, providing improved lower bounds, optimal constructions for several infinite families, as well as asymptotically optimal constructions and approximations. We have also studied the case $C > 3$, focusing specifically on the case $C = k(k + 1)/2$ for some $k \geq 1$. We have also proposed optimal decompositions for several congruence classes of $N$ using the existence of some combinatorial designs.
6.1.1.2. Routing Reconfiguration and its Links with Graph Searching

In production networks, traffic evolution, failures and maintenance operations force to adapt regularly the current configuration of the network (virtual topology, routing of connections). The routing reconfiguration problem in WDM networks is thus to schedule the migration of established lightpaths from current routing to a new pre-computed one while minimizing service disruptions. We have shown in the past the relations between this problem and the graph searching problem (see also Section 6.4.3).

This year, we have continued studying the tradeoffs between the total number and the number of simultaneous interruptions that occurs during the reconfiguration process, proving in particular that the knowledge of one parameter does not help to optimize the other [28], [15]. We have also started investigating the influence of physical layer impairment constraints on the reconfiguration problem [74]. More precisely, using a new wavelength in a fiber of a WDM network forces to tune or recalibrate all already used wavelengths. We thus model the cost of using a new wavelength with a linear function of the number of already used wavelengths. We have then studied the problem of minimizing the cost of the reconfiguration according to this function. We have shown that this optimization problem is already NP-complete in a two-node network. We have also obtained general bounds and characterized instances for which the problem can be solved in polynomial time. We have additionally proposed and evaluated heuristics.

6.1.1.3. Green Networking

The minimization of ICT (Information and Communications Technologies) energy consumption has become a priority with the recent increase of energy cost and the new sensibility of public, governments and corporations towards energy consumption. ICT alone is responsible of 2% to 10% (depending on the estimations) of the world power consumption. For example, it is estimated that switches, hubs, routers account for 6 TWh per year in the US.

Several studies exhibit that the traffic load of the routers only has a small influence on their energy consumption. Hence, the power consumption in networks is strongly related to the number of active network elements, such as interfaces, line cards, base chassis, etc. In [78], [15], we have defined and modeled formally the problem of finding a routing that minimizes the (weighted) number of active network elements. We have proved that this problem is not in APX, that is there is no polynomial-time constant-factor approximation algorithm to solve it. We have obtained general bounds for this problem, and bounds for particular topologies such as trees, grids, and cliques. We have also proposed a heuristic algorithm offering good performance on real topologies. Last, we have analyzed the impact of energy efficient routing on the stretch factor and on fault tolerance.

We have also studied potential energy savings in fixed broadband wireless networks [77], [61]. See Section 6.1.2.1 for more details.

6.1.1.4. Xcast6 Treemap Islands

IP multicast is a protocol that deals with group communications with the aim of reducing traffic redundancy in the network. However, due to difficulty in deployment and poor scalability with a large number of multicast groups, IP multicast is still not widely deployed nor used on the Internet. Recently, Xcast6 and Xcast6 Treemap, the two network layer multicast protocols, have been proposed with complementary scaling properties to IP multicast: they support a very large number of active multicast sessions. However, the key limitation of these protocols is that they only support small multicast groups. To overcome this limitation, we have proposed the Xcast6 Treemap Island [96], a hybrid model of Application Layer Multicast (ALM) and Xcast6 that can work for large multicast groups. Our model has several advantages: ease of deployment, efficiency in bandwidth savings, no control message between end-host and router, zero multicast forwarding state at router and no need for a multicast address allocation protocol. In addition, this model is a potential service from which an ISP (Internet Service Provider) can get new revenue. We have shown the feasibility of our model by simulation and comparison with IP multicast and NICE protocols.

6.1.1.5. Time-Dependent Graphs - Applications to Transport Networks

In [70], we focus on time-dependent graphs which seem to be a good way to model transport networks. In the first part, we remind some notations and techniques related to time-dependent graphs. In the second one, we
introduce new algorithms to take into account the notion of probability related to paths in order to guarantee travelling times with a certain accuracy. We also discuss different probabilistic models and show the links between them.

Other results on multi-interface networks were obtained outside of MASCOTTE [37], [36], [65], [63], [66], [52], [20].

6.1.2. Wireless Networks

MASCOTTE has conducted an intense research effort on wireless access networks. From the technological and architectural point of view, the field is broad, from mesh (or multi-hop cellular) networks to ad-hoc and sensor networks. Nevertheless, many questions and approaches are generic from an algorithmic and structural prospect. In particular, we have considered three of the most prominent performance metrics for radio networks. Using combinatorial optimization and centralized algorithmic with a network design flavor, fast data gathering, call scheduling, transport capacity and energy consumption of the networks have been studied. Our approach is complementary with those developed in other INRIA project-teams such as PLANETE, MAESTRO, SWING, or POPS. The complementarity has been exploited through a joint Ph.D. between MAESTRO and MASCOTTE [15], through an ANR VERSO project in which MAESTRO, MASCOTTE, and SWING are involved, and through regular collaborations with POPS. At the international level, we cooperate with some groups in renowned research centers such as CTI of Patras in Greece, RWTH Aachen in Germany, Universities of Roma or Salerno in Italy, the Technion Institute in Israël, SFU in Vancouver, Canada, UFC Universidade Federal do Ceará, Fortaleza, Brazil, or the University of Sao Paulo in Brazil. We studied a wide range of issues of wireless networks, from the design of efficient cross-layer medium access, call scheduling and routing techniques to energy efficient optimization. We developed theoretical tools for integrating dynamic characteristics of the networks in the optimization models, and analyzing and evaluating dynamic networks. Some graph coloring problems motivated by channel assignment in wireless networks are detailed in Section 6.3.

6.1.2.1. Wireless Backhaul

We have investigated network optimization problems related to the design and configuration of fixed wireless microwave backhaul - the portion of the network infrastructure that provides interconnectivity between the access and the core networks. Unlike wired networks, the capacity of a microwave radio link is prone to variations, either due to external factors (e.g., weather) or by the action of the network operator. This fundamental difference raises a variety of new issues to be addressed appropriately. We concentrated on conceiving reliable fixed broadband wireless networks under outage probability constraints [60], [59]. We have developed a joint optimization of data routing and bandwidth assignment that minimizes the total renewal fees of licenses, while handling all the traffic requirements simultaneously. We have proposed a chance-constrained mathematical program taking into account unreliable channel conditions. This approach remains one of the main challenges of modern stochastic programming and it is still considered as very difficult and widely intractable. We have derived integer linear programming (ILP) counterparts for these chance-constrained programs and propose cutset-based valid inequalities to enhance the performance of ILP solvers. Computational results illustrate the price of reliability and present a comparative study on the performance of the different formulations. Moreover, we have been interested in potential energy savings in fixed broadband wireless networks by selectively turning off idle communication devices in low-demand scenarios [77], [61]. We have proposed a mathematical formulation of the problem relying on a fixed-charge capacitated network design (FCCND) problem, which is very hard to optimize. We have derived from this modeling heuristic algorithms producing feasible solutions in a short time. This work was done in collaboration with the SME 3Roam, and partially developed within the scope of the joint project RAISOM (Réseaux de collecte IP sans fil optimisés).

6.1.2.2. Wireless Mesh Networks

We have addressed the problem of computing the transport capacity of Wireless Mesh Networks (WMNs) dedicated to Internet access [26]. Routing and transmission scheduling have a major impact on the capacity provided to the clients. A cross-layer optimization of these problems allows the routing to take into account
contentions due to radio interference. We have presented a generic Mixed Integer Linear Programming (MILP) addressing gateway placement, routing, and scheduling optimizations in a WMN. We have then derived new optimization models that can take into account a large variety of radio interference models, and QoS requirements on the routing. We also provide efficient resolution methods that deal with realistic size instances. It allows to work around the combinatoric of simultaneously achievable transmissions and point out a critical region in the network bounding the network achievable capacity. Based upon strong duality arguments, it is then possible to restrict the computation to a bounded area. It allows for computing solutions very efficiently on large networks. We have then extended our models to deal with the dynamic characteristics of the network [75]. We have proposed a new robust optimization model that considers traffic demand uncertainty, in order to compute an optimal robust routing and bandwidth allocation in WMNs. We have presented a linear program efficiently solved by column generation, and we have quantified the price of robustness, i.e. the additional cost to pay in order to obtain a feasible solution for the robust scheme.

We have additionally investigated on the feasibility of providing network connectivity to vehicles over a predefined trajectory (trains, metros, urban buses, etc.) [14]. The communication between the vehicle and the infrastructure network is based only on WiFi technology. The contributions of this work are two-fold: 1) the horizontal handover (between WiFi access points) and 2) the design and analysis of an infrastructure network (backbone network plus WiFi access network) deployed along the trajectory of the vehicle.

6.1.2.3. Data Gathering

We have studied algorithmic and complexity issues originating from the problem of data gathering in wireless networks [56]. We give an algorithm to construct minimum makespan transmission schedules for data gathering when the communication graph is a tree network, the interference range is any integer \( m \geq 2 \), and no buffering is allowed at intermediate nodes. In the interesting case in which all nodes have to deliver an arbitrary non-zero number of packets, we provide a closed formula for the makespan of the optimal gathering schedule. Additionally, we consider the problem of determining the computational complexity of data gathering in general graphs and show that the problem is weakly NP-complete. On the positive side, we design a simple \((1 + 2/m)\) factor approximation algorithm for general networks. We have also considered the data gathering process in multi-hop wireless sensor networks [76], [57]. Wireless sensors networks (WSNs) are deployed to collect huge amounts of data from the environment. This produced data has to be delivered through sensor’s wireless interface using multi-hop communications toward a sink. The position of the sink impacts the performance of the wireless sensor network regarding delay and energy consumption especially for relaying sensors. Optimizing the data gathering process in multi-hop wireless sensor networks is, therefore, a key issue. We have addressed the problem of data collection using mobile sinks in a WSN. We provide a multi-objective optimization framework that studies the trade-off between energy consumption and delay of data collection. This framework provides solutions that allow decision makers to optimally design the data gathering plan in wireless sensor networks with mobile sinks.

6.1.3. P2P Networks

6.1.3.1. Performance Analysis of Distributed Storage Systems

Distributed or peer-to-peer storage solutions rely on the introduction of redundant data to be fault-tolerant and to achieve high reliability. To ensure long-term fault tolerance, the storage system must have a self-repair service that continuously reconstructs lost fragments of redundancy. The speed of this reconstruction process is crucial for the data survival. In [93], we propose a new analytical framework, based on queuing models, to estimate the repair time and the probability of data loss. This model takes into account the correlation of concurrent repairs. The models and schemes proposed are validated by mathematical analysis, extensive set of simulations, and experimentation using the Grid’5000 test-bed platform. Recently, the Regenerating Codes were proposed as an improvement over classical replication and erasure codes to introduce redundancy. These codes make a better use of the available bandwidth when reconstructing the missing information. In [50], we propose a new code based on a hybrid approach, Double Coding, and compare it to existing codes from the point of view of availability, durability and storage space.
6.1.3.2. Well Balanced Designs for Data Placement

In collaboration with MAESTRO. The problem we consider in [88] is motivated by data placement, in particular, data replication in video on-demand systems. We are given a set $V$ of $n$ servers and $b$ files (data, documents). Each file is replicated on exactly $k$ servers. A placement consists in finding a family of $b$ subsets of $V$ (representing the files) called blocks each of size $k$. Each server has some probability to fail and we want to find a placement which minimizes the variance of the number of available files. It was conjectured that there always exists an optimal placement (with variance better than that of any other placement for any value of the probability of failure). We show that the conjecture is true if there exists a well balanced design, that is a family of blocks, such that each $j$-element subset of $V$, $1 \leq j \leq k$, belongs to the same or almost the same number of blocks (difference at most one). The existence of well balanced designs is a difficult problem as it contains as subproblem the existence of Steiner systems. We completely solve the case $k = 2$ and give bounds and constructions for $k = 3$ and some values of $n$ and $b$.

6.1.3.3. Peer-Assisted Time-shifted Streaming Systems: Design and Promises

Time-shifted streaming (or catch-up TV) allows viewers to watch their TV programs within an expanded time window. In [71], we emphasize the challenging characteristics of time-shifted TV systems that prevent known delivery systems to be used. We model time-shifted TV as multiple-interval graph, then we present a Peer-Assisted Catch-Up Streaming system, namely PACUS, where a set of end users’ computers assists the server for the content delivery. We show in particular how the PACUS tracker server can be efficiently implemented for catch-up TV. We demonstrate the benefits of PACUS by simulations. We especially highlight that PACUS reduces the traffic at the server side with the advantages of lightweight and self-adaptive unstructured peer-to-peer systems.

6.2. Simulation and Optimization Tools

Participants: Olivier Dalle, Luc Hogie, Aurélien Lancin, Emilio Mancini, Juan-Carlos Maureira, Philippe Mussi, Van Dan Nguyen, Judicaël Ribault, Issam Tahiri.

The works related to simulation and optimization tools address two kinds of issues: issues related to the development of the tools and their associated methodology, and issues related to the use of these tools in order to investigate a particular problem or assess the performances or properties of a particular system.

Since 2005, MASCOTTE has been developing a discrete event simulation architecture, named OSA, whose aim is to investigate how new software engineering techniques, such as component-based frameworks or Aspect Oriented Programming can help improving the simulation methodology, especially in terms of software reuse [16], [46], [47]. After six years of research development, OSA entered in the process of being diffused in 2011. This process is supported by a two-year INRIA “Development Action” (ADT) funding. This first year was devoted to cleaning the code base and produce a public release with a significant effort placed on user documentation and tutorial (cf http://osa.inria.fr/).

Aside our efforts on the OSA project, we are strongly involved in the USS-SimGrid ANR funded project, whose aim is at developing an efficient simulation platform geared at Grid Computing and very large scale distributed computing architectures. In this project, we worked on two tasks:

- Monitoring and characterization of the workload of large scale distributed applications [68];
- Support for modeling Peer-to-peer applications in the SimGrid simulator (originally designed for modeling grid-computing platforms).

We also pursued our involvement in the Discrete Event Systems Specification (DEVS) standardization effort [81], [82]. This formalism has reached a strong agreement amongst the community, but it still lacks implementation standard. Since OSA is aimed at providing better support for methodology, we consider necessary to support DEVS and participate to this effort. Our particular focus was on techniques and Architecture Description Languages (ADLs) for describing very large models of distributed applications [69].
Regarding our works on simulation studies and application-oriented developments, this year was the conclusion of our effort on the Internet on Rails project [14]. In this project, we studied and designed, both by means of simulations and experimentations, a low cost communication architecture based on IEEE 802.11 WiFi to provide high quality Internet access onboard high speed trains.

6.3. Graph Theory

Participants: Julio Araújo, Jean-Claude Bermond, Nathann Cohen, Frédéric Giroire, Frédéric Havet, František Kardoš, Ana Karolinna Maia, Dorian Mazauric, Remigiusz Modrzejewski, Leonardo Sampaio, Michel Syska.

MASCOTTE principally investigates applications in telecommunications via Graph Theory (see other objectives). However it also studies a number of theoretical problems of general interest. Our research mainly focused on graph coloring and some other problems arising from networks problems.

6.3.1. Graph Coloring

Coloring and edge-coloring are two central concepts in Graph Theory. There are many important and long-standing conjectures in these areas. We are trying to make advances towards such conjectures, in particular Steinberg’s conjecture, the List coloring Conjecture and the Acyclic Edge-Coloring Conjecture.

We are also interested in coloring problems arising from some practical problems: improper coloring, \( L(p,q) \)-labeling, directed star arboricity and good edge-labelling. The first two are both motivated by channel assignment and the last two by problems arising in WDM networks. For many practical problems are posed in a dynamic setting, we study on-line coloring and list coloring.

We also study some other variants of coloring like non-repetitive coloring or frugal coloring.

For all the coloring problems, we also consider the associated algorithmic problem, which consists in designing algorithms for finding the minimum number of colors of a coloring of a given graph. Algorithmic results on graph coloring are presented in Section 6.4.

The most classical notion of coloring (of edges or vertices) is the one of proper coloring, in which we insist on two adjacent elements to have distinct colors. However, it is usual to consider additional constraints, as well as relaxed constraints. For each variant of coloring, one can consider, its list version in which every element \( x \) is given a list \( L(x) \) of prescribed colors. A graph is said to be \( L \)-colorable if is has an \( L \)-coloring (fulfilling the constraints) such that \( x \in L(x) \) for all element \( x \). The choosability of a graph \( G \) is the smallest integer \( k \) for which \( G \) has an \( L \)-coloring whenever \( |L(x)| \geq k \) for all elements \( x \).

6.3.1.1. Coloring Graphs with Few Crossings

The famous Four Color Theorem states that every planar graph can be properly colored with 4 colors and Thomassen Five Color Theorem states that the choosability of every planar graph is at most 5. Hence, a natural question is to ask about the chromatic number and choosability of graphs with few crossings. In [38], we disprove a conjecture of Oporowski and Zhao stating that every graph with crossing number at most 5 and clique number at most 5 is 5-colorable. However, we show that every graph with crossing number at most 4 and clique number at most 5 is 5-colorable. We also show some colorability results on graphs that can be made planar by removing few edges. In particular, we show that if there exists three edges whose removal leaves the graph planar then it is 5-colorable. In [90], we show that every graph with two crossings is 5-choosable. We also prove that every graph which can be made planar by removing one edge is 5-choosable.

Another famous theorem on planar graphs is the one of Grötzsch, which says that every planar graph with no cycle of length 3 can be properly 3-colored. Steinberg’s Conjecture (1976) asserts that a graph with no cycles of length 4 or 5 is 3-colorable. Many approaches have been used towards this conjecture. We considered the following one in which, we relax the constraints on the color classes. Instead of insisting on them be independent sets, we allow them to induce a graph with some bounded degree. A graph \( G = (V,E) \) is said to be \((i,j,k)\)-colorable if its vertex set can be partitioned into three sets \( V_1, V_2, V_3 \) such that the graphs \( G[V_1], G[V_2], G[V_3] \) induced by the sets \( V_1, V_2, V_3 \) have maximum degree at most \( i,j,k \) respectively. Under
this terminology, Steinberg’s Conjecture says that every graph with no cycle of length 4 or 5 is \((0, 0, 0)\)-colorable. In [91], we prove that every graph of \(\mathcal{G}\) is \((2, 1, 0)\)-colorable and \((4, 0, 0)\)-colorable.

6.3.1.2. Acyclic, Linear and Frugal Colorings

A classical constraint added to a proper coloring is that at least three colors appears on each cycle, in which case we speak about acyclic coloring. In other words, the graph induced by the elements of any two color classes is a forest. The acyclic chromatic index of a graph \(G\), denoted \(\chi'_{ac}(G)\) is the minimum \(k\) such that \(G\) admits an acyclic edge-coloring with \(k\) colors. The famous Acyclic Edge-Coloring Conjecture asserts that \(\chi'_{ac}(G) = \Delta(G) + 2\), where \(\Delta(G)\) is the maximum degree of the graph. In [21], we conjecture that if \(G\) is planar and \(\Delta(G)\) is large enough then \(\chi'_{ac}(G) = \Delta(G)\). We settle this conjecture for planar graphs with girth at least 5. We also show that \(\chi'_{ac}(G) \leq \Delta(G) + 12\) for all planar \(G\).

Even stronger constraints are the following: a proper coloring of a graph is 2-frugal (resp. linear) if the graph induced by the elements of any two color classes is of maximum degree 2 (resp. is a forest of paths). In [29], we improve some bounds on the 2-frugal choosability and linear choosability of graphs with small maximum average degree.

6.3.1.3. Coloring of Plane Graphs with Constraints on the Faces

We studied several variants of vertex and edge colorings of plane graphs insisting some constraints on the faces.

A face of a vertex colored plane graph is called loose if the number of colors used on its vertices is at least three. The looseness of a plane graph \(G\) is the minimum \(k\) such that any surjective \(k\)-coloring involves a loose face. In [35], we prove that the looseness of a connected plane graph \(G\) equals the maximum number of vertex disjoint cycles in a dual graph \(G^*\) increased by 2. We also show upper and lower bounds on the looseness of graphs based on the number of vertices, the edge connectivity, and the girth of the dual graph. These bounds improve the result of Negami for the looseness of plane triangulations. We also present infinite classes of graphs where the equalities are attained.

A vertex coloring of a 2-connected plane graph \(G\) is a strong parity vertex coloring if for every face \(f\) and each color \(c\), the number of vertices incident with \(f\) colored by \(c\) is either zero or odd. Czap et al. [Discrete Math. 311 (2011) 512–520] proved that every 2-connected plane graph has a proper strong parity vertex coloring with at most 118 colors. In [34], we improve this upper bound for some classes of plane graphs.

A facial parity edge coloring of a connected bridgeless plane graph is such an edge coloring in which no two face-adjacent edges (consecutive edges of a facial walk of some face) receive the same color, in addition, for each face \(f\) and each color \(c\), either no edge or an odd number of edges incident with \(f\) is colored with \(c\). From Vizing’s theorem it follows that every 3-connected plane graph has a such coloring with at most \(\Delta^* + 1\) colors, where \(\Delta^*\) is the size of the largest face. In [33] we prove that any connected bridgeless plane graph has a facial parity edge coloring with at most 92 colors.

A sequence \(r_1, r_2, \ldots, r_{2n}\) such that \(r_i = r_{n+i}\) for all \(1 \leq i \leq n\), is called a repetition. A sequence \(S\) is called non-repetitive if no block (i.e. subsequence of consecutive terms of \(S\)) is a repetition. Let \(G\) be a graph whose edges are colored. A trail is called non-repetitive if the sequence of colors of its edges is non-repetitive. If \(G\) is a plane graph, a facial non-repetitive edge-coloring of \(G\) is an edge-coloring such that any facial trail (i.e. trail of consecutive edges on the boundary walk of a face) is non-repetitive. We denote \(\pi_f'(G)\) the minimum number of colors of a facial non-repetitive edge-coloring of \(G\). In [41], we show that \(\pi_f'(G) \leq 8\) for any plane graph \(G\). We also get better upper bounds for \(\pi_f'(G)\) in the cases when \(G\) is a tree, a plane triangulation, a simple 3-connected plane graph, a hamiltonian plane graph, an outerplanar graph or a Halin graph. The bound 4 for trees is tight.

6.3.1.4. Improper Coloring

In [85] and [48], we study a coloring problem motivated by a practical frequency assignment problem and up to our best knowledge new. In wireless networks, a node interferes with the other nodes the level of interference depending on numerous parameters: distance between the nodes, geographical topography, obstacles, etc. We model this with a weighted graph \(G\) where the weights on the edges represent the noise (interference) between
the two end-nodes. The total interference in a node is then the sum of all the noises of the nodes emitting on the same frequency. A weighted $t$-improper $k$-coloring of $G$ is a $k$-coloring of the nodes of $G$ (assignment of $k$ frequencies) such that the interference at each node does not exceed some threshold $t$. The Weighted Improper Coloring problem, that we consider here consists in determining the weighted $t$-improper chromatic number defined as the minimum integer $k$ such that $G$ admits a weighted $t$-improper $k$-coloring. We also consider the dual problem, denoted the Threshold Improper Coloring problem, where given a number $k$ of colors (frequencies) we want to determine the minimum real $t$ such that $G$ admits a weighted $t$-improper $k$-coloring. We show that both problems are NP-hard and first present general upper bounds; in particular we show a generalization of Lovász’s Theorem for the weighted $t$-improper chromatic number. We then show how to transform an instance of the Threshold Improper Coloring problem into another equivalent one where the weights are either 1 or $M$, for a sufficient big value $M$. Motivated by the original application, we study a special interference model on various grids (square, triangular, hexagonal) where a node produces a noise of intensity 1 for its neighbors and a noise of intensity $1/2$ for the nodes that are at distance 2. Consequently, the problem consists of determining the weighted $t$-improper chromatic number when $G$ is the square of a grid and the weights of the edges are 1, if their end nodes are adjacent in the grid, and $1/2$ otherwise. Finally, we model the problem using linear integer programming, propose and test heuristic and exact Branch-and-Bound algorithms on random cell-like graphs, namely the Poisson-Voronoi tessellations.

6.3.1.5. On-line Coloring

Several on-line algorithms producing colorings have been designed. The most basic and most widespread one is the greedy algorithm. The largest number of colours that can be given by the greedy algorithm on some graph, is called its Grundy number. Determining the Grundy number of a graph is NP-hard even for $P_4$-free graphs, while it is polynomial-time solvable for $P_4$-free graphs. In [19], we define a new class of graphs, namely the fat-extended $P_4$-laden graphs, which intersects the class of $P_5$-free graphs and strictly contains the one of $P_4$-free. We show a polynomial-time algorithm to determine the Grundy number of such graphs. It implies that the Grundy number can be computed in polynomial time for most graph classes defined in terms of containing few $P_4$’s: $P_4$-reducible, extended $P_4$-reducible, $P_4$-sparse, extended $P_4$-sparse, ...

In [94], we study a game version of greedy coloring. Given a graph $G = (V,E)$, two players, Alice and Bob, alternate their turns in choosing uncolored vertices to be colored. Whenever an uncolored vertex is chosen, it is colored by the least positive integer not used by any of its colored neighbors. Alice’s goal is to minimize the total number of colors used in the game, and Bob’s goal is to maximize it. The game Grundy number of $G$ is the number of colors used in the game when both players use optimal strategies. It is proved in this paper that the maximum game Grundy number of forests is 3, and the game Grundy number of any partial 2-tree is at most 7. We also gave some complexity results on $b$-colorings, which is a manner of improving colorings on-line [43].

6.3.1.6. Other Results on Graph Coloring

In [18], we aim at characterizing the class of graphs that admit a good edge-labelling. Such graphs are interesting, as they correspond to set of requests in UPP-digraphs (those in which there is at most one dipath from a vertex to another) for which the minimum number of wavelengths is equal to the maximum load. This implies that the problem can be solved efficiently. First, we exhibit infinite families of graphs for which no good edge-labelling can be found. We then show that deciding if a graph admits a good edge-labelling is NP-complete. Finally, we give large classes of graphs admitting a good edge-labelling: $C_3$ -free outerplanar graphs, planar graphs of girth at least 6, subcubic $C_3$, $K_2$, 3 -free graphs.

A wheel is a graph formed by a chordless cycle and a vertex that has at least three neighbors in the cycle. We prove in [83] that every 3-connected graph that does not contain a wheel as a subgraph is in fact minimally 3-connected. We prove that every graph that does not contain a wheel as a subgraph is 3-colorable. We were then told that this result was already proved by Thomassen, though with a different proof.

Gallai-Hasse-Roy-Vitaver Theorem states that every $n$-chromatic digraph contains a directed path of order $n$. Let $f(k)$ be the smallest integer such that every $f(k)$-chromatic digraph contains every oriented tree of order $k$. Burr proved that $f(k) \leq (k - 1)^2$ and conjectured $f(k) = 2n - 2$. In [84], we give some sufficient conditions
for an $n$-chromatic digraphs to contains some oriented tree. In particular, we show that every acyclic $n$-chromatic digraph contains every oriented tree of order $n$. We also show that $f(k) \leq k^2/2 - k/2 + 1$. Finally, we consider the existence of antidirected trees in digraphs. We prove that every antidirected tree of order $k$ is contained in every $(5k - 9)$-chromatic digraph. We conjecture that if $|E(D)| > (k - 2)|V(D)|$, then the digraph $D$ contains every antidirected tree of order $k$. This generalizes Burr’s conjecture for antidirected trees and the celebrated Erdős-Sós Conjecture. We give some evidences for our conjecture to be true.

6.3.2. Matchings and Independent Sets

Matchings and independent sets are important substructures which appears in many problems. In particular, color classes of vertex-colorings and edge-colorings are independent sets and matchings, respectively.

In [45], we show that every (sub)cubic $n$-vertex graph with sufficiently large girth has fractional chromatic number at most 2.2978 which implies that it contains an independent set of size at least 0.4352$n$. Our bound on the independence number is valid to random cubic graphs as well as it improves existing lower bounds on the maximum cut in cubic graphs with large girth.

In [39], we show that every cubic bridgeless graph $G$ has at least $2|V(G)|/3656$ perfect matchings. This confirms an old and celebrated conjecture of Lovász and Plummer in the 1970’s. This improves the first superlinear bound given in [40].

6.3.3. Hypergraphs

Hypergraphs, also called set systems, are a natural generalization of graphs. In a graph an edge is set of two vertices, while in a hypergraph an edge is a set of any size. It turns out to be an important notion in database theory. A digraph is $\alpha$-acyclic if it can be reduced to the null hypergraph by successively removing either a vertex which in at most one edge or an edge included in another. It is one of the possible generalizations of forest to hypergraphs. The $\alpha$-arboricity of a hypergraph $H$ is the minimum number of $\alpha$-acyclic hypergraphs that partition the edge set of $H$. In [23], the $\alpha$-arboricity of the complete 3-uniform (every edge is a set of 3 vertices) hypergraph is determined completely.

In [80], we generalize the concept of line digraphs to line dihypergraphs. We give some general properties in particular concerning connectivity parameters of dihypergraphs and their line dihypergraphs, like the fact that the arc connectivity of a line dihypergraph is greater than or equal to that of the original dihypergraph. Then we show that the De Bruijn and Kautz dihypergraphs (which are among the best known bus networks) are iterated line digraphs. Finally we give short proofs that they are highly connected.

6.3.4. Miscellaneous

6.3.4.1. Zagreb Indices

The first and second Zagreb indices of a graph are defined by $M_1 = \sum_{v \in V(G)} d(v)^2$ and $M_2 = \sum_{uv \in E(G)} d(u)d(v)$, respectively. They are used in chemistry where it represents properties of molecules. In [17], we present some classes of graphs with prescribed degrees that satisfy $M_1/n \leq M_2/m$, where $M_1$ and $M_2$ are the first and second Zagreb indices. We also prove that for any $\Delta \geq 5$, there is an infinite family of graphs of maximum degree $\Delta$ such that the inequality is false. Moreover, we give alternative and slightly shorter proof of this inequality for trees and unicyclic graphs.

6.3.4.2. Induced Decomposition

An induced $H$-decomposition of a graph $G$ is a partition $(E_1, \cdots, E_k)$ of its edge set $E(G)$, such that the graph induced by each $E_i$, $1 \leq i \leq k$, is a copy of $H$. Bondy and Szwarcfiter asked for the maximum number $\text{ex}(n, H)$ of edges in a graph on $n$ vertices which admits an induced $H$-decomposition. In [13], we prove that for every non-empty graph $H$, $\text{ex}(n, H) = n(n - 1)/2 - o(n^2)$.

6.4. Algorithms

Participants: Julio Araújo, Janna Burman, Nathann Cohen, David Coudert, Gianlorenzo d’Angelo, Frédéric Giroire, František Kardoš, Dorian Mazaure, Nicolas Nisse, Stéphane Pérennes, Ronan Pardo Soares, Leonardo Sampaio.
MASCOTTE is also interested in the algorithmic aspects of Graph Theory. In general we try to find the most efficient algorithms to solve various problems of Graph Theory and telecommunication networks.

6.4.1. Coloring Graphs

Almost all graph coloring problems are NP-hard and most of them are even hard to approximate. Hence, to solve them efficiently, we aim at designing general exponential-time algorithms as well as polynomial-time algorithms for special classes. This is exemplified by the following results.

6.4.1.1. L(p, q)-labeling

An L(p, q)-labeling of G is an integer assignment f to the vertex set V(G) such that |f(u) − f(v)| ≥ p, if u and v are adjacent, and |f(u) − f(v)| ≥ q, if u and v have a common neighbor. Such a concept is a modeling of a simple channel assignment, in which the separation between channels depends on the distance. The goal is to find an L(p, q)-labeling f of G with minimum span (i.e. max {f(u) − f(v), u, v ∈ V(G)}). It is well known that for all k ≥ 4, deciding if a graph has an L(p, 1)-labeling with minimum span k is NP-complete. In [42], we present exact exponential time algorithms that are faster than existing ones.

6.4.1.2. Counting and Enumerating Total and Edge Colorings

In [89], we are interested in computing the number of edge colorings and total colorings of a graph. We prove that the maximum number of k-edge-colorings of a k-regular graph on n vertices is \(k \cdot (k - 1)^{n/2}\). Our proof is constructible and leads to a branching algorithm enumerating all the k-edge-colorings of a k-regular graph using a time \(O^*((k - 1)^{n/2})\) and polynomial space. In particular, we obtain a algorithm on time \(O^*(2^{n/2}) = O^*(1.4143^n)\) and polynomial space to enumerate all the 3-edge colorings of a cubic graph, improving the running time of \(O^*(1.5423^n)\) of the algorithm due to Golovach et al. We also show that the number of 4-total-colorings of a connected cubic graph is at most \(3.2^{3n/2}\). Again, our proof yields a branching algorithm to enumerate all the 4-total-colorings of a connected cubic graph.

6.4.1.3. Coloring Graphs of Special Classes

For some coloring problems that are known to be NP-hard for general graphs, we give some polynomial-time algorithms for the restriction to some graph classes. These graph classes defined in terms of forbidden induced subgraphs. In [95], [79], we provide linear algorithms for coloring \(P_3\)-free graphs. In [58], we obtain polynomial time algorithms to determine the acyclic chromatic number, the star chromatic number and the harmonious chromatic number of \((q, q - 4)\)-graphs. Such graphs are those such that no set of at most \(q\) vertices induces more than \(q - 4\) distinct \(P_4\)'s.

6.4.2. Complexity and Computation of Graph Parameters

We used graph theory to model various networks’ problems. In general we study their complexity and then we investigate the structural properties of graphs that make these problems hard or easy. In particular, we try to find the most efficient algorithms to solve the problems, sometimes focusing on specific graph classes where the problems are polynomial-time solvable.

6.4.2.1. Path Vertex Cover

A subset \(S\) of vertices of a graph \(G\) is called a k-path vertex cover if every path of order \(k\) in \(G\) contains at least one vertex from \(S\). The k-path vertex cover problem consists in finding such a set with minimum cardinality in \(G\). In [25], it is shown that this problem is NP-complete for each \(k \geq 2\) while it can be solved in linear-time in trees. The particular case of \(k = 3\) is studied in [44], where an exact algorithm is given with running time \(O^*(1.5171^n)\) in \(n\)-node graphs. In [44], we also design a polynomial time randomized approximation algorithm with an expected approximation ratio of \(\frac{\sqrt{33}}{3}\) for the minimum 3-path vertex cover.
6.4.2.2. Convexity in Graphs

The geodesic convexity of graphs naturally extends the notion of convexity in euclidean metric spaces. A set $S$ of vertices of a graph $G = (V, E)$ is convex if any vertex on a shortest path between two vertices of $S$ also belongs to $S$. The convex hull of $S \subseteq V$ is the smallest convex set containing $S$. Finally, a hull set of a graph is a set of vertices the convex hull of it is $V$. The hull number of a graph $G$ is the minimum size of a hull set in $G$. In [86], [49], we prove that computing the hull number is NP-complete in bipartite graphs. We also provide bounds and design various polynomial-time algorithms for this problem in different graph classes such as co-bipartite graphs, $P_4$-sparse graphs, etc.

6.4.2.3. Induced Subdivision in Digraphs

In [51], we consider the following problem for oriented graphs and digraphs: Given an oriented graph (digraph) $G$, does it contain an induced subdivision of a prescribed digraph $D$? The complexity of this problem depends on $D$ and on whether $H$ must be an oriented graph or is allowed to contain 2-cycles. We give a number of examples of polynomial instances as well as several NP-completeness proofs.

6.4.2.4. Circuits in Grids

A circuit in a simple undirected graph $G = (V, E)$ is a sequence of vertices $\{v_1, v_2, \ldots, v_{k+1}\}$ such that $v_1 = v_{k+1}$ and $\{v_i, v_{i+1}\} \in E$ for $i = 1, \ldots, k$. A circuit $C$ is said to be edge-simple if no edge of $G$ is used twice in $C$. In [30], we study the following problem: which is the largest integer $k$ such that, given any subset of $k$ ordered vertices of an infinite square grid, there exists an edge-simple circuit visiting the $k$ vertices in the prescribed order? We prove that $k = 10$. To this end, we first provide a counterexample implying that $k < 11$. To show that $k \geq 10$, we introduce a methodology, based on the notion of core graph, to reduce drastically the number of possible vertex configurations, and then we test each one of the resulting configurations with an ILP solver.

6.4.3. Graph Searching, Cops and Robber Games

Pursuit-evasion encompasses a wide variety of combinatorial problems related to the capture of a fugitive residing in a network by a team of searchers. The goal consists in minimizing the number of searchers required to capture the fugitive in a network and in computing the corresponding capture strategy. This can also be viewed as cleaning the edges of a contaminated graph. We investigated several variants of these games.

6.4.3.1. Process Number and Routing Reconfiguration in WDM Networks

Graph searching, where the fugitive is arbitrary fast and moves simultaneously to the searchers, has been widely studied for its close relationship with graph decompositions. More recently, a variant of graph searching, namely the graph processing game, has been widely studied as a model for the routing reconfiguration in WDM networks (see Section 6.1.1.2). In [32], we give a linear time (resp. polynomial-time) algorithm to recognize graphs (resp., digraphs) with process number at most 2, along with a characterization in terms of forbidden minors, and a structural description. In [31], we give a polynomial (both in terms of time complexity and in the number of exchanged messages) distributed algorithm to compute the process number of trees. By slightly modifying the initial parameter of the algorithm, it also allows to compute various parameters of trees as pathwidth, search number, etc.

6.4.3.2. Cops and Robber Games

The "Cops and Robber" games are turn-by-turn games where a team of cops purchase a robber in a graph. We investigated two generalizations of the game introduced by Quilliot, Nowakoski and Winkler in 1983. We provided structural characterizations of graphs where one cop is sufficient to capture a fast fugitive able to hide [27]. In particular, one of these characterizations relies on hyperbolicity of the considered graph.

A surprising application of "Cops and Robber"-like games is the problem for a web-browser to download documents in advance while an internaut is surfing on the Web. In [92], we provide a modelling of the prefetching problem in terms of Cops and Robber games. The parameter to be optimized is then the download-speed necessary for the Internaut only accesses to already downloaded webpages. This allows us to provide several complexity results and polynomial-time algorithms in some graph classes.
6.4.4. Distributed Algorithms

We investigated algorithmic problems arising in complex networks like the Internet or social networks. In this kind of networks, problems are becoming harder or impracticable because of the size and the dynamicity of these networks. One way to handle the dynamicity is to provide (distributed) fault tolerant algorithms. Studying the mobile agents paradigm seems to be a promising approach (somehow related to Cops and Robber in Section 6.4.3) to address some models of distributed computing. We considered self-stabilizing algorithms for the gathering problem, and algorithms for updating routing tables.

Besides, the more an algorithm uses local information, the easier it is to update/correct the behaviour of the algorithm. In this direction, we investigated communication problems through game theory. We also studied the power of a communication model using only localized information, i.e., we study what can be computed using this communication model.

6.4.4.1. Mobile Agents and Self-stabilization

In [64], we consider a recent model of robot-based computing which makes use of identical, memoryless mobile robots placed on nodes of anonymous graphs. The robots operate in Look-Compute-Move cycles that are performed asynchronously for each robot. In particular, we consider the case of gathering robots on an anonymous ring. We provide a new distributed approach which turns out to be very interesting as it neither completely falls into symmetry-breaking nor into symmetry-preserving techniques.

We address dynamic large scale emerging networks, e.g., mobile sensor (agent) networks. The agents are resource limited and prone to failures. They move almost unpredictably and communicate in pairs. Population Protocol model is a communication model suited for such networks. We use a recently proposed version of this model where every agent is associated with a parameter called Cover Time. Cover Times abstract the interaction characteristics of mobile agents and allow the design of fast converging protocols and the evaluation of their convergence times (this is impossible in the original model). We take advantage of this model and perform first analytical analysis of a data collection protocol used in the ZebraNet project for the wild-life tracking of zebras. We propose alternative data collection protocols for ZebraNet and we analysis their time complexities [72], [53], [54]. To achieve fault-tolerance in population protocols, we develop a generic self-stabilizing transformer [22]. This is an automatic technique to convert a protocol to its self-stabilizing version.

In addition, we address important problems of coordination and synchronization. We present and prove correct two self-stabilizing deterministic protocols solving the classical mutual exclusion problem and the group mutual exclusion one [54].

6.4.4.2. Distributed Update of Routing Tables

In [62], we propose a simple and practical distributed algorithm for computing and updating routing tables for shortest path routing. This algorithm can be combined with every distance vector shortest paths routing algorithm, and allows to reduce the total number of messages sent. We give experimental evidence that it leads to an important gain in terms of the number of messages sent at the price of a little increase in terms of space occupancy per node.

Arc-Flags is a data structure used to speed-up the shortest paths computation in a graph. In [67], we introduce a new data structure, named Road-Signs, which allows us to efficiently update the Arc-Flags of a graph in a dynamic scenario. Road-Signs can be used to compute Arc-Flags, can be efficiently updated and do not require large space consumption for many real-world graphs.

6.4.4.3. Models of Distributed Computation

Since, we need to face both locality and dynamicity issues, we are developing new techniques allowing to obtain global structural information from local (partial) views of the network. In [55], [73], we has investigated the question of determining which graph properties can or cannot be computed using only local information. We consider the following model: each of the $n$ nodes of a graph which only knows its own ID and the IDs of its neighbours is allowed to send a message of $O(\log n)$ bits to some central entity, called the referee. We then investigate whether the referee is able to decide some basic structural properties of the
network topology $G$ or not. We show that simple questions like, "does $G$ contain a square?", "does $G$ contain a triangle?" or "Is the diameter of $G$ at most 3?" cannot be solved in general [55], [73]. On the other hand, the referee can decode the messages in order to have full knowledge of $G$ when $G$ belongs to many graph classes such as planar graphs, bounded treewidth graphs and, more generally, bounded degeneracy graphs [55], [73]. Following our framework, we are able to simulate asynchronicity of the network. In particular, we have exhibited a hierarchy of problems and distributed models of computation [87].
6. New Results

6.1. Perfect simulation

We have proposed a new approach for sampling the stationary distribution of general Markov chains that only needs to consider two trajectories. We show that this new approach is particularly effective when the state space can be partitioned into pieces where envelopes can be easily computed [26]. We further show that most Markovian queuing networks have this property and we propose efficient algorithms for some of them, in particular when the rates of events range over several orders of magnitude [45]. We also provided a novel approach for efficient sampling of queues with phase type servers [37] (this paper has received the best paper award at ASMTA 2011) and Markov chains with infinite state spaces (but with a known bounding process). Perfect sampling has been used for model checking of probabilistic models in [14].

6.2. Economic models for clouds

Recently introduced spot instances in the Amazon Elastic Compute Cloud (EC2) offer low resource costs in exchange for reduced reliability; these instances can be revoked abruptly due to price and demand fluctuations. Mechanisms and tools that deal with the cost-reliability trade-offs under this scheme are of great value for users seeking to lessen their costs while maintaining high reliability. We study how mechanisms, namely, checkpointing and migration, can be used to minimize the cost and volatility of resource provisioning. Based on the real price history of EC2 spot instances, we compare several adaptive checkpointing schemes in terms of monetary costs and improvement of job completion times. We evaluate schemes that apply predictive methods for spot prices. Furthermore, we also study how work migration can improve task completion in the midst of failures while maintaining low monetary costs. Trace-based simulations show that our schemes can reduce significantly both monetary costs and task completion times of computation on spot instance [25].

6.3. Game theory and networks

We studied the traffic routing problem in networks whose users try to minimize their latencies by employing a distributed learning rule inspired by the replicator dynamics of evolutionary game theory. The stable states of these dynamics coincide with the network’s (Wardrop) equilibrium points. Despite this abundance of stable states, we find that (almost) every solution trajectory converges to an equilibrium point at an exponential rate. When network latencies fluctuate unpredictably we show that the time-average of the traffic flows of sufficiently patient users is still concentrated in a neighborhood of evolutionarily stable equilibria and we estimate the corresponding stationary distribution and convergence times [42].

We also analyzed the distributed power allocation problem in parallel multiple access channels (MAC) by studying an associated non-cooperative game which admits an exact potential function. We show that the parallel MAC game admits a unique equilibrium almost surely. Furthermore, if the network’s users employ a distributed learning scheme based on the replicator dynamics, we show that they converge to equilibrium from almost any initial condition, even though users only have local information at their disposal [41].

Using a large deviations approach we calculate the probability distribution of the mutual information of MIMO channels in the limit of large antenna numbers. We calculate the full distribution, including its tails which strongly deviate from the Gaussian behavior near the mean. This calculation provides us with a tool to obtain outage probabilities analytically at any point in the parameter space, as long as the number of antennas is not too small [20].
6.4. Mean field analysis for networks

We have studied the deterministic limits of Markov processes made of several interacting objects. While most classical results assume that the limiting dynamics has Lipschitz properties, we show that these conditions are not necessary to prove convergence to a deterministic system.

We show that under mild assumptions, the stochastic system converges to the set of solutions of a differential inclusion and we provide simple way to compute the limiting inclusion. When this differential inclusion satisfies a one-sided Lipschitz condition, there exists a unique solution of this differential inclusion and we show convergence in probability with explicit bounds.

This extends the applicability of mean field techniques to systems exhibiting threshold dynamics such as queuing systems with boundary conditions or controlled dynamics. This is illustrated by applying our results to several types of systems: fluid limits of priority queues, best response dynamics in games, push-pull queues with a large number of sources and a large number of servers and self-adapting computing systems [65].

6.5. Idleness and failure prediction in large infrastructures

We have proposed a method to discover statistical models of availability in large distributed systems and applied it to run an enlightening study of SETI@home [19]. This was also used to make long-term availability predictions for groups of desktop grid resources [21]. We have used statistically based models of heterogeneous failures in parallel systems and assessed their tolerance [39]. A similar approach was used to design correlated resource models of Internet end hosts [38], [17].

6.6. Scheduling and Game Theory

A stochastic model of failures has been used to optimize the scheduling of checkpoints on desktop grids [28].

We have also shown that non-cooperative scheduling can be considered harmful in collaborative volunteer computing environments [33].

Optimal scheduling and route selection have been investigated using a novel approach based on Lagrangian optimization. This result is inspired from flow control in multi-path networks and was used for multiple mag-of-tasks application scheduling on grids [61].

In the similar context of broker-based networks of non-observable parallel queues, we provide lower bounds on the minimum response time. We introduce the “Price of Forgetting” (PoF), the ratio between the minimum response times achieved by a probabilistic broker and a broker with memory, that is shown to be unbounded or arbitrarily close to one depending on the coefficient of variation of the service time distributions. We also put our results in the context of game theory revisiting the “Price of Anarchy” (PoA) of parallel queues: It can be decomposed into the product of the PoA achieved by a probabilistic broker (already well understood) and the PoF [10].

6.7. Validity study of flow-based network models.

Researchers in the area of distributed computing conduct many of their experiments in simulation. While packet-level simulation is often used to study network protocols, it can be too costly to simulate network communications for large-scale systems and applications. The alternative chosen in SimGrid and a few other simulation frameworks is to simulate the network based on less costly flow-level models. Surprisingly, in the literature, validation of these flow-level models is at best a mere verification for a few simple cases. Consequently, although distributed computing simulators are widely used, their ability to produce scientifically meaningful results is in doubt. In [9], [70] we focus on the validation of state-of-the-art flow-level network models of TCP communication, via comparison to packet-level simulation. While it is straightforward to show cases in which previously proposed models lead to good results, instead we systematically seek cases that lead to invalid results. Careful analysis of these cases reveal fundamental flaws and also suggest improvements. One
contribution of this work is that these improvements lead to a new model that, while far from being perfect, improves upon all previously proposed models. A more important contribution, perhaps, is provided by the pitfalls and unexpected behaviors encountered in this work, leading to a number of enlightening lessons. In particular, this work shows that model validation cannot be achieved solely by exhibiting (possibly many) “good cases.” Confidence in the quality of a model can only be strengthened through an invalidation approach that attempts to prove the model wrong.

6.8. Visualization

We have proposed a methodology for detecting resource usage anomalies in large scale distributed systems. The methodology relies on four functionalities: characterized trace collection, multi-scale data aggregation, specifically tailored user interaction techniques, and visualization techniques. We have shown the efficiency of this approach through the analysis of simulations of the volunteer computing Berkeley Open Infrastructure for Network Computing architecture (BOINC). Three scenarios have been analyzed in [48], [23]: analysis of the resource sharing mechanism, resource usage considering response time instead of throughput, and the evaluation of input file size on Berkeley Open Infrastructure for Network Computing architecture. The results show that our methodology enables to easily identify resource usage anomalies, such as unfair resource sharing, contention, moving network bottlenecks, and harmful short-term resource sharing. Triva, the resulting software, has been demonstrated at the SuperComputing conference.

We also have investigated how to use trace-based visualization to understand applications I/O performance [49] and how to visually compare two traces [70] and highlight differences.

6.9. Experimental methodology

In the scientific experimentation process, an experiment result needs to be analyzed and compared with several others, potentially obtained in different conditions. Several tools are dedicated to the control of the experiment input parameters and the experiment replay. In parallel, concurrent and distributed systems, experiment conditions are not only restricted to the input parameters, but also to the software environment in which the experiment was carried out. It is therefore essential to be able to reconstruct this type of environment. This can quickly become complex for experimenters, particularly on research platforms dedicated to scientific experimentation, where both hardware and software are in constant rapid evolution. We study the concept of the reconstructability of software environments and propose a tool for dealing with this problem in [64].

We have also started investigating the systematic use of Design of Experiments to computer studies (see [61]). Nonetheless such approach provides results that are much more trustworthy than what is generally done in the parallel and distributed computing community but it also enables to shorten the experiments cycle and to use less computing resources.

6.10. Multi-core platforms

We have used memory access traces to map threads on hierarchical multi-core platforms [13]. We have also used software transactional memory to analyze and trace applications running on multi-core architectures [30].

An approach based on machine learning was used to map threads on transactional memory applications in [31].

The impact of CPU and memory affinity on multi-core platforms was investigated in [46] using numerical scientific multi-threaded applications as a typical case study. This resulted in improvement of the performance of parallel systems using a NUMA-aware load balancer [68].

We have also carried a performance evaluation of WiNoCs for parallel workloads based on collective communications [43] as well as for Infiniband networks [40].
6.11. High performance computing

We have developed a runtime system, named SGPU 2, that enable large applications to run on clusters of hybrid nodes [44].

BigDFT is a parallel simulator of the matter at the nano scale. It uses Daubechies Wavelets for High Performance Electronic Structure Calculations [16]. This tool is shown to make efficient use of massive parallel hybrid architectures [57].

6.12. Input-Output

Atmospheric models usually demand high processing power and generate large amounts of data. As the degree of parallelism grows, the I/O operations may become the major impacting factor of their performance. In [27], we evaluate the Ocean-Land-Atmosphere Model (OLAM) on the PVFS file system in order to point the I/O characteristics of the application. We show that storing the files on PVFS has lower performance than using the local disks of the cluster nodes due to file creation and network concurrency. Additionally, we study the performance of a new version of OLAM that used MPI associated with OpenMP and show that the combined strategy presents I/O times 20 times shorter than the original MPI-only version and 9 times shorter on total execution time. Finally, a survey on I/O Characterization of several applications is given in [51].
6. New Results

6.1. Kaapi

New version of Kaapi, called X-Kaapi, has been released. The kernel is written in C for hypothetical required from embedded system. On top of the kernel, several APIs co-exist: a template based C++ library called Kaapi++; a C API; a Fortran API; and a compiler that transform a source code annotated with pragma directive to a source code with calls to the runtime library function. The compiler works with C and a subset of C++. [http://kaapi.gforge.inria.fr](http://kaapi.gforge.inria.fr)

6.2. Multi-criteria optimization

The main idea is the development of a methodology to generate a reasonable set of approximated Pareto’ solutions (closed to the best achievable solutions). Especially, we have applied this methodology to better take into account users’ criteria than the other existing methods offer. We have also studied the problem of selection of best algorithms in a portfolio. This research axis is currently enforced by the INRIA postdoc position of Joachim Lepping where we have started to include a learning process to select the best algorithm on a given instance.

6.3. Stochastic models for optimizing checkpoint protocol

After our past studied on design of origin checkpoint protocols, we have proposed a new stochastic performance model of the parallel execution in presence of failures. Thanks to this formulation, we are able to optimize several criteria (the time lost due to failure; the expected completion time) by making right decision of the date of each checkpoint. The model is general and it does not take into account the failure distribution law and accept variable checkpoint time estimation, which is important for dynamic parallelism applications.

6.4. Work stealing scheduling algorithm taking care of communication

On some applications, the amount of data transfers can be high. To minimize the amount of data transfers during the execution, Jean-Noel Quintin has developed an algorithm called WSCOM which uses the DAG structure of the application. For each steal request, the work-stealing algorithm tries to balance the load between the thief and the stolen processor. Thus, WSCOM tries to divide the work on the stolen processor into two parts with a small number of edges between the two parts. This cutting is done with a negligible overhead at each steal request. This algorithm has been implemented in a tool called DSSSكل. This tool executes the set of tasks described by a Makefile on a distributed platform. In addition, I have developed a simulator to validate algorithm performance and its behavior. We compared WSCOM and several static list-scheduling algorithms. The comparison shows that WSCOM outperforms list-scheduling algorithms, on clusters with some network congestion.

Besides, based on SIPS analysis of work stealing, Stefano Mor in his thesis compared the influence of the choice of the stolen tasks on the number of steal operations, distinguishing unsuccessful and successful steals. While standard bounds are related to unsuccessful steals, they are pessimistic with respect to the number of successful steals that define intensive data communications.
6.5. Homomorphic coding for soft error resilience

We extended our results for fault-tolerant modular computations in two directions. To improve the correction rate of Reed-Solomon codes, power-decoding techniques consist in augmenting the number of syndrom equations by raising the received word to successive powers. The correction is done by a generalization of Berlekamp-Massey algorithm acting on multiple sequences. This method is, if not equivalent, at least very close to the list-decoding proposed by Sudan in its first version, in particular, error correction rates are identical. We improve the power-decoding method by reformulation into a vector rational function reconstruction, with benefit from fast polynomial matrix arithmetic. Besides, for basic exact linear algebra computations (eg dense linear system), we designed interactive protocols between a trusted platform and a non trusted one for resilience to soft-errors.

6.6. Chimeric algorithms design

To reach provable multicriteria performance, we used the coupling of various algorithms that adapt in several contexts: recursive cascading of both sequential and parallel algorithms with work-stealing; coupling specific algorithms on heterogeneous platforms (eg CPU/GPU); interactive distributed computations; fault-tolerant computations by coupling both a trustfully platform with low computation bandwidth and and an unreliable computing platform with high bandwith. A unification work is currently developed for the design of a chimeric algorithms that is composed of the parts of multiple algorithms, interactively cascaded to achieve provable multicriteria performance.
6. New Results

6.1. Introduction

Research results are presented according to the research directions of the MYRIADS team.

6.2. Autonomous Management of Virtualized Infrastructures


6.2.1. Cloud Federations


6.2.1.1. Virtual Execution Platforms in Cloud Federations

In the context of the Contrail European project, we have defined the overall architecture of the Contrail software stack for cloud computing on top of cloud federations [51]. We have focused on the design and the implementation of a first basic prototype of the Virtual Execution Platform (VEP) component [52]. It is in charge of provisioning hardware resources from Cloud providers and to deploy and run distributed applications submitted by users under the control of a negotiated Service Level Agreement (SLA) [16]. Within VEP software, REST interface, OVF parsing, SSL security, Authorization modules are under active development and at various levels of integration. A first demo version of VEP running on top of OpenNebula IaaS cloud has been successfully demonstrated at the first annual project review.

6.2.1.2. Efficient virtual cluster migration

We continued our work on Shrinker, a system providing efficient live migration of virtual clusters on wide area networks. The design has been improved to coordinate the deduplication on the source site of the migration. Deduplication is now performed only within an individual virtual cluster, in order to reduce security issues and avoid performance impact on virtual machines of other users. We performed a comprehensive performance evaluation of Shrinker. An article presenting the design, implementation, and performance of Shrinker was published in [28].

6.2.1.3. Elastic Map/Reduce over Cloud Federations

We worked on the development of Resilin, a system to easily create execution platforms over distributed cloud resources for executing MapReduce computations. Resilin implements the Amazon Elastic MapReduce web service API and uses resources from private and community clouds. Resilin takes care of provisioning, configuring and managing cloud-based Hadoop execution platforms, potentially using multiple clouds. An initial implementation of Resilin was presented at the CCA ’11 workshop [36]. Further development was performed in the context of Ancuta Iordache’s master internship. The results of this work were published as a research report [45].

6.2.1.4. Sky Computing Experiments

We continued our collaboration with the University of Florida on sky computing experiments, which led to the publication of a book chapter [38].

6.2.2. Infrastructure as a Service Clouds

Participants: Stefania Costache, Eugen Feller, Yvon Jégou, Christine Morin, Nikos Parlavantzas, Pierre Riteau.
6.2.2.1. Large scale Energy aware self-healing IaaS

The research done in 2011 was two fold. A prototype of the previously proposed scalable, fault-tolerant and energy-aware virtual machine (VM) management framework called Snooze was implemented and evaluated on the Grid5000 testbed [41]. In 2011, we have focused on the implementation of the self-healing mechanisms and protocols, and on integrating in Snooze the system-level mechanisms (e.g. for automatically switching on/off cluster nodes) to support energy aware resource management algorithms. Our experimental results show that the fault-tolerance features of the framework do not impact application performance. Moreover, negligible cost is involved in performing distributed VM management and the system remains highly scalable with increasing amounts of resources. A nature-inspired VM placement algorithm [24] based on the Ant Colony Optimization (ACO) meta-heuristic was developed and evaluated by means of simulations.

6.2.2.2. Resource Management in Private Clouds

We focused on the design of a resource management system for private clouds that provides support for different application SLAs while maximizing the resource utilization of the infrastructure. As we also considered the need of providing users the incentives to truthfully express their valuation for the performance of their application we investigated existing economic models of allocating resources. As a result, we proposed a novel resource management architecture based on a virtual economy. In this system, independent agents monitor the application’s performance and dynamically provision virtual machines from the infrastructure under user’s budget constraints. To provision virtual machines, a proportional share auction is used, allowing a fine-grain resource sharing at a low complexity cost. This work was done as part of a collaboration with EDF R&D and was published at the VHPC 2011 workshop [22]. We have also implemented a first prototype of this proposal. In collaboration with Vydia Rajagopalan (Master student at VU Amsterdam) we have implemented the proportional-share auction scheduler and integrated it with the OpenNebula Virtual Infrastructure Manager. Then, we have extended this work with the design of agents that execute scientific applications (MPI and Bag-of-Task applications) under deadline and budget constraints. Experimental evaluations are currently performed on Grid5000.

6.2.2.3. Resilience

We initiated a collaboration with Box Leangsuksun’s group on high availability of cloud infrastructures. We carried out a preliminary study on how the HA-OSCAR environment developed at the Louisiana Tech University could be used to ensure the high availability of critical services in Nimbus IaaS clouds [30].

6.2.3. XtreemOS Grid Distributed Operating System

Participants: Amine Belhaj, Jérôme Gallard, Rémy Garrigue, Yvon Jégou, Christine Morin, Yann Radenac.

6.2.3.1. Facilitating Experiments with XtreemOS Grid System

XtreemOS Grid system that has been developed in the framework of the XtreemOS European project is now evolving as an open source software in a community driven by INRIA in the framework of the XtreemOS Easy ADT. We have provided first level support to users and maintained XtreemOS website, wiki and mailing-lists. We have updated XtreemOS documentation to reflect the evolution of XtreemOS system. We facilitated the access to XtreemOS in three different ways: maintaining an open public testbed running XtreemOS, providing ready-to-use XtreemOS virtual machines and developing tools to automatically deploy XtreemOS Grid system on the Grid’5000 large-scale experimentation platform. In 2011, we have finalized a new 3.0 version of XtreemOS system and ported it on top of the OpenSuse 11.4 Linux distribution. We performed a number of tests to validate the installation, configuration and execution of the new XtreemOS version based on openSuse Linux distribution. An incremental integration process has been set up to facilitate the integration of patches and bug fixes. We have run a number of experiments with XtreemOS 3.0 based on Mandriva Linux distribution: MPI programs, Salomé numerical simulation platform, bio-informatics applications. Yann Radenac, in the framework of the COOP project funded by ANR contributed to XtreemOS’s code by fixing bugs, cleaning the source code to improve maintainability, and adding a few minor features.
6.2.3.2. Resource Management for Dynamic Applications

In the framework of the COOP project funded by ANR, we compared the features offered by the CooRM resource manager for dynamic applications developed by Christian Perez and Cristian Klein at ENS Lyon with those provided by the XtrememOS Grid system. A plan has been set to adapt CooRM to XtrememOS system and to extend XtrememOS’s API to include a CooRM-like interface [53]. We worked on the definition of a variant of CooRM that can work in the context of XtrememOS Grid operating system.

6.3. Dynamic Adaptation of Service-based Applications

Participants: Françoise André, Djawida Dib, Erwan Daubert, Guillaume Gauvrit, André Lage, Christine Morin, Nikos Parlavantzas, Jean-Louis Pazat, Chen Wang, Mohamed Zouari.

6.3.1. Dynamic Adaptation in a Distributed Operating System

Participants: Françoise André, Djawida Dib, Christine Morin, Nikos Parlavantzas.

We have studied the feasibility to dynamic adapt the features of a distributed operating system using a framework for self-adaptation of service oriented distributed applications [46]. We have focused on the consistency protocols for replicated data in distributed shared memory systems. We have considered two strict consistency protocols, one based on invalidation and one based on broadcast on write operations. The adaptation framework selects one of these two algorithms based on the inter-node data transfer delay. We have implemented a prototype based on Kerrighed single system image operating system for clusters and the SAFDIS adaptation framework. We have integrated a broadcast based consistency protocol in Kerrighed that already implements a write invalidation consistency protocol. We have implemented the adaptation policy in the SAFDIS framework and the needed adaptation mechanisms in Kerrighed as well as a component for monitoring data transmission delays. An experimental evaluation is being carried out.

6.3.2. Adaptation for Data Management

Participants: Françoise André, Mohamed Zouari.

The usage of context-aware data management in mobile environments has been previously investigated by Françoise André in collaboration with Mayté Segarra and Jean-Marie Gilliot from Telecom Bretagne Brest (previously known as ENST Bretagne). This work focuses on data management in grid and mobile environments; an ambient assisted living application illustrates the approach. This work was realized in the context of the ALORAD project (Architecture LOgicielle pour la Réplication Adaptative de Données), financed by the Brittany council. Mayté Segarra from Telecom Bretagne Brest was co-adviser for the PhD thesis of M. Zouari [12].

6.3.3. Adaptation for Service-Oriented Architectures

Participants: Françoise André, Erwan Daubert, Guillaume Gauvrit, André Lage, Nikos Parlavantzas, Jean-Louis Pazat, Chen Wang.

Service-Oriented Computing is a paradigm that is rapidly spreading in all application domains and all environments - grids, clusters of computers, mobile and pervasive platforms. The following works take place in the context of the S-CUBE European Network of Excellence.

6.3.3.1. Services adaptation in distributed and heterogeneous systems

We are still studying service adaptation in distributed and heterogeneous systems. This work covers different aspects such as structural, behavioral and environmental adaptation, distributed decision and planification of adaptation actions, adaptive allocation of resources for services. A framework called SAFDIS for "Self Adaptation For Distributed Services" has been defined and implemented. It is built as a set of services, providing functionalities useful to build an adaptation system. The analysis phase can take reactive as well as proactive decisions. This gives the ability to either react fast or to take decisions for the long term. This implies the ability to analyze the context with a variable depth of reasoning. Our implementation of the SAFDIS analysis phase also distributes and decentralizes its analysis process to spread the computational
load and make the analysis process scalable. The planning phase seeks the set of actions (the plan) needed to adapt the system according to the strategy chosen by the analysis phase. It also schedules the selected actions to ensure a coherent and efficient execution of the adaptation. The planning topic is a well known subject in AI research works and many algorithms already exist in that field to produce efficient schedules. With our SAFDIS framework, the planning phase is able to reuse these algorithms. The resulting plan of actions can have actions that can be executed in parallel.

6.3.3.2. Quality Assurance for Distributed Services

In the context of the service-centric paradigm, we have designed and developed the Qu4DS (Quality Assurance for Distributed Services) research prototype. Qu4DS is a cloud PaaS solution which fills the gap between the conception of higher-level SaaS service providers over the resource-level PaaS layer. Qu4DS provides an automatic support for service execution management by aiming at increasing the service provider’s profit. More specific, Qu4DS dynamically acquires resources according to the customer demand, deploys service instances and implements QoS assurance mechanisms in order to prevent SLA violations. Moreover, Qu4DS has been evaluated on Grid5000 and showed to be effective on reducing service provider’s costs [27].

6.3.3.3. Self-configuration for Cloud Platforms

By definition, cloud computing offers an abstraction to manage various needs and concepts such as distributed software design, the deployment of such software on dynamic resources and the management of this kind of resources. Thus it is possible to reconfigure (adapt) according to some needs the software as well as the use of the resources. However these reconfigurations that are used on different layers may also have impacts on the others. Moreover these layers are independent and so are able to adapt themselves independently of the others. In our work, we propose to use some adaptation capabilities offered for example by the infrastructure (IaaS) that manages the resources to adapt the software (SaaS). We also propose to use planning algorithms to coordinate the adaptations between them to avoid inconsistency or inefficiency due to concurrent adaptations.

6.3.3.4. Dynamic Adaptation of Chemical services

We have proposed a QoS-aware middleware for dynamic service execution. In the context of dynamic execution, a workflow is defined by composing a set of abstract activities as place holders. Each activity is bound to a suitable partner service, which is selected at run-time from a set of functional equivalent candidates with different non-functional properties such as quality of service (QoS). The service selection process is modeled as a series of chemical reactions.

6.3.3.5. Prediction of SLA violations and dynamic adaptation in workflows

During execution, run-time QoS is determined by the dynamic execution environment and thus the expected QoS is not always ensured. In addition, infrastructure failures can make a service undeliverable. The adaptive execution reflects the capability to recompose a (part of a) workflow on-the-fly in case that global SLA violation is predicted. Most techniques for predicting global SLA violation require past experiences on executions of a business process. All historical execution instances have the same structure as well as the same bindings. These solutions do not adapt to the case of dynamic execution, where for each execution, partner services are selected and bound at run time.

In order to predict global SLA violation in the context of dynamic service execution, we proposed a 2-phase prediction technique, which is fit for generic workflow composition. The prediction method works with a high accuracy for simple workflows, but when the workflow composes complicated structures (such as loops and exclusive branches), the performance degrades. The reason is that the estimation of global SLA is based on the critical path, which is not definitely executed. To solve this problem, we propose to use data mining technique to predict workflow branches and the number of loop execution. Based on predicted branches, the prediction of global SLA violation is much more accurate. The numerical evaluation will be carried out in the near future.

6.4. A Chemical Approach for Autonomous Service Computing

Participants: Héctor Fernandez, Marko Obrovac, Thierry Priol, Cédric Tedeschi.
6.4.1. Chemistry-Inspired Workflow Management System for e-Science Applications

Participants: Héctor Fernandez, Cédric Tedeschi, Thierry Priol.

In the research track that aims at leveraging the properties of the chemical Programming models for autonomic computing, we have built a software based on the HOCL compiler (part of the HOCL-tools) that was actually deployed and experimented over the Grid5000 platform. The experiments have shown, that envisioning the execution workflow as an autonomic chemical process is actually viable in practice. Experimented with different well known workflow-based e-Science applications, the software showed a performance level comparable to current top-rated scientific workflow management systems [25].

6.4.2. Solving Workflow Patterns Through Molecular Composition

Participants: Héctor Fernandez, Cédric Tedeschi, Thierry Priol.

In the same area, but on a more conceptual point of view, we have shown how the expressive power of the chemical model can be leveraged to solve complex workflow patterns. This aspect was also integrated into the HOCL-tools and experimented over the Grid5000 platform, following two architectures with a different level of decentralization, showing the advantages and drawbacks of decentralizing the workflow execution using a chemical workflow engine [26].

6.4.3. Scalable Atomic Capture of Molecules

Participants: Marko Obrovac, Cédric Tedeschi.

Capturing the reactants involved in a reaction constitutes one of the main challenges in the execution of chemical programs. Doing it at large scale is one of the essential barriers hindering the actual execution of chemical programs at large scale. We proposed a protocol solving this issue on top of a distributed hash table (DHT). The DHT secures the scalability of the communications and provides a scalable discovery of reactants. Our protocol is triggered once reactants are found. It is made of two sub-protocols being used at different stages of the computation, according to the density of possible reactions. The protocol is validated through its proof of liveness and simulations showing that it is able to switch from one sub-protocol to the other efficiently, according to the execution’s conditions [18].
6. New Results

6.1. Distributed Programming Models

6.1.1. Multi-active Objects

Participants: L. Henrio, I. Zsolt, F. Huet.

The active object programming model is particularly adapted to easily program distributed objects: it separates objects into several activities, each manipulated by a single thread, preventing data races. However, this programming model has its limitations in terms of expressiveness – risk of deadlocks – and of efficiency on multicore machines. We proposed extends active objects with local multi-threading. We rely on declarative annotations for expressing potential concurrency between requests, allowing easy and high-level expression of concurrency. This year contribution includes

- publication of the basic principles of the new model [25]
- refinement of the proposal and adding dynamic compatibility
- operational semantic for multi-active objects
- extensive experiments

6.2. Component-oriented Distributed and Large-Scale Programming

6.2.1. Behavioural models for Distributed Components

Participants: E. Madelaine, R. Halalai, A. Savu, M. Alexe, L. Henrio.

In the past, we defined the behavioural semantics of active objects and components, in [3]; we extended last year this work to take group communications. On the practical side, this work contributes to the Vercors platform; the overall picture being to provide tools to the programmer for defining his application, including its behavioural specification. Then some generic properties like absence of deadlocks, but also application specific properties can be validated on the composed model using an existing model-checker. We mainly use CADP model-checker, that also supports distributed generation of state-space. This year our main achievements are the following:

- We provided model for one-to-many component communication
- We provided a model for Byzantine failures, specified a component application supporting some Byzantine faults, and proved its correctness;
- We conducted heavy experiments on distributed model-checking in this context;
- We worked on the formal specification of the behavioural model generation for component systems.

Most of those results were published in [22] and [35].

6.2.2. Enacting large-scale service orchestration using a component-based approach

Participants: F. Baude, V. Legrand.
The distribution of business processes encompasses the inclusion of external service providers in the overall process as well as the usage of external infrastructures like clouds. Both of these approaches lead to decentralization and outsourcing of a part of the global workflow, resulting in a complexified management of the global orchestration. As a matter of fact, the overall data are decentralized among different domains and must, most of the time, be gathered manually. To this extent, we continue our work on agile and distributed orchestration, showing that the framework we develop eases multidomain orchestration management. Our approach extracts, gathers and digests data from the decentralized processes in order to provide an unified and global view of a distributed orchestration. This year we focussed in particular on:

- *The specification of the execution framework* extending the SCA specification by adding temporal dependencies
- *The definition of a use-case* allowing the provisioning of a large set of OSGi gateways.

This work resulted in the PhD thesis of Virginie Legrand [12].

### 6.2.3. Autonomic Monitoring and Management of Components

**Participants:** F. Baude, C. Ruz, B. Sauvan, R. Dib.

We have completed the design of a framework for autonomic monitoring and management of component-based applications. We have provided an implementation using GCM/ProActive taking advantage of the possibility of adding components in the membrane, and we have tested it in simple applications. Our implementation allows the designer to describe in a separate way each phase of the MAPE autonomic control loop (Monitoring, Analysis, Planning, and Execution), and to plug them or unplug them dynamically [16].

- We presented the general description of the framework and its capability to support autonomic behaviour in [30]. This work takes advantage of the componentized membrane of GCM/ProActive, and of the PAGCMSScript reconfiguration extensions made in our team.
- We used our proposition to design an integrated framework to cover the life-cycle of a service application from business and design level, to deployment and execution concerns in a Cloud environment, in a work done in conjunction with Adrian Mos and Alain Boulze formerly leading the INRIA ADT Galaxy from INRIA Rhône-Alpes. This work was presented in the SoEA4EE workshop [29].
- We experimented with the use of our autonomic framework to integrate autonomic behaviour into skeletons. This work was taken during the engineering internship of Rima Dib, and included the collaboration with Marco Danelutto from Università di Pisa.

### 6.3. Middleware for Grid and Cloud computing

#### 6.3.1. RDF Data Storage and Retrieval In P2P Systems

**Participants:** I. Filali, F. Huet, F. Baude, F. Bongiovanni, L. Pellegrino, B. Sauvan, I. Alshabani, A. Bourdin.

We have proposed in the context of the SOA4ALL FP7-IP project [8.3.1.1] the design and the implementation of a hierarchical Semantic Space infrastructure based on Structured Overlay Networks (SONS) [46],[10]. It aims at the storage and the retrieval of the semantic description of services at the Web scale [47]. This infrastructure combines the strengths of both P2P paradigm at the architectural level and the Resource Description Framework (RDF) data model at the knowledge representation level. As it is designed, the proposed infrastructure enables the processing of simple and complex queries. This year, the following achievements have been realised.

- A thorough survey of the existing works that have adapted the combination of RDF data model and the P2P communication model to build distributed infrastructures for RDF data storage and retrieval has been performed. This effort was published in a journal [34]. A previous but more complete version of this work can be found in a research report [45] and was used extensively in [39], [38],[36].
• We provided and presented in [23] an implementation of a three dimensional CAN overlay network for storing and retrieving RDF triples. At the implementation level, a modular and flexible architecture for the Semantic Space infrastructure has been proposed. The implementation relies on the ProActive Grid middleware and provides a clear separation between its sub-components (overlay, storage, query engine, etc.). The modularity of the architecture is combined with the decentralized aspect of the infrastructure enabling the RDF data storage and retrieval at large scale. The evaluation of the infrastructure through micro-benchmarks experiments on clusters and grids shows the impact of the architecture and data distribution on the performance of the storage and processing mechanisms.

In the context of the FP7 Strep PLAY (8.3.1.2) and French ANR SocEDA (8.2.2) research projects, we have extended the aforementioned work with a content-based Publish/Subscribe abstraction in order to support asynchronous queries for RDF-based events in large scale settings. In order to support these queries efficiently, we worked on an efficient broadcast primitive on top of CAN which we formalized and implemented (see section 6.3.2). We are also working towards a generalization of this broadcast algorithm to a multicast one, and contribute intensively to the general integration effort for offering such innovative semantic described event marketplace platform at cloud scale [41].

6.3.2. An algorithm for efficient broadcast over CAN-like P2P networks

Participants: L. Henrio, F. Bongiovani, A. Craciun.

The nature of some large-scale applications such as content delivery systems or publish/subscribe systems, built on top of structured overlay networks, demands application-level dissemination primitives which do not overwhelm the overlay and which are also reliable. Building such communication primitives in a reliable manner on top of such networks would increase the confidence regarding their behavior prior to deploying them in real settings. In order to come up with real efficient primitives, we take advantage of the underlying geometric topology of the overlay network and we also model the way peers communicate with one another. For this our objective is to design and prove an efficient and reliable broadcast algorithm for CAN-like P2P networks. To this aim, this year we:

• Formalised in Isabelle/HOL a CAN-like P2P system, devised formalised tools to reason on CAN topologies, and on communication protocols on top of CANs. We proved first completeness and correctness properties on some classes of broadcast algorithms.

• Designed an efficient broadcast algorithm on top of CAN and implemented it.

Preliminary results were presented at CFSE and published as a research report [37]; another publication is under way.

6.3.3. Matlab/Scilab parallel programming

Participant: F. Viale.

Matlab & Scilab, with millions of users around the world, are industry standards for numerical computing. They both lack a powerful and modern parallel computing framework to meet the industry’s growing demand in terms of parallel processing. This activity is intended to integrate into both softwares a toolbox for parallel processing, based on ProActive.

• This year, we implemented a ProActive Scilab toolbox with the same functionalities as the ProActive Matlab toolbox we built last year.

• We added in the Matlab toolbox a disconnected mode to allow closing the Matlab session while uncomplete Matlab jobs are still running on the scheduler side, and retrieving the job results at the next connection.

• We reorganized both Matlab and Scilab toolboxes with a cleaner and more intuitive API, a stronger and stabler implementation and an extensive documentation. We designed as well unit-tests to make the toolboxes fully usable for production standards.

• The Scilab toolbox is now deployed on PACAGrid cluster and used extensively by other partners of the OMD2 ANR.
6.3.4. Network Aware Cloud Computing

Participants: S. Malik, F. Huet.

We have proposed a cloud scheduler module named Network Awareness Module (NAM), which helps the scheduler to take more efficient scheduling decisions on the basis of resource features, such as network latency, reliability, environment compatibility and monetary cost issues.

- Currently we are working on Reliability Assessment based Scheduling on Cloud Infrastructure. We are building a model, which enables a scheduler to schedule the tasks on cloud infrastructure, on the basis of adaptive reliability of nodes (virtual machine). The core of this model is a reliability assessment algorithm, which computes the reliability for individual resources and for the group of resources.

- We have proposed, designed and implemented an algorithm for the grouping of nodes on the basis of inter-node latencies. This algorithm can do the dynamic grouping and work with the incomplete latency information available. It groups the nodes on the basis of node latency instead of neighbor count. It produces mutually exclusive groups and can perform group reconfiguration.

- We have designed a model of the Virtual Cloud [27]. The concept of Virtual cloud revolves around the concept, “Rent Out the Rented Resources”. In this model, cloud vendors offer low cost cloud services by acquiring underutilized resources from some big third-party enterprise. The cloud vendor then further rents out those resources/services to the cloud users. The upfront and administrative costs for the Virtual cloud vendor are lower, and the cloud users access services at a cheaper rate.

- We have proposed a fault tolerance model for real time cloud computing [27]. In this model, the system tolerates the faults and makes the decision on the basis of reliability of the virtual machines. The reliability of the virtual machines is adaptive, which changes after every computing cycle. A metric model is given for the reliability assessment. The system provides both the forward and backward recovery mechanisms.
6. New Results

6.1. Leveraging Software Architectures to Guide and Verify the Development of Sense/Compute/Control Applications

A software architecture describes the structure of a computing system by specifying software components and their interactions. Mapping a software architecture to an implementation is a well known challenge. A key element of this mapping is the architecture’s description of the data and control-flow interactions between components. The characterization of these interactions can be rather abstract or very concrete, providing more or less implementation guidance, programming support, and static verification.

In this work, we have introduced a notion of behavioral contract that expresses the set of allowed interactions between components, describing both data and control-flow constraints [15]. This declaration is part of the architecture description, allows generation of extensive programming support, and enables various verifications. We have instantiated our approach in an architecture description language for the domain of Sense/Compute/Control (SCC) applications, and described associated compilation and verification strategies.

The main contributions of this work are the following:

- We have introduced a language for behavioral contracts dedicated to SCC applications.
- We have shown that behavioral contracts can effectively guide the implementation of SCC applications by enabling the generation of highly customized programming frameworks using a dedicated compiler. This approach ensures the conformance between the architecture and the implementation, while facilitating software evolution.
- We have shown that such descriptions are precise enough to verify safety properties such as information flow reachability or behavioral invariants.
- Based on an implementation of behavioral contracts in a design language targeting SCC applications, we have assessed the benefit of behavioral contracts at a conceptual level and in terms of metrics on the resulting code.

6.2. A Step-wise Approach for Integrating QoS throughout Software Development

Non-functional requirements are used to express the quality to be expected from a system. For real-time systems such as avionics, it is critical to guarantee this quality, in particular time-related performance properties. In this domain, deterministic QoS is generally ensured at the execution platform level (e.g., operating systems, distributed systems technologies, hardware specificities), independently of a particular application. When addressing the QoS requirements of a given application, these platform-specific guarantees are not sufficient.

In this work, we have proposed a step-wise QoS approach integrated through all development phases and development artifacts [17]. This approach is dedicated to control-loop systems. Control-loop systems are systems that sense the external environment, compute data, and eventually control the environment accordingly. This kind of systems can be found in a range of domains, including avionics, robotics, and pervasive computing. For example, in the avionics domain, a flight management application is a control-loop system that (1) senses the environment for location and other navigation information, (2) computes the trajectory and (3) modifies the wings configuration accordingly.
The main contributions of this work are the following:

- We have developed a step-wise approach that systematically processes QoS requirements throughout software development. This integrated approach is dedicated to control-loop systems, allowing to rely on a particular architectural pattern and thus enhancing the design and programming support level for non-functional aspects. For now, we focus on time-related performance but the approach could be generalized to other non-functional properties (e.g., CPU or memory consumption).

- Our approach has been integrated into DIASUITE, a tool-based development methodology dedicated to control-loop systems. DIASUITE is based on a dedicated design language that we have enriched with time-related performance properties. This non-functional extension has been used to offer verification and programming support at each development stage.

- Our approach has been applied to the development of avionics applications such as a flight management system and a collision avoidance system. These experiments have demonstrated that our step-wise approach can effectively guide the avionics certification process.

### 6.3. Architecturing Conflict Handling of Pervasive Computing Resources

The rapid development of new devices (further resources) and development tools being opened to third-parties have paved the way to an increasing number of applications being deployed in pervasive computing environments. These applications anarchically access resources. In this situation, it is very common for a resource to be accessed by multiple applications, potentially leading to conflicts. For example, in a building management system, a security application that grants access inside the building, can conflict with an application dealing with emergency situations like fires, preventing the building to be evacuated.

Managing conflicts consists of three main parts, detection, resolution and prevention. These parts crosscut the development cycle of applications and pervasive computing systems. In this work, we have proposed a conflict management process that cleanly separates conflict management tasks by providing a design method and supporting tools [18].

The main contributions of this work are the following:

- We have identified requirements at different stages during the development cycle that are necessary to detect, resolve, and prevent conflicts. We have assigned duties and responsibilities to existing roles, that are carried out during the conflict management process without interfering with the standard application development.

- We have extended a domain-specific design language to declare conflict resolution at an architectural level. During the conflict management process conditions are specified and prioritized. Afterwards conflicting applications (inter application) or modules (intra application) are linked to these conditions.

- The declared information is used to generate code dedicated to conflict handling. On the one hand, a compiler generates a dedicated framework that guides the implementation of the conflict handling logic at application and system level. On the other hand, it generates code that orchestrates resource accesses and prevents conflicts.
6. New Results

6.1. Towards Data-Centric Networking


- Disruption Tolerant Networking

We designed an efficient message delivery framework, called MeDeHa, which enables communication in an internet connecting heterogeneous networks that is prone to disruptions in connectivity[24]. MeDeHa is complementary to the IRTF’s Bundle Architecture: besides its ability to store messages for unavailable destinations, MeDeHa can bridge the connectivity gap between infrastructure-based and multi-hop infrastructure-less networks. It benefits from network heterogeneity (e.g., nodes supporting more than one network and nodes having diverse resources) to improve message delivery. For example, in IEEE 802.11 networks, participating nodes may use both infrastructure- and ad-hoc modes to deliver data to otherwise unavailable destinations. It also employs opportunistic routing to support nodes with episodic connectivity. One of MeDeHa’s key features is that any MeDeHa node can relay data to any destination and can act as a gateway to make two networks inter-operate or to connect to the backbone network. The network is able to store data destined to temporarily unavailable nodes till the time of their expiry. This time period depends upon current storage availability as well as quality-of-service needs (e.g., delivery delay bounds) imposed by the application. We showcase MeDeHa’s ability to operate in environments consisting of a diverse set of interconnected networks and evaluate its performance through extensive simulations using a variety of scenarios with realistic synthetic and real mobility traces. Our results show significant improvement in average delivery ratio and a significant decrease in average delivery delay in the face of episodic connectivity. We also demonstrate that MeDeHa supports different levels of quality-of-service through traffic differentiation and message prioritization.

Then, we have extended the MeDeHa framework to support multihop mobile ad-hoc networks (or MANETs). Integrating MANETs to infrastructure-based networks (wired or wireless) allows network coverage to be extended to regions where infrastructure deployment is sparse or nonexistent as well as a way to cope with intermittent connectivity. Indeed, to date there are no comprehensive solutions that integrate MANETs to infrastructure-based networks. We have proposed a message delivery framework that is able to bridge together infrastructure-based and infrastructure-less networks. Through extensive simulations, we have demonstrated the benefits of the extended MeDeHa architecture especially in terms of the extended coverage it provides as well as its ability to cope with arbitrarily long-lived connectivity disruptions. Another important contribution of this work is to deploy and evaluate our message delivery framework on a real network testbed as well as conduct experiments in "hybrid" scenarios running partly on simulation and partly on real nodes[32].

Finally, we have proposed a naming scheme for heterogeneous networks composed of infrastructure-based and infrastructure-less networks where nodes may be subject to intermittent connectivity. The proposed scheme, called Henna, aims at decoupling object identification from location and is designed to operate with status-quo Internet routing. We evaluated the proposed naming scheme using the ns-3 network simulator and demonstrated that nodes were able to receive messages in both infrastructure-based and infrastructure-less networks despite frequent disconnections and changing location identifiers (i.e., IP address), while visiting different networks[31].

Another important contribution of this work is to deploy and evaluate our message delivery framework on a real network testbed as well as conduct experiments in “hybrid” scenarios running partly on simulation and partly on real nodes. This was demonstrated at the ACM Sigcomm conference in Toronto on August 2011[74].
These different works are the result of collaborations with Katia Obraczka and Marc Mendonca from University of California Santa Cruz (UCSC) in the context of the COMMUNITY Associated Team, see URL [http://inrg.cse.ucsc.edu/community/](http://inrg.cse.ucsc.edu/community/).

Another activity in the same domain relates to efficient scheduling and drop policies in DTNs. We remind that Delay Tolerant Networks are wireless networks where disconnections may occur frequently. In order to achieve data delivery in such challenging environments, researchers have proposed the use of store-carry-and-forward protocols: there, a node may store a message in its buffer and carry it along for long periods of time, until an appropriate forwarding opportunity arises. Multiple message replicas are often propagated to increase delivery probability. This combination of long-term storage and replication imposes a high storage and bandwidth overhead. Thus, efficient scheduling and drop policies are necessary to:

1. decide on the order by which messages should be replicated when contact durations are limited, and
2. which messages should be discarded when nodes' buffers operate close to their capacity.

We worked on an optimal scheduling and drop policy that can optimize different performance metrics, such as the average delivery rate and the average delivery delay. First, we derived an optimal policy using global knowledge about the network, then we introduced a distributed algorithm that collects statistics about network history and uses appropriate estimators for the global knowledge required by the optimal policy, in practice. At the end, we are able to associate to each message inside the network a utility value that can be calculated locally, and that allows to compare it to other messages upon scheduling and buffer congestion. Our solution called HBSD (History Based Scheduling and Drop) integrates methods to reduce the overhead of the history-collection plane and to adapt to network conditions. The first version of HBSD and the theory behind have been published in 2008. A recent paper [27] provides an extension to a heterogenous mobility scenario in addition to refinements to the history collection algorithm. An implementation is proposed for the DTN2 architecture as an external router and experiments have been carried out by both real trace driven simulations and experiments over the SCORPION testbed at the University of California Santa Cruz. We refer to the web page of HBSD for more details [http://planete.inria.fr/HBSD_DTN2/](http://planete.inria.fr/HBSD_DTN2/).

HBSD in its current version is for point-to-point communications. Another interesting schema is to consider one-to-many communications, where requesters for content express their interests to the network, which looks for the content on their behalf and delivers it back to them. We are working on this extension within a new framework called MobiTrade, which provides a utility driven trading system for efficient content dissemination on top of a disruption tolerant network. While simple tit-for-tat (TFT) mechanisms can force nodes to give one to get one, dealing with the inherent tendency of peers to take much but give back little, they can quickly lead to deadlocks when some (or most) of interesting content must be somehow fetched across the network. To resolve this, MobiTrade proposes a trading mechanism that allows a node (merchant) to buy, store, and carry content for other nodes (its clients) so that it can later trade it for content it is personally interested in. To exploit this extra degree of freedom, MobiTrade nodes continuously profile the type of content requested and the collaboration level of encountered devices. An appropriate utility function is then used to collect an optimal inventory that maximizes the expected value of stored content for future encounters, matched to the observed mobility patterns, interest patterns, and collaboration levels of encountered nodes. Using ns-3 simulations based on synthetic and real mobility traces, we show that MobiTrade achieves up to 2 times higher query success rates compared to other content dissemination schemes. Furthermore, we show that MobiTrade successfully isolates selfish devices. For further details on MobiTrade, we refer to [41] and to the web page of the project [http://planete.inria.fr/MobiTrade/](http://planete.inria.fr/MobiTrade/).
Content distribution prevails in today's Internet and content-oriented networking proposes to access data directly by their content name instead of their location, changing so the way routing must be conceived. We worked a routing mechanism that faces the new challenge of interconnecting content-oriented networks. Our solution relies on a naming resolution infrastructure that provides the binding between the content name and the content networks that can provide it. Content-oriented messages are sent encapsulated in IP packets between the content-oriented networks. In order to allow scalability and policy management, as well as traffic popularity independence, binding requests are always transmitted to the content owner. The content owner can then dynamically learn the caches in the network and adapt its binding to leverage the cache use.

The work done so far is related to routing between content-oriented networks. We are starting an activity on how to provide routing inside a content network. To that aim, we are investigating on the one hand probabilistic routing and, on the other hand, deterministic routing and possible extension to Bellman-Ford techniques. In addition to routing, we are investigating the problem of congestion in content-oriented networks. Indeed, in this new paradigm, congestion must be controlled on a per-hop basis, as opposed to the end-to-end congestion control that prevails today. We think that we can combine routing and congestion control to optimize resource consumption. Finally, we are studying the implications of using CCN from an economical perspective. This activity was started in October 2011 by Damien Saucez.

**Application-Level Forward Error Correction Codes (AL-FEC) and their Applications to Broadcast/Multicast Systems**

With the advent of broadcast/multicast systems (e.g., DVB-H/SH), large scale content broadcasting is becoming a key technology. This type of data distribution scheme largely relies on the use of Application Level Forward Error Correction codes (AL-FEC), not only to recover from erasures but also to improve the content broadcasting scheme itself (e.g., with FLUTE/ALC).

Our recent activities, in the context of the PhD of F. Mattoussi, included the design, analysis and improvement of GLDPC-Staircase codes, a "Generalized" extension to LDPC-Staircase codes. We have shown in particular that these codes: (1) offer small rate capabilities, i.e. can produce a large number of repair symbols 'on-the-fly', when needed; (2) feature high erasure recovery capabilities, close to that of ideal codes. Therefore they offer a nice opportunity to extend the field of application of existing LDPC-Staircase codes, while keeping backward compatibility (LDPC-Staircase "codewords" can be decoded with a GPLDPC-Staircase codec).

Our LDPC-Staircase codes, that offer a good balance in terms of performance, have been included as the primary AL-FEC solution for ISDB-Tmm (Integrated Services Digital Broadcasting, Terrestrial Mobile Multimedia), a Japanese standard for digital television (DTV) and digital radio. This is the first adoption of these codes in an international standard.

This success has been made possible, on the one hand, by major efforts in terms of standardization within IETF: the RFC 5170 (2008) defines the codes and their use in FLUTE/ALC, a protocol stack for massively scalable and reliable content delivery services, an active Internet-Draft published last year describes the use of these AL-FEC codes in FECFRAME, a framework for robust real-time streaming applications, and a recent Internet-Draft [66] defines the GOE (Generalized Object Encoding) extension of LDPC-Staircase codes for UEP (Unequal Erasure Protection) and file bundle protection services.

This success has also been made possible, on the other hand, by our efforts in terms of design and evaluation of two efficient software codecs of LDPC-Staircase codes. One of them is distributed in open-source, as part of our OpenFEC project (http://openfec.org), a unique initiative that aims at promoting open and free AL-FEC solutions. The second one, a highly optimized version with improved decoding speed and reduced memory requirements, will be commercialized in 2012.
through an industrial partner. This codec proves that LDPC-Staircase codes can offer erasure recovery performances close to ideal codes in many circumstances while keeping decoding speeds over 1Gbps.

The fact that LDPC-Staircase codes have been preferred to a major AL-FEC competitor for the ISDB-Tmm standard, is the recognition of their intrinsic qualities and of an appropriate balance between several technical and non technical criteria.

- **Unequal Erasure Protection (UEP) and File bundle protection through the GOE (Generalized Object Encoding) scheme**

  This activity has been initiated with the PostDoc work of Rodrigue IMAD. It focuses on Unequal Erasure Protection capabilities (UEP) (when a subset of an object has more importance than the remaining) and file bundle protection capabilities (e.g. when one want to globally protect a large set of small objects).

  After an in-depth understanding of the well-known PET (Priority Encoding Technique) scheme, and the UOD for RaptorQ (Universal Object Delivery) initiative of Qualcomm, which is a realization of the PET approach, we have designed the GOE FEC Scheme (Generalized Object Encoding) alternative. The idea, simple, is to decouple the FEC protection from the natural object boundaries, and to apply an independent FEC encoding to each “generalized object”. The main difficulty is to find an appropriate signaling solution to synchronize the sender and receiver on the exact way FEC encoding is applied. In [65] we show this is feasible, while keeping a backward compatibility with receivers that do not support GOE FEC schemes. Two well known AL-FEC schemes have also been extended to support this new approach, with very minimal modifications, namely Reed-Solomon and LDPC-Staircase codes [66], [65].

  During this work, we compared the GOE and UOD/PET schemes, both from an analytical point of view (we use an N-truncated negative binomial distribution to that purpose) and from an experimental, simulation based, point of view [67]. We have shown that the GOE approach, by the flexibility it offers, its simplicity, its backward compatibility and its good recovery capabilities (under finite of infinite length conditions), outperforms UOD/PET for practical realizations of UEP/file bundle protection systems. See also [http://www.ietf.org/proceedings/81/slides/rmt-2.pdf](http://www.ietf.org/proceedings/81/slides/rmt-2.pdf).

- **Application-Level Forward Error Correction Codes (AL-FEC) and their Applications to Robust Streaming Systems**

  AL-FEC codes are known to be useful to protect time-constrained flows. The goal of the IETF FECFRAME working group is to design a generic framework to enable various kinds of AL-FEC schemes to be integrated within RTP/UDP (or similar) data flows. Our contributions in the IETF context are three fold. First of all, we have contributed to the design and standardization of the FECFRAME framework, now published as a Standards Track RFC [68].

  Secondly, we have proposed the use of Reed-Solomon codes (with and without RTP encapsulation of repair packets) and LDPC-Staircase codes within the FECFRAME framework: [59][60][61].

  Finally, in parallel, we have started an implementation of the FECFRAME framework in order to gain an in-depth understanding of the system. Previous results showed the benefits of LDPC-Staircase codes when dealing with high bit-rate real-time flows.

  A second type of activity, in the context of robust streaming systems, consisted in the analysis of the Tetrys approach, in [29]. Tetrys is a promising technique that features high reliability while being independent from RTT, and performs better than traditional block FEC techniques in a wide range of operational conditions.
• A new File Delivery Application for Broadcast/Multicast Systems

FLUTE has long been the one and only official file delivery application on top of the ALC reliable multicast transport protocol. However FLUTE has several limitations (essentially because the object meta-data are transmitted independently of the objects themselves, in spite of their inter-dependency), features an intrinsic complexity, and is only available for ALC.

Therefore, we started the design of FCAST, a simple, lightweight file transfer application, that works both on top of both ALC and NORM. This work is carried out as part of the IETF RMT Working Group, in collaboration with B. Adamson (NRL). This document has passed WG Last Call and is currently considered by IESG[56], [57], [58].

• Security of the Broadcast/Multicast Systems

We believe that sooner or later, broadcasting systems will require security services. This is all the more true as heterogeneous broadcasting technologies will be used, for instance hybrid satellite-based and terrestrial networks, some of them being by nature open, as wireless networks (e.g., wimax, wifi). Therefore, one of the key security services is the authentication of the packet origin, and the packet integrity check. A key point is the ability for the terminal to perform these checks easily (the terminal often has limited processing and energy capabilities), while being tolerant to packet losses.

The TESLA (Timed Efficient Stream Loss-tolerant Authentication) scheme fulfills these requirements. We are therefore standardizing the use of TESLA in the context of the ALC and NORM reliable multicast transport protocols, within the IETF MSEC working group. This document has been published as RFC 5776.

In parallel, we have specified the use of simple authentication and integrity schemes (i.e., group MAC and digital signatures) in the context of the ALC and NORM protocols in [62], [63], [64]. This activity is also carried out within the IETF RMT working group.

• High Performance Security Gateways for High Assurance Environments

This work focuses on very high performance security gateways, compatible with 10Gbps or higher IPsec tunneling throughput, while offering a high assurance thanks in particular to a clear red/black flow separation. In this context we have studied last year the feasibility of high-bandwidth, secure communications on generic machines equipped with the latest CPUs and General-Purpose Graphical Processing Units (GPGPU).

The work carried out in 2011 has consisted in setting up and evaluating the high performance platform. This platform heavily relies on the Click modular TCP/IP protocol stack implementation, which turned out to be a key enabler both in terms of specialization of the stack and parallel processing. Our activities also consisted in analyzing the PMTU discovery aspect since it is a critical factor in achieving high bandwidths. To that goal we have designed a new approach for qualifying ICMP blackholes in the Internet, since PMTUD heavily relies on ICMP.

6.2. Network Security and Privacy

Participants:  Sana Ben Hamida, Claude Castelluccia, Walid Dabbous, Mohamed Ali Kaafar, Arnaud Legout, Stevens Le Blond, Daniele Perito.

• Online users tracking and profiling techniques
Usernames are ubiquitously used for identification and authentication purposes on web services and the Internet at large, ranging from the local-part of email addresses to identifiers in social networks. Usernames are generally alphanumerical strings chosen by the users and, by design, are unique within the scope of a single organization or web service. In this work, we investigate the feasibility of using usernames to trace or link multiple profiles across services that belong to the same individual. The intuition is that the probability that two usernames refer to the same physical person strongly depends on the entropy of the username string itself. Our experiments, based on usernames gathered from real web services, show that a significant portion of the users’ profiles can be linked using their usernames. In collecting the data needed for our study, we also show that users tend to choose a small number of related usernames and use them across many services. This work is the first to consider usernames as a source of information when profiling users on the Internet. It has been published in PETS 2011 [47], one of the most prestigious conference in the area of Computer Privacy, and has been awarded the Andreas Pfitzmann award for the best contribution.

- **Online Privacy measurements and threats identification in online social networks**

  In this work, we show how these seemingly harmless interests (e.g., Music Interests) can leak privacy-sensitive information about users. In particular, we infer their undisclosed (private) attributes using the public attributes of other users sharing similar interests. In order to compare user-defined interest names, we extract their semantics using an ontologized version of Wikipedia and measure their similarity by applying a statistical learning method. Besides self-declared interests in Music, our technique does not rely on any further information about users such as friends relationship or group belongings. Our experiments, based on more than 104K public profiles collected from Facebook and more than 2000 private profiles provided by volunteers, show that our inference technique efficiently predicts attributes that are very often hidden by users. To the best of our knowledge, this is the first time that user interests are used for profiling, and more generally, semantics-driven inference of private data is addressed. This work has been published in the prestigious Network & Distributed System Security Symposium (NDSS) 2012 [37].

- **Privacy Enhancing Technologies**

  The increasing amount of personal and sensitive information disseminated over the Internet prompts commensurately growing privacy concerns. Digital data often lingers indefinitely and users lose its control. This motivates the desire to restrict content availability to an expiration time set by the data owner. This work presents and formalizes the notion of Ephemeral Publishing (EphPub), to prevent the access to expired content. We propose an efficient and robust protocol that builds on the Domain Name System (DNS) and its caching mechanism. With EphPub, sensitive content is published encrypted and the key material is distributed, in a steganographic manner, to randomly selected and independent resolvers. The availability of content is then limited by the evanescence of DNS cache entries. The EphPub protocol is transparent to existing applications, and does not rely on trusted hardware, centralized servers, or user proactive actions. We analyze its robustness and show that it incurs a negligible overhead on the DNS infrastructure. We also perform a large-scale study of the caching behavior of 900K open DNS resolvers. Finally, we propose an Android application, Firefox and Thunderbird extensions that provide ephemeral publishing capabilities, as well as a command-line tool to create ephemeral files. This work has been published in ICNP 2011 [36].

- **Differentially private smart metering**
Several countries throughout the world are planning to deploy smart meters in households in the very near future. The main motivation, for governments and electricity suppliers, is to be able to match consumption with generation. Traditional electrical meters only measure total consumption on a given period of time (i.e., one month or one year). As such, they do not provide accurate information of when the energy was consumed. Smart meters, instead, monitor and report consumption in intervals of few minutes. They allow the utility provider to monitor, almost in realtime, consumption and possibly adjust generation and prices according to the demand. Although smart metering might help improving energy management, it creates many new privacy problems. Smart meters provide very accurate consumption data to electricity providers. As the interval of data collected by smart meters decreases, the ability to disaggregate low-resolution data increases.

We developed a new privacy-preserving smart metering system. Our scheme is private under the differential privacy model and therefore provides strong and provable guarantees. With our scheme, an (electricity) supplier can periodically collect data from smart meters and derive aggregated statistics while learning only limited information about the activities of individual households. For example, a supplier cannot tell from a user’s trace when he watched TV or turned on heating. Our scheme is simple, efficient and practical. Processing cost is very limited: smart meters only have to add noise to their data and encrypt the results with an efficient stream cipher.

This work was presented at IH’11 (the Information Hiding Conference, 2011) [34].

- **Protecting against Physical Resource Monitoring**

This work considers the problem of resource monitoring. We consider the scenario where an adversary is physically monitoring on the resource access, such as the electricity line or gas pipeline, of a user in order to learn private information about his victim. Recent works, in the context of smart metering, have shown that a motivated adversary can basically profile a user or a family solely from his electricity traces. However, these works only consider the case of a semi-honest-but-non-intrusive adversary that is only trying to learn information from the consumption reports sent by the user. This work, instead, considers the much more challenging case of a intrusive semi-honest adversary, i.e. a semi-honest adversary that is in addition physically monitoring the resource by modifying the distribution network. We aim at answering to the following question: is it possible to design a resource distribution scheme that prevents resource monitoring and provides strong protection? We propose and analyze several possible solutions. The proposed solutions provide different privacy bounds and performance results. This work was presented at WPES’11 (ACM Workshop on Privacy in the Electronic Society) [35].

- **The Failure of Noise-Based Non-Continuous Audio Captchas**

CAPTCHAs, which are automated tests intended to distinguish humans from programs, are used on many web sites to prevent bot-based account creation and spam. To avoid imposing undue user friction, CAPTCHAs must be easy for humans and difficult for machines. However, the scientific basis for successful CAPTCHA design is still emerging. This project examines the widely used class of audio CAPTCHAs based on distorting non-continuous speech with certain classes of noise and demonstrates that virtually all current schemes, including ones from Microsoft, Yahoo, and eBay, are easily broken. More generally, we describe a set of fundamental techniques, packaged together in our Decaptcha system, that effectively defeat a wide class of audio CAPTCHAs based on non-continuous speech. Decaptcha’s performance on actual observed and synthetic CAPTCHAs indicates that such speech CAPTCHAs are inherently weak and, because of the importance of audio for various classes of users, alternative audio CAPTCHAs must be developed.

This work was presented at IEEE Security and Privacy 2011 [33].
• **BlueBear: Privacy in P2P systems**

We have started a new project called bluebear on privacy threats in the Internet. Indeed, the Internet has never been designed with privacy in mind. For instance, the Internet is based on the IP protocol that exposes the IP address of a user to any other users it is communicating with. However, we believe that current users of the Internet do not realize how much they compromise their privacy by using the Internet. Indeed, the common wisdom is that there are so many users in the Internet that it is not feasible for an attacker, apart may be for national agencies, to globally compromise the privacy of a large fraction of users. Therefore, finding a specific user is like looking for a needle in a haystack. The goal of the bluebear project is to raise attention on privacy issues when using the Internet. In particular, we want to show that without any dedicated infrastructure, it is possible to globally compromise the privacy of Internet users. BitTorrent is arguably the most efficient peer-to-peer protocol for content replication. However, BitTorrent has not been designed with privacy in mind and its popularity could threaten the privacy of millions of users.

In a first study we showed that it is possible to continuously monitor from a single machine most BitTorrent users and to identify the content providers (also called initial seeds). We performed a very large monitoring operation continuously “spying” on most BitTorrent users of the Internet from a single machine and for a long period of time. During a period of 103 days, we collected 148 million IP addresses downloading 2 billion copies of contents. We then identified the IP address of the content providers for 70% of the BitTorrent contents we spied on. We showed that a few content providers inject most contents into BitTorrent and that those content providers are located in foreign data centres. We also showed that an adversary could compromise the privacy of any peer in BitTorrent and identify the big downloaders that we define as the peers who subscribe to a large number of contents. This is a major privacy threat as it is possible for anybody in the Internet to reconstruct all the download and upload history of most BitTorrent users. This work was published in LEET 2010.

To circumvent this kind of monitoring, BitTorrent users are increasingly using anonymizing networks such as TOR to hide their IP address from the tracker and, possibly, from other peers. We explored in a second study whose goal was to Exploit P2P Applications to Trace and Profile Tor Users, to which extent a P2P protocol such as BitTorrent, when not designed to protect users information, leak information that may compromise the identity of users. We quantified such an issue with BitTorrent on top of anonymizing networks. We also designed an attack that reveals the identity of Tor users (We showed that it is possible to retrieve the IP address for more than 70% of BitTorrent users on top of TOR). Moreover, once the IP address of a peer is retrieved, it is possible to link to the IP address other applications used by this peer on top of TOR [45].

The fact that it is hard for a person to map an IP address to an identity mitigates the impact of the privacy attacks we described. However, we show that we can exploit a peer-to-peer VoIP system to associate a social identity (name, email address, etc.) to an IP address [46]. This means that anybody can now find this mapping that was only known by ISPs or big companies (like Google and Facebook), but never communicated unless in case of a legal action. The privacy threat is thus very high because this mapping enables blackmail, social attacks, targeted phishing attacks, etc.

As a proof of concept, we show that it is possible to track VoIP users mobility and BitTorrent downloads [46] using Skype, one of the most popular VoIP system with more than 500 millions registered users.

All these works received a very large media coverage (see [http://www-sop.inria.fr/members/Arnaud.Legout/Projects/bluebear.html](http://www-sop.inria.fr/members/Arnaud.Legout/Projects/bluebear.html)).
6.3. Network measurement, modeling and understanding


The main objective of our work in this domain is a better monitoring of the Internet and a better control of its resources. We work on new measurement techniques that scale with the fast increase in Internet traffic and growth of its size. We propose solutions for a fast and accurate identification of Internet traffic based on packet size statistics and host profiles. Within the ECODE FP7 project, we work on a network-wide monitoring architecture that, given a measurement task to perform, tune the monitors inside the network optimally so as to maximize the accuracy of the measurement results while limiting the overhead resulting from collected traffic. Within the ANR CMON project, we work on monitoring the quality of the Internet access by end-to-end probes, and on the detection and troubleshooting of network problems by collaboration among end users.

Next, is a sketch of our main contributions in this area.

- **Internet traffic classification by means of packet level statistics**
  One of the most important challenges for network administrators is the identification of applications behind the Internet traffic. This identification serves for many purposes as in network security, traffic engineering and monitoring. The classical methods based on standard port numbers or deep packet inspection are unfortunately becoming less and less efficient because of encryption and the utilization of non standard ports. In this activity, we come up with an online iterative probabilistic method that identifies applications quickly and accurately by only using the size of packets. Our method associates a configurable confidence level to the port number carried in the transport header and is able to consider a variable number of packets at the beginning of a flow. By verification on real traces we observe that even in the case of no confidence in the port number, a very high accuracy can be obtained for well known applications after few packets were examined. In another work [39], we make a complete study about the inter-packet time to prove that it is also a valuable information for the classification of Internet traffic. We discuss how to isolate the noise due to the network conditions and extract the time generated by the application. We present a model to preprocess the inter-packet time and use the result as input to the learning process. We discuss an iterative approach for the online identification of the applications and we evaluate our method on two different real traces. The results show that the inter-packet time is an important parameter to classify Internet traffic.

  We pursued this activity further by accounting for the communication profiles of hosts for the purpose of a better traffic classification [39], [38], [40]. We use the packet size and the inter-packet time as the main features for the classification and we benefit from the traffic profile of the host (i.e. which application and how much) to refine the classification and decide in favor of this or that application. The host profile is then updated online based on the result of the classification of previous flows originated by or addressed to the same host. We evaluate our method on real traces using several applications. The results show that leveraging the traffic pattern of the host ameliorates the performance of statistical methods. They also prove the capacity of our solution to derive profiles for the traffic of Internet hosts and to identify the services they provide.

  For a more thorough study of the traffic classification problem by means of packet statistics and host profiles, we refer to the PhD dissertation of Mohamad Jaber who was the main contributor to this activity inside the EPI Planete.

- **Adaptive network-wide traffic monitoring**
  The remarkable growth of the Internet infrastructure and the increasing heterogeneity of applications and users’ behavior make more complex the manageability and monitoring of ISP networks and raises the cost of any new deployment. The main consequence of this trend is an inherent disagreement between existing monitoring solutions and the increasing needs of management applications.
In this context, we work on the design of an adaptive centralized architecture that provides visibility over the entire network through a network-wide cognitive monitoring system. Given a measurement task, the proposed system drives its own configuration, typically the packet and flow sampling rates in routers, in order to address the tradeoff between monitoring constraints (processing and memory cost, collected data) and measurement task requirements (accuracy, flexibility, scalability). We motivate our architecture with an accounting application: estimating the number of packets per flow, where the flow can be defined in different ways to satisfy different objectives (e.g., Domain-to-Domain traffic, all traffic originated from a domain, destined to a domain). The architecture and the algorithms behind it are explained in paper published in 2010 for the case of a proactive control and in [43] for the case of a reactive control. In [44] the architecture and its algorithms are specified to a flow counting application. In all these works, the performances of our architecture are being validated in typical scenarios over an experimental platform we developed for the purpose of the study. Our platform is called MonLab (Monitoring Lab) and is described with more details in the Section on produced softwares. For now, MonLab presents a new approach for the emulation of Internet traffic and for its monitoring across the different routers. It puts at the disposal of users a real traffic emulation service coupled to a set of libraries and tools capable of Cisco NetFlow data export and collection, the overall destined to run advanced applications for network-wide traffic monitoring and optimization.

The activities in this direction are funded by the ECODE FP7 STREP project (Sep. 2008 - Dec. 2011). The dissertation of Imed Lassoued [21] provides an introduction to the field in addition to details on our contributions and the MonLab emulation platform.

- **Spectral analysis of packet sampled traffic**

  In network measurement systems, packet sampling techniques are usually adopted to reduce the overall amount of data to collect and process. Being based on a subset of packets, they hence introduce estimation errors that have to be properly counteracted by a fine tuning of the sampling strategy and sophisticated inversion methods. This problem has been deeply investigated in the literature with particular attention to the statistical properties of packet sampling and the recovery of the original network measurements. Herein, we propose a novel approach to predict the energy of the sampling error on the real time traffic volume estimation, based on a spectral analysis in the frequency domain. We start by demonstrating that errors due to packet sampling can be modeled as an aliasing effect in the frequency domain. Then, we exploit this theoretical finding to derive closed-form expressions for the Signal-to-Noise Ratio (SNR), able to predict the distortion of traffic volume estimates over time. The accuracy of the proposed SNR metric is validated by means of real packet traces. The analysis and the expressions of the SNR that stemmed from are described in [26]. In [52], we adopt such a model to design a real-time algorithm, that sets the IPFIX counter export timers in order to grant, to each flow, a target estimation accuracy. The work within this direction has been partially supported by the FP7 ECODE project.

- **Monitoring the quality of the Internet access by end-to-end probes**

  The detection of anomalous links and traffic is important to manage the state of the network. Existing techniques focus on detecting the anomalies but little attention has been devoted to quantify to which extent network anomaly affects the end user access link experience. We refer to this aspect as the local seriousness of the anomaly. In order to quantify the local seriousness of an anomaly, we consider the percentage of affected destinations, that we call the impact factor. In order to measure it, a host should monitor all possible routes to detect any variation in performance, but this is not practical in reality. In this activity, funded by the ANR CMON project, we work on finding estimates for the impact factor and the local seriousness of network anomalies through a limited set of measurements to random nodes we call landmarks.
We initially study the user access network to understand the typical features of its connectivity tree. Then, we define an unbiased estimator for the local seriousness of the anomaly and a framework to achieve three main results: (i) the computation of the minimum number of paths to monitor, so that the estimator can achieve a given significance level, (ii) the localization of the anomaly in terms of hop distance from the local user, and (iii) the optimal selection of landmarks. We are using real data to evaluate the local seriousness of the anomaly and to determine the sufficient number of landmarks to select randomly without knowing anything on the Internet topology. The localization mechanism leverages the study on the connectivity tree and the relationship between the impact factor and the minimum hop distance of an anomaly. Our first results show that the impact factor is a meaningful metric to evaluate the quality of Internet access. The current work focuses on extending this solution towards a collaborative setting where different end users collaborate together by exchanging the results of their observations. The objective will be a better estimation of the impact factor by each of them and a finer localization of the origin of any network problem.

On the experimental side, we have implemented the solution in a tool called ACQUA, which stands for Application for Collaborative Estimation of QUality of Internet Access. We design an anomaly detection mechanism based on the histogram of delay measurements and the likelihood of observations. Then, we give to ACQUA a pipeline based software architecture, and we go deeply into experimentation inside and outside Planetlab. We show what the properties and usage of the algorithm are, focusing also on how this tool can help us to get information about the network anomalies detected. Later we extend the idea of Impact Factor Estimation (IFE) by using what we call Inverse IFE from Planetlab, where the computer of the user whose connectivity is tested has a completely passive role in the measurements procedure. We study its strong and weak points, and we show conditions under which Inverse IFE from Planetlab gives similar results to traditional IFE.

- **Applied Internet Measurements**

  The performance of several Internet applications often relies on the measurability of path similarity between different participants. In particular, the performance of content distribution networks mainly relies on the awareness of content sources topology information. It is commonly admitted nowadays that, in order to ensure either path redundancy or efficient content replication, topological similarities between sources is evaluated by exchanging raw traceroute data, and by a hop by hop comparison of the IP topology observed from the sources to the several hundred or thousands of destinations. In this project, based on real data we collected, we advocate that path similarity comparisons between different Internet entities can be much simplified using lossy coding techniques, such as Bloom filters, to exchange compressed topology information. The technique we introduce to evaluate path similarity enforces both scalability and data confidentiality while maintaining a high level of accuracy. In addition, we demonstrate that our technique is scalable as it requires a small amount of active probing and is not targets dependent. This work has been published in [25].

- **Reliability of Geolocation Databases**

  In this project, we question the reliability of geolocation databases, the most widely used technique for IP geolocation. It consists in building a database to keep the mapping between IP blocks and a geographic location. Several databases are available and are frequently used by many services and web sites in the Internet. Contrary to widespread belief, geolocation databases are far from being as reliable as they claim. We conduct a comparison of several current geolocation databases -both commercial and free- to have an insight of the limitations in their usability. First, the vast majority of entries in the databases refer only to a few popular countries (e.g., U.S.). This creates an imbalance.

---

2 [http://planete.inria.fr/acqua/](http://planete.inria.fr/acqua/)
in the representation of countries across the IP blocks of the databases. Second, these entries do not reflect the original allocation of IP blocks, nor BGP announcements. In addition, we quantify the accuracy of geolocation databases on a large European ISP based on ground truth information. This is the first study using a ground truth showing that the overly fine granularity of database entries makes their accuracy worse, not better. Geolocation databases can claim country-level accuracy, but certainly not city-level. This study has been published in CCR [28].

- **Impact of Live Streaming Traffic**

  Video streaming is the most popular traffic in the Internet and a strong case for content centric networks. Therefore, it is fundamental to understand the network traffic characteristics of video streaming. In this work [49], we extensively studied the network traffic characteristics of YouTube and Netflix (the most popular video streaming traffic in the USA). We have shown that the traffic characteristics vastly depends on the type of browser, mobile application, and container (Flash, Silverlight, HTLM5) used.

6.4. Experimental Environment for Future Internet Architecture

**Participants:** Walid Dabbous, Thierry Parmentelat, Baris Metin, Frédéric Urbani, Daniel Camara, Alina Quereilhac, Shafqat Ur-Rehman, Thierry Turletti, Julien Tribino.

- **SFA Federation of experimental testbeds**

  The OneLab2 project has come to its end in spring 2010. We are now involved in the NOVI (E.U. STREP) project, the F-Lab (French A.N.R.) project, and have the lead of the “Federation” WorkPackage of OpenLab (E.U. IP) project. Within these frameworks, we are codevelopping with Princeton University a reference implementation for the Testbed-Federation architecture known as SFA for Slice-based Federation Architecture. As a sequel of former activities we also keep a low-noise maintenance activity of the PlanetLab software, which has been running in particular on the PlanetLab global testbed since 2004, with an ad-hoc federated model in place between PlanetLab Central (hosted by Princeton University) and PlanetLab Europe (hosted at INRIA) since 2007.

  During 2011 we have focused on the maturation of the SFA codebase, with several objectives in mind. Firstly we have contributed to a major overhaul of the specification as defined essentially within the GENI (N.S.F.) Project, with participations from all over the world. These changes, that affected both the core API and the schema used to expose and manage resource specifications, aimed at reaching a mature level of interoperability between the PlanetLab world and the EmuLab a.k.a. ProtoGeni world that has its own implementation, and are now available in SFA-2.0 issued late 2011.

  Secondly, the SFA codebase has been redesigned to provide a more generic shelter that other testbeds can easily leverage in order to come up with their own SFA-compliant wrapper. This is perceived as a powerful means to foster further adoption of the architecture, and the Planète team has been instrumental in bringing two entirely different testbeds to the federation, namely Senslab - developped in other INRIA Project-teams - and FEDERICA, the outcome of another E.U.-funded Project. Along the same lines we are working, although more remotely, with NICTA in Australia that publishes the O.M.F. testbed for running wireless testbeds, and who are interested in adopting this federation paradigm.

  Finally, as part of the pure PlanetLab development, we have added a feature for running nodes in a 'reservable’ mode, which breaks the usual best-effort PlanetLab model, but turned out very helpful both for making experiments possible, that needed a more reproducible behaviour of experiments, and also in a federation perspective, for closing the usage gap with, notably wireless testbeds, that typically have a reservable-only provisioning mechanisms.
- **Content Centric Networks Simulation**

We worked this year on the extension of the DCE framework for ns-3 in order to run CCN implementation under the ns-3 simulator. DCE stands for Direct Code Execution, its goal is to execute unmodified C/C++ binaries under ns-3 network simulator. With this tool researchers and developers can use the same code to do simulation and real experiments. DCE operation principle is to catch the standard systems calls done by the real application in the experiment and to emulate them within the ns-3 virtual network topology. Concerning CCN we use the PARC implementation named CCNx which is a well working open source software reference implementation of Content Centric Network protocol. As promised by DCE this integration of CCNx requires no modification of its code, it requires 'only' working on adding the system calls used by CCN that are not already supported by DCE. The advantage of this approach is that the integration work of CCN advanced DCE and will be useful in others completely different experiments. Another great advantage is that every evolution of the CCNx implementation is very easy to integrate, all what is needed is to compile the new source code. The next steps will be naturally to use DCE/ns-3 to evaluation CCN protocols in specific scenarios, to improve the coverage of systems calls supported by DCE, and to improve the DCE scheduler to be more realistic and to take into account CPU time spent in router queues. This work is done in the context of the ANR CONNECT project.

- **ns-3 Module store**

Bake is an integration tool which is used by software developers to automate the reproducible build of a number of projects which depend on each other and which might be developed, and hosted by unrelated parties. This software is being developed with the participation of the Planète group and is intended to be the automatic building tool adopted by the ns-3 project.

The client version of Bake is already working and the Planète group had a significant participation in its development. The contributions were in the context the addition of new functionalities, bug fixing and in the development of the regression tests. We are now starting the development of the ns-3 modules repository, which is a web portal to store the meta-information of the available modules. In the present state we have already designed and implemented the portal data basis and the main interface. It is already possible to register new modules and browse among the already registered ones.

The web portal has to be finished, notably the part that will create the xml file that will be used to feed the bake’s client. We also need to add new functionalities to the client part, to enable incremental build over partially deployed environments. As it is today, bake does not enable the user to add just one new module to an already deployed version of the ns-3 simulator. This work is done in the context of the ADT MobSim in collaboration with Hipercom and Swing Inria project-teams.

- **The ns-3 consortium**

We have founded last year a consortium between INRIA and University of Washington. The goals of this consortium are to (1) provide a point of contact between industrial members and the ns-3 project, to enable them to provide suggestions and feedback about technical aspects, (2) guarantee maintenance of ns-3’s core, organize public events in relation to ns-3, such as users’ day and workshops and (3) provide a public face that is not directly a part of INRIA or NSF by managing the [http://www.nsnam.org](http://www.nsnam.org) web site. This web site is now finalized. However, activities related to developing the consortium have slowed down during 2011 due to the leave of Mathieu Lacage. We plan to put more resources on this aspect in 2012.
• **Using Independent Simulators, Emulators, and Testbeds for Easy Experimentation**

Evaluating new network protocols, applications, and architectures uses many kinds of experimentation environments: simulators, emulators, testbeds, and sometimes, combinations of these. As the functionality and complexity of these tools increases, mastering and efficiently using each of them is becoming increasingly difficult.

We designed the preliminary prototype of the Network Experiment Programming Interface (NEPI) whose goal is to make easier the use of different experimentation environments, and switch among them easily. NEPI intends to make it possible to write a single script to control every aspect of a potentially mixed experiment, including a hierarchical network topology description, application-level setup, deployment, monitoring, trace setup, and trace collection. We showed how a single object model which encompasses every aspect of a typical experimentation workflow can be used to completely describe experiments to be run within very different experimentation environments.

The development of NEPI started in 2009 with the implementation of the core API, an address allocator, a routing table configurator, but also a prototype ns-3 backend driven by a simple graphical user interface based on QT. Last year, we validated and evolved the core API with the addition of a new backend based on Linux network namespace containers and stabilized the existing ns-3 backend. This year we have enhanced the design of NEPI and provided experiment validation, distributed experiment control, and failure recovery functionalities. In particular, we enforced separation between experiment design and execution stages, with off-line experiment validation. We also introduced a hierarchical distributed monitoring scheme to control experiment execution. We implemented a stateless message-based communication scheme, and added failure recovery mechanisms to improve robustness. The NEPI approach has been validated by implementing support for three complementary environments: a physical testbed, a network emulator, and a network simulator. Furthermore, we showed with a concrete experiment use case, available online for reproduction, how easy it is with NEPI to integrate these environments for hybrid-experimentation [48].

• **Guidelines for the accurate design of empirical studies in wireless networks**

Traditionally, wireless protocol proposals have been often tested and validated using only analytical and simulation models [73]. However, as the wireless environment is very complex to model accurately, and since the cost of wireless cards has decreased in an exponential way, today more and more research papers include evaluation of new proposals using experimentation on real devices. Indeed, experimentation is a mandatory step before possible deployment of new network protocols with real users. However, wireless experimentation is much more complex to set up and run than simulation, and it is important to avoid many pitfalls that can occur during experimentation. The objectives of this work are twofold. First, we described typical problems currently encountered in wireless-based experimentation, and we presented simple guidelines to avoid them [50]. Second, we proposed an experimental methodology where the detection of anomalies, calibration of the measurement setup, and clear definition of the scenario (among others) make easier the repeatability of results [55]. This work has been done in collaboration with Cristian Tala, Luciano Ahumada and Diego Dujovne from the Universidad Diego Portales of Chili, in the context of the WELCOME STIC AMSud 2011.

• **Multicast Video Streaming over WiFi Networks: Impact of Multipath Fading and Interference**
We conducted an experimental study in order to analyse the impact of interference, multipath fading and path loss on multicast video streaming (i.e., goodput, packet loss and ultimately on the video quality) using off-the-shelf fixed WiFi equipment in a wireless (802.11 b/g) local area network (WLAN) environment. We used the ricean K-factor as a measure of multipath fading, spectrum analyzer to estimate channel interference and received signal strength indicator (RSSI) as indication of signal power and attenuation. In order to realistically measure aforementioned metrics, we conducted extensive wireless experiments against six test cases representing common real-world situations using off-the-shelf wireless equipment.

We showed that interference has more impact on performance than multipath fading. Multipath fading can result in considerable performance degradation in environments where moving objects cause perturbation. On the contrary, channel interference is more frequent and more prominent cause of performance degradation in wireless networks because ISM 2.4 GHz band is increasingly being utilized in homes and work places. Being able to quantify the impact of multipath fading and interference is crucial in planning, troubleshooting, managing as well as benchmarking and optimizing wireless networks. This study has been published in MediaWin 2011 [51].

- **Making easier Experimentation**

  Wireless experimentations are challenging to evaluate due to the high variability of the channel characteristics and its sensitivity to interferences.

  Merging traces represents a complex problem especially in wireless experimentations, due to packet redundancy in multiple probes. Merging traces solutions need to be efficient in order to process the large amount of generated traces. These solutions should provide an output data structure that allows easy and fast analysis and must be scalable in order to be used in large and various experimental settings. We have designed an algorithm that performs trace synchronization and merging in a scalable way. The algorithm output is stored in a configured MYSQL database allowing for smart packet trace storage. This solution reduces processing time by 400% and storage space by 200% with regard to raw trace file solutions. It has been implemented in an open source software called CrunchXML, available under the GNU General Public License v2 at http://twiki-sop.inria.fr/twiki/bin/view/Projets/Planete/CrunchXML.

- **An Integration Framework for Network Experimentation**

  Many different experimentation environments address complementary aspects of network protocol evaluation, but because of their disparities and complexities it is often hard to use them to reproduce the same experiment scenario.

  Simulation is often used for the evaluation of new network protocols and architectures. In order to perform more realistic simulations, modern simulators such as ns-3 integrate more detailed models and even support direct execution of real protocol code. However, such complex models require more computational and memory requirements. We have studied the feasibility of a hybrid approach based on distributing a complex simulation scenario on several nodes in a grid network. We showed that by exploiting the real time operation of the ns-3 simulator, it is possible to map such complex scenarios on grid nodes. We also proposed a basic mapping algorithm to distribute a simulation scenario in several nodes [42].
5. New Results

5.1. RFID and Internet of Things

Participants: Roudy Dagher, Nathalie Mitton, Roberto Quilez, Loic Schmidt, David Simplot-Ryl, Lei Zhang.

5.1.1. Reader anti-collision protocol

In a Radio-Frequency IDentification network, while several readers are placed close together to improve coverage and consequently read rate, reader-reader collision problems happen frequently and inevitably. High probability of collision not only impairs the benefit of multi-reader deployment, but also results in misreadings in moving RFID tags. In order to eliminate or reduce reader collisions, we propose in [28] an Adaptive Color based Reader Anti-collision Scheduling algorithm (ACoRAS) for 13.56 MHz RFID technology where every reader is assigned a set of colors that allows it to read tags during a specific time slot within a time frame. Only the reader holding a color (token) can read at a time. Due to application constraints, the number of available colors should be limited, a perfect coloring scheme is not always feasible. ACoRAS tries to assign colors in such a way that overlapping areas at a given time are reduced. To the best of our knowledge ACoRAS is the first reader anti-collision algorithm which considers, within its design, both application and hardware requirements in reading tags. We show, through extensive simulations, that ACoRAS outperforms several anticollision methods and detects more than 99% of mobile tags while fitting application requirements.

5.1.2. Distributed ALE

Following the Internet of Things concept [14], each object is associated with a unique identifier which will allow to retrieve information about it in large databases. In the process of managing a large amount of objects, and consequently a large amount of events from readers, without overloading the network, these events have to be filtered and aggregated. This is the aim of the Application Level Events (ALE) standard from EPCGlobal, which receives events from readers and sends a useful and well constructed report to the business application. The ALE may be connected to several hundreds of readers. As the number of readers may increase with the increase of the company, a bottleneck may appear with all readers events sent to the ALE. A solution for scalability is to distribute the ALE. In [37], we propose an efficient way to solve this problem based on a Distributed Hash table (DHT). One role of the ALE is to insulate business application from technical concern so in our solution, we present a mechanism to distribute the ALE using Chord, a well-known peer-to-peer lookup system, and being transparent for business applications. This solution is compliant with the EPCglobal existing standard, scalable, robust and transparent for other layers of the middleware. We show that our solution generates only 10% overhead than in a nominal case while offering a better robustness and scalability when numbers of tags and readers increase significantly.

5.1.3. Advance Internet of Things

The Internet of Things (IoT) is a network of Internet-enabled objects, whose original purpose would be to interconnect all things in our daily life to build an always connected world. However, most of studies in the current IoT scientific community only focus on the radio-frequency identification (RFID) and wireless sensor network (WSN) based objects and lose the generality features endowed by the original definition of IoT. Furthermore, the emergence and proliferation of smart objects have been significantly changing our daily lives. It has been becoming evident that the objects should far beyond only "be identified and interconnected", but can also be controlled in an intelligent and transparent way independent of third party object (user) profiles and space and time span. In [39], we proposes a standardization scheme for a new paradigm: Advanced Internet of Things (AIoT), which is based on our proposed Unified Object Description Language (UODL) and allows to identify and interconnect every object and event with a standard format, and makes it easier and flexible for the third party control and management by integrating multiple services issued from cloud computing.
The purpose of our proposed AIoT scheme is to build a smart world of always on, always-awareness, always-connected, always-controllable, and establish an “intelligent networking” based relationship among the objects, service suppliers and the third party users. In the scope of AIoT, all the objects are transparent across the networks and can be identified and controlled (with security guarantees) via a standard prototype anytime and anywhere.

5.2. Topology control and neighbor discovery

Participants: Xu Li, Nathalie Mitton, Jovan Radak, David Simplot-Ryl, Isabelle Simplot-Ryl.

5.2.1. Topology control

Topology control is a tool for self-organizing wireless networks locally. It allows a node to consider only a subset of links/neighbors in order to later reduce computing and memory complexity. Topology control in wireless sensor networks is an important issue for scalability and energy efficiency. It is often based on graph reduction performed through the use of Gabriel Graph or Relative Neighborhood Graph. This graph reduction is usually based on geometric values.

In [35] we tackle the problem of possible connectivity loss in the reduced graph by applying a battery level based reduction graph. Experiments are conducted to evaluate our proposition. Results are compared with RNG [52] reduction which takes into account only the strength of the received signal (RSSI). Results show that our algorithm maintains network connectivity longer than solutions from the literature and balances the energy consumption over nodes.

In [31], we propose a radically new family of geometric graphs, i.e., Hypocomb, Reduced Hypocomb and Local Hypocomb for topology control. The first two are extracted from a complete graph; the last is extracted from a Unit Disk Graph (UDG). We analytically study their properties including connectivity, planarity and degree bound. All these graphs are connected (provided the original graph is connected) planar. Hypocomb has unbounded degree while Reduced Hypocomb and Local Hypocomb have maximum degree 6 and 8, respectively. To our knowledge, Local Hypocomb is the first strictly-localized, degree-bounded planar graph computed using merely 1-hop neighbor position information. We present a construction algorithm for these graphs and analyze its time complexity. Hypocomb family graphs are promising for wireless ad hoc networking. We report our numerical results on their average degree and their impact on FACE [49] routing. We discuss their potential applications and some open problems.

5.2.2. Neighbor discovery

To perform topology control, a node needs to discover its neighbors. Hello protocol is the basic technique for neighborhood discovery in wireless ad hoc networks. It requires nodes to claim their existence/aliveness by periodic "hello" messages. Central to a hello protocol is the determination of hello message transmission rate. No fixed optimal rate exists in the presence of node mobility. The rate should in fact adapt to it, high for high mobility and low for low mobility. In [30], we propose a novel mobility prediction based hello protocol, named ARH (Autoregressive Hello protocol). Each node predicts its own position by an ever-updated autoregression-based mobility model, and neighboring nodes predict its position by the same model. The node transmits "hello" message (for location update) only when the predicted location is too different from the true location (causing topology distortion), triggering mobility model correction on both itself and each of its neighbors. ARH evolves along with network dynamics, and seamlessly tunes itself to the optimal configuration on the fly using local knowledge only. Through simulation, we demonstrate the effectiveness and efficiency of ARH, in comparison with the only competitive protocol TAP (Turnover based Adaptive hello Protocol). With a small model order, ARH achieves the same high neighborhood discovery performance as TAP, with dramatically reduced message overhead (about 50% lower hello rate).
5.2.3. Address allocation

In [9], we propose a localized address autoconfiguration (LaConf) scheme for wireless ad hoc networks. Address allocation information is maintained on the network border nodes, called addressing agents (AAs), which are locally identified by a geographic routing protocol GFG (Greedy-FACE-Greedy). When a node joins the network, it acquires an address from a neighboring AA (if any exists) by local communication or from the head AA (a geographic extreme AA) by GFG-based multi-hop communication. A Geographic Hash Table (GHT) is adopted for duplicate address detection. Each address is hashed to a unique location in the network field, and the associated assignment information is stored along the face perimeter enclosing that location (in the planar graph). When a node receives an address assignment, it consults with the perimeter nodes around the hash location of the assigned address about any conflicts. AAs detects network partitions and mergers locally according to neighborhood change and triggers AA re-selection and network re-configuration (if necessary). We propose to apply a Connected Dominating Set (CDS) to improve the performance. We also evaluate LaConf through simulation using different planar graphs.

5.3. Routing

Participants: Nicolas Gouvy, Xu Li, Nathalie Mitton, David Simplot-Ryl.

In mobile wireless sensor networks, flows sent from data collecting sensors to a sink could traverse inefficient resource expensive paths. Such paths may have several negative effects such as devices battery depletion that may cause the network to be disconnected and packets to experience arbitrary delays. This is particularly problematic in event-based sensor networks (deployed in disaster recovery missions) where flows are of great importance. In [27], we use node mobility to improve energy consumption of computed paths. Mobility is a two-sword edge, however. Moving a node may render the network disconnected and useless. We propose CoMNet (Connectivity preservation Mobile routing protocol for actuator and sensor NETworks), a localized mechanism that modifies the network topology to support resource efficient transmissions. To the best of our knowledge, CoMNet is the first georouting algorithm which considers controlled mobility to improve routing energy consumption while ensuring network connectivity. CoMNet is based on (i) a cost to progress metric which optimizes both sending and moving costs, (ii) the use of a connected dominating set to maintain network connectivity. CoMNet is general enough to be applied to various networks (actuator, sensor). Our simulations show that CoMNet guarantees network connectivity and is effective in achieving high delivery rates and substantial energy savings compared to traditional approaches. CoMNet has then been extended in [26] to multi-hop movement.

In [12] we propose a novel localized Integrated Location Service and Routing (ILSR) scheme, based on the geographic routing protocol GFG, for data communications from sensors to a mobile sink in wireless sensor networks. The objective is to enable each sensor to maintain a slow-varying routing next hop to the sink rather than the precise knowledge of quick-varying sink position. In ILSR, sink updates location to neighboring sensors after or before a link breaks and whenever a link creation is observed. Location update relies on flooding, restricted within necessary area, where sensors experience (next hop) change in GFG routing to the sink. Dedicated location update message is additionally routed to selected nodes for prevention of routing failure. Considering both unpredictable and predictable (controllable) sink mobility, we present two versions. We prove that both of them guarantee delivery in a connected network modeled as unit disk graph. ILSR is the first localized protocol that has this property. We further propose to reduce message cost, without jeopardizing this property, by dynamically controlling the level of location update. A few add-on techniques are as well suggested to enhance the algorithm performance. We compare ILSR with an existing competing algorithm through simulation. It is observed that ILSR generates routes close to shortest paths at dramatically lower (90% lower) message cost.

In [29], we propose a novel trust management scheme for improving routing reliability in wireless ad hoc networks. It is grounded on two classic autoregression models, namely Autoregressive (AR) model and Autoregressive with exogenous inputs (ARX) model. According to this scheme, a node periodically measures the packet forwarding ratio of its every neighbor as the trust observation about that neighbor. These measurements constitute a time series of data. The node has such a time series for each neighbor. By applying
an autoregression model to these time series, it predicts the neighbors future packet forwarding ratios as their trust estimates, which in turn facilitate it to make intelligent routing decisions. With an AR model being applied, the node only uses its own observations for prediction; with an ARX model, it will also take into account recommendations from other neighbors. We evaluate the performance of the scheme when AR, ARX or a previously proposed Bayesian model is used. Simulation results indicate that the ARX model is the best choice in terms of accuracy.

5.4. Self-deployment, localization and area coverage

**Participants:** Milan Erdelj, Xu Li, Enrico Natalizio, Nathalie Mitton, Tahiry Razafindralambo, David Simplot-Ryl, Isabelle Simplot-Ryl.

5.4.1. Deployment

First steps in order to perform any task, a network needs to be deployed and nodes need to discover each others. To the best of our knowledge, very few scenarios when robots self-deploy to afterwards themselves constitute the network nodes or drop off sensor nodes have been investigated so far and none of them ensure the network connectivity at every step. In [15], we consider the self-deployment of wireless sensor networks. We present a mechanism which allows to preserve network connectivity during the deployment of mobile wireless sensors. Our algorithm is localized and is based on a subset of neighbors for motion decision. Our algorithm maintains a connected topology regardless of the direction chosen by each sensor. To preserve connectivity, the distance covered by the mobile nodes is constrained by the connectivity of the node to its neighbors in a connected subgraph like the relative neighborhood graph. We show the connectivity preservation property of our algorithm through analysis and present some simulation results on different deployment schemes such as full coverage, point of interest coverage or barrier coverage.

Another approach is the one proposed in [34] in which node placement is performed off-line with objective to optimize a criterion. Based on the specific application, different objectives can be taken into account such as energy consumption, throughput, delay, coverage, etc. Also many schemes have been proposed in order to optimize a specific quality of service (QoS) parameter. Power consumption is an essential issue in wireless multimedia sensor networks (WMSNs) due to the elevated processing capabilities requested by the video acquisition hardware installed on the generic sensor node. Hence, node placement scheme in WMSNs greatly impacts the overall network lifetime. [34] first proposes a suitable hardware architecture to implement a feasible WMS node based on off-the-shelf technology, then it evaluates the energy consumption obtained throughout a wise "energy-spaced" placement of the wireless nodes without affecting the video quality of multimedia traffic. In [4], we propose to use a neural network as a controller for nodes mobility and a genetic algorithm for the training of the neural network through reinforcement learning. This kind of scheme is extremely adaptive, since it can be easily modified in order to consider different objectives and QoS parameters. In fact, it is sufficient to consider a different kind of input for the neural network to aim for a different objective. All things considered, we propose a new method for programming a WSRN and we show practically how the technique works, when the coverage of the network is the QoS parameter to optimize. Simulation results show the flexibility and effectiveness of this approach even when the application scenario changes (e.g., by introducing physical obstacles).

5.4.2. Coverage

The coverage of Points of Interest (PoI) is a classical requirement in mobile wireless sensor applications. Optimizing the sensors self-deployment over a PoI while maintaining the connectivity between the sensors and the sink is thus a fundamental issue. [22] addresses the problem of autonomous deployment of mobile sensors that need to cover a predefined PoI with a connectivity constraints and provides the solution to it using Relative Neighborhood Graphs (RNG) [52]. Our deployment scheme minimizes the number of sensors used for connectivity thus increasing the number of monitoring sensors. Analytical results, simulation results and real implementation are provided to show the efficiency of our algorithm. To the best of our knowledge, only [21] both preserves the network connectivity and validates its proposition through experimentations with real wireless robots. This work has been extended to discovery and coverage of multi-point of interested in [
Indeed, the problems of multiple PoI coverage, environment exploration and data report are still solved separately and there are no works that combine the aforementioned problems into a single deployment scheme. In [21], we present a novel approach for mobile sensor deployment, where we combine multiple PoI coverage with network connectivity preservation and environment exploration in order to capture the dynamics of the monitored area. We examine the performance of our scheme through extensive simulation campaigns.

As sensors are energy constrained devices, one challenge in wireless sensor networks (WSNs) is to guarantee coverage and meanwhile maximize network lifetime. In [7], we leverage prediction to solve this challenging problem, by exploiting temporal-spatial correlations among sensory data. The basic idea lies in that a sensor node can be turned off safely when its sensory information can be inferred through some prediction methods, like Bayesian inference. We adopt the concept of entropy in information theory to evaluate the information uncertainty about the region of interest (RoI). We formulate the problem as a minimum weight submodular set cover problem, which is known to be NP hard. To address this problem, an efficient centralized truncated greedy algorithm (TGA) is proposed. We prove the performance guarantee of TGA in terms of the ratio of aggregate weight obtained by TGA to that by the optimal algorithm. Considering the decentralization nature of WSNs, we further present a distributed version of TGA, denoted as DTGA, which can obtain the same solution as TGA. The implementation issues such as network connectivity and communication cost are extensively discussed. We perform real data experiments as well as simulations to demonstrate the advantage of DTGA over the only existing competing algorithm and the impacts of different parameters associated with data correlations on the network lifetime.

5.4.3. Localization

In mobile-beacon assisted sensor localization, beacon mobility scheduling aims to determine the best beacon trajectory so that each sensor receives sufficient beacon signals with minimum delay. We propose a novel DeteRministic bEAcon Mobility Scheduling (DREAMS) algorithm [32], [10], without requiring any prior knowledge of the sensory field. In this algorithm, beacon trajectory is defined as the track of depth-first traversal (DFT) of the network graph, thus deterministic. The mobile beacon performs DFT under the instruction of nearby sensors on the fly. It moves from sensor to sensor in an intelligent heuristic manner according to RSS (Received Signal Strength)-based distance measurements. We prove that DREAMS guarantees full localization (every sensor is localized) when the measurements are noise-free. Then we suggest to apply node elimination and topology control (Local Minimum Spanning Tree) to shorten beacon tour and reduce delay. Through simulation we show that DREAMS guarantees full localization even with noisy distance measurements. We evaluate its performance on localization delay and communication overhead in comparison with a previously proposed static path based scheduling method.

5.5. Platforms and Substitution Networks


5.5.1. Platforms

In the framework of the ANR SensLAB project, a wireless sensor testbed has been set up in Lille in order to allow the evaluation through experimentations of scalable wireless sensor network protocols and applications. All functionalities offered by the platform have then been presented in [17], [16], [42]. SensLAB’s main and most important goal is to offer an accurate open access multisusers scientific tool to support the design, the development tuning, and the experimentation of real large-scale sensor network applications. The SensLAB testbed is composed of 1024 nodes over 4 sites. Each site hosts 256 sensor nodes with specific characteristics in order to offer a wide spectrum of possibilities and heterogeneity. Within a given site, each one of the 256 nodes is able both to communicate via its radio interface to its neighbors and to be configured as a sink node to exchange data with any other “sink node”. The hardware and software architectures that allow to reserve, configure, deploy firmwares and gather experimental data and monitoring information are described. We also present demonstration examples to illustrate the use of the SensLAB testbed and encourage researchers to test
and benchmark their applications/protocols on a large scale WSN testbed. A survey of platforms similar to SensLAB can be found in [6].

5.5.2. Emulation

Although some platforms like SensLAB are very convenient, they do not always fit the application requirements and setting up experimental testbeds of large scale wireless sensor networks requires huge cost, space and human resources. A more affordable approach is needed to provide preliminary insights on network protocols performance. To overcome the need for significant number of sensors required to perform a realistic experiment, and/or to experiments with high density networks, we introduce in [43] a novel approach: emulation by using all available sensors as candidate forwarding neighbors of the node S currently holding the packet. Destination position is virtual. After successfully sending message to forwarding node B over realistic wireless channel, the position of virtual destination is adjusted by translating it for vector BS and possibly rotating it to change the neighborhood configuration. The same node S then again selects new forwarding neighbor. Such selection of best forwarding neighbor continues until virtual destination appears close to a real node, and the later then becomes final destination node. Compared to real testbeds, our emulation has advantages of testing networks with very large densities (which may not be possible in small scale implementations), and in unlimited scalability of our physical implementations (e.g. we can emulate network with a million nodes).

5.5.3. Substitution network

A substitution network is a rapidly deployable backup wireless solution to quickly react to network topology changes due to failures or to flash crowd effects on the base network. Unlike other ad hoc and mesh solutions, a substitution network does not attempt to provide new services to customers but rather to restore and maintain at least some of the services available before the failure. Furthermore, a substitution network is not deployed directly for customers but to help the base network to provide services to customers. Therefore, a substitution network is not, by definition, a stand-alone network. [36] describes the quality of service architecture for substitution networks and discusses provisioning, maintenance, and adaptation of QoS inside and between the base network and the substitution network. In the same context, [33] shows the impact of the router mobility on the QoS of such networks.

5.6. Data collection and management

Participants: Thierry Delot, Geoffroy Cogniaux, Arnaud Fontaine, Alia Ghaddar, Michael Hauspie, Samuel Hym, Xu Li, Nathalie Mitton, Tahiry Razafindralambo, David Simplot-Ryl, Isabelle Simplot-Ryl.

5.6.1. Data collection

Wireless sensors networks (WSNs) are deployed to collect huge amounts of data from the environment. This produced data has to be delivered through sensor’s wireless interface using multi-hop communications toward a sink. The position of the sink impacts the performance of the wireless sensor network regarding delay and energy consumption especially for relaying sensors. Optimizing the data gathering process in multi-hop wireless sensor networks is, therefore, a key issue. [19] and [18] address the problem of data collection using mobile sinks in a WSN. We provide a framework that studies the trade-off between energy consumption and delay of data collection. This framework provides solutions that allow decision makers to optimally design the data collection plan in wireless sensor networks with mobile sinks.

In [20], [5], we focus on information gathering in vehicular ad hoc networks. Until now, only a few research works have addressed this problem. They have lead to solutions relying on push models, where potentially useful data are pushed towards vehicles. To the best of our knowledge, no work has tackled the use of pull models in VANETs. Such models would allow users to send queries to a set of cars in order to find the desired information. In order to propose such a query processing scheme, the main challenge to address is to route the different results towards their recipient in a highly dynamic network where the nodes move very quickly. To solve this issue, we propose GeoVanet, a DHT-based geographic routing protocol which ensures that the sender of a query can get a consistent answer. Our goal is not to compute the query result “instantaneously” but to ensure that the user will be able to retrieve it within a bounded time. To prove the effectiveness of GeoVanet,
an experimental evaluation is provided in the paper. It shows that up to 80% of the available query results are delivered to the user.

Another way to optimize data collection is to send data only when necessary. Knowledge discovery and data analysis in resource constrained wireless sensor networks faces different challenges. One of the main challenges is to identify misbehaviors or anomalies with high accuracy while minimizing energy consumption in the network. In [25], we extend a previous work of us and we present an algorithm for temporal anomalies detection in wireless sensor networks. Our experiments results show that our algorithm can efficiently and accurately detect anomalies in sensor measurements. It also produces low false alarm rate for slow variation time series measurements without harvesting the source of energy.

In data aggregation, sensor measurements from the whole sensory field or a sub-field are collected as a single report at an actor using aggregate functions such as sum, average, maximum, minimum, count, deviation, etc. We propose a localized Delay-bounded and Energy-efficient Data Aggregation (DEDA) protocol [11], [38] for request-driven wireless sensor networks with IEEE 802.11 CSMA/CA MAC layer. This protocol uses a novel two-stage delay model, which measures end-to-end delay using either hop count or degree sum along a routing path depending on traffic intensity. It models the network as a unit disk graph (UDG) and constructs a localized minimal spanning tree (LMST) sub-graph. Using only edges from LMST, it builds a shortest path (thus energy-efficient) tree rooted at the actor for data aggregation. The tree is used without modification if it generates acceptable delay, compared with a given delay bound. Otherwise, it is adjusted by replacing LMST sub-paths with UDG edges. The adjustment is done locally on the fly, according to the DESired Progress (DEP) value computed at each node. We further propose to integrate DEDA with a localized sensor activity scheduling algorithm and a localized connected dominating set algorithm, yielding two DEDA variants, to improve its energy efficiency and delay reliability. Through an extensive set of simulation, we evaluate the performance of DEDA with various network parameters. Our simulation results indicate that DEDA far outperforms the only existing competing protocol.

5.6.2. Data management

The use of reliable high-level languages based on virtual machines, such as java, is now possible on systems as small as smart cards or sensors. However, the potential of these languages is widely limited by hardware constraints as memory storage capacity etc. We claim that is lock may be leveraged by coupling cache mechanisms with external memory storages. [40] is a preliminary study of the set up of such an approach. Thanks to simulation based results, we identify three main factors which tend to decrease the performances of cache setting code in Java.

5.6.3. Data security

[41], [24] presents the enforcement of control flow policies for Java bytecode devoted to open and constrained devices. On-device enforcement of security policies mostly relies on run-time monitoring or inline checking code, which is not appropriate for strongly constrained devices such as mobile phones and smart-cards. We present a proof-carrying code approach with on-device lightweight verification of control flow policies statically at loading time. Policies are expressed by finite automata, the technique is in-between security automata and control flow security policies of Jensen et al. Our approach is suitable for evolving, open and constrained Java-based systems as it is compositional, to avoid re-verification of already verified bytecode upon loading of new bytecode, and it is regressive, to cleanly support bytecode unloading.

While mobile devices have become ubiquitous and generally multi-application capable, their operating systems provide few high level mechanisms to protect services offered by application vendors against potentially hostile applications coexisting on the device. In [23], we tackle the issue of controlling application interactions including collusion in Java-based systems running on open, constrained devices such as smart cards or mobile phones. We present a model specially designed to be embedded in constrained devices to verify at install-time that interactions between applications abide by the security policies of each involved application without resulting in run-time computation overheads; this models deals with application (un)installations and policy changes in an incremental fashion. We show the feasibility of our approach and its security
enhancements on a multi-application use case for GlobalPlatform/Java Card smart cards. This approach is developed in EVE - TCF.

Telecommunication software systems, containing security vulnerabilities, continue to be created and released to consumers. We need to adopt improved software engineering practices to reduce the security vulnerabilities in modern systems. Contracts can provide a useful mechanism for the identification, tracking, and validation of security vulnerabilities. In [8], we propose a new contract-based security assertion monitoring framework (CB SAMF) that is intended to reduce the number of security vulnerabilities that are exploitable across multiple software layers, and to be used in an enhanced systems development life cycle (SDLC). We show how contract-based security assertion monitoring can be achieved in a live environment on Linux. Through security activities integrated into the SDLC we can identify potential security vulnerabilities in telecommunication systems, which in turn are used for the creation of contracts defining security assertions. Our contract model is then exercised, as runtime probes, against two common security related vulnerabilities in the form of a buffer overflow and a denial of service.
4. New Results

4.1. Algorithms: Bandwidth Allocation in Optical Networks
Participants: Christine Fricker, Philippe Robert, James Roberts.

The development of dynamic optical switching is widely recognized as an essential requirement to meet anticipated growth in Internet traffic. Since September 2009, RAP has begun an investigation into the traffic management and performance evaluation issues that are particular to this technology. A first analysis of passive optical networks used for high speed Internet access has led to the proposal of an original dynamic bandwidth allocation algorithm and to an evaluation of its traffic capacity. Our activity on optical networking is carried out in collaboration with Orange Labs with whom we have had a research contract and are currently finalizing a new one. We have also established contacts with Alcatel-Lucent Bell Labs and had fruitful exchanges with Iraj Saniee and his team on their proposed time-domain wavelength interleaved networking architecture (TWIN).

We have also analyzed the traffic capacity of wavelength division multiplexing (WDM), passive optical networks (PONs) where user stations (optical network units) are equipped with tunable transmitters. For these systems users can use any of the multiple wavelengths to transmit their data but only within the limit determined by the number of transmitters they possess. A mean field approximation is investigated to estimate the capacity of a limited-gated multiserver polling system with a limit on the number of servers a given station can use simultaneously. The approximation provides an expression for the stability limit under very general assumptions about the traffic process and system configuration.

More generally, motivated by these next generation passive optical networks, a multi-server polling system has been studied where the number of servers that can attend to a queue simultaneously is limited. The stability condition is investigated for this model under quite general assumptions. The result is proved for unlimited service policies. The paper [1] presents a conjecture for the case of limited service policies and general service limits. A simulation study shows that the stability conditions may hold.

In 2011, we have worked on bandwidth allocation in meshed networks. A first study applies the TWIN architecture for a metropolitan area network but with an original medium access control (MAC) algorithm. This algorithm is inspired by our prior work on access networks and ensures an efficient and fair allocation of bandwidth to flows between network nodes. The paper [9] describes this network architecture and presents a performance evaluation using analytical models backed up by simulations.

The TWIN architecture is not extensible to a wide area for reasons of scalability and the excessive signalling delay between geographically distant nodes. We have therefore invented a new notion of a multipoint-to-multipoint lightpath that avoids these problems. A patent application relating to this invention has been submitted. This patent is owned by Orange following the terms of our contract with them. The second patent (that simply perfects the first invention) is jointly owned since the research was performed after the end of this contract. The submitted paper [13] describes the invention and its evaluation. A major advantage demonstrated in this paper is the energy saving achieved by the use of the proposed optical technology in place of electronic routers.

Ongoing research seeks to apply this type of networking solution to data centres, on one hand, and to geographically spread tier 1 Internet carrier networks, on the other. This work is performed in collaboration with Orange Labs and will be covered by a contract that is close to being finalized. We have also participated in the preparation of a European CELTIC project proposal that includes a work package dedicated to the development and experimentation of the network proposed in [13].

4.2. Algorithms: Content-Centric Networking
Participants: Mathieu Feuillet, Christine Fricker, Philippe Robert, James Roberts, Nada Sbihi.
RAP is participating in an ANR project named CONNECT which will contribute to the definition and evaluation of a new paradigm for the future Internet: a content-centric network (CCN) where, rather than interconnecting remote hosts like IP, the network directly manages the information objects that users publish, retrieve and exchange. CCN has been proposed by Van Jacobson and colleagues at the Palo Alto Research Center (PARC). In CCN, content is divided into packet-size chunks identified by a unique name with a particular hierarchical structure. The name and content can be cryptographically encoded and signed, providing a range of security levels. Packets in CCN carry names rather than addresses and this has a fundamental impact on the way the network works. Security concerns are addressed at the content level, relaxing requirements on hosts and the network. Users no longer need a universally known address, greatly facilitating management of mobility and intermittent connectivity. Content is supplied under receiver control, limiting scope for denial of service attacks and similar abuse. Since chunks are self-certifying, they can be freely replicated, facilitating caching and bringing significant bandwidth economies. CCN applies to both stored content and to content that is dynamically generated, as in a telephone conversation, for example. RAP is contributing to the design of CCN in two main areas:

- the design and evaluation of traffic controls recognizing that TCP is no longer applicable and queue management will require new, name-based criteria to ensure fairness and to realize service differentiation;
- the design and evaluation of replication and caching strategies that realize an optimal trade-off of expensive bandwidth for cheap memory.

The team will also contribute to the development of efficient forwarding strategies and investigate economic arguments that make CCN a viable replacement for IP.

The ANR project began in January 2011 and several task meetings have taken place. We have also held meetings with PARC establishing close cooperation with them and with some participants in the NSF project "Named Data Networking". We also participated in the CCN Community meeting in Palo Alto where we presented our work on traffic control. A paper describing the proposed flow-aware approach and results of a performance evaluation has been accepted for the conference Infocom 2012 [15].

Work on the performance of caching in CCN is ongoing. We have investigated popularity distributions for various types of content and evaluated their impact on the memory bandwidth tradeoff to be realized by CCN.

4.3. Algorithms: Channel Access algorithms in wireless networks

**Participants:** Mathieu Feuillet, Philippe Robert.

This is a collaboration with Thomas Bonald (Telecom ParisTech) and Alexandre Proutière (Microsoft Research). In wireless networks, to share available bandwidth between users is necessary. The bandwidth can be divided in several channels (frequency division) or the users can share the whole bandwidth by transmitting in different time slots (time division). We are studying different algorithms that allow users of a wireless network to access the channel. Those algorithms must avoid collisions and use the available bandwidth in the most efficient way. More and more wireless networks are decentralized and those algorithms must be distributed. Moreover, in order to use bandwidth in an efficient way, it is necessary to take the network topology into account. Recent studies have shown that it is possible to use the available bandwidth in a distributed and efficient way without message passing.

We studied a simplified version of the 802.11 channel access algorithm: CSMA/CA (Carrier Sense Multiple Access With Collision Avoidance). We proved that this algorithm does not use the bandwidth in the most efficient way and we proposed in 2010 a modification of this algorithm that is efficient. This result has been extended to multi-channels networks in [8] and its extended version [12].

4.4. Scaling Methods: Fluid limits in wireless networks

**Participant:** Philippe Robert.
This is a collaboration with Amandine Veber (CMAP, École Polytechnique). The goal is to investigate the stability properties of wireless networks when the bandwidth allocated to a node is proportional to a function of its backlog. This is, in some sense, a generalization of processor-sharing policies. We have investigated the fluid limits of simple examples of star topologies when the function used is \( \log \). We have shown that, under this scaling, some new phenomenon occurs, namely that a node may stabilize for some time at some very high level while the number of jobs of other nodes decreases at some fixed rate or remains finite, i.e. lives in the neighborhood of 0. An averaging phenomenon plays an important role for the return to equilibrium.

4.5. Algorithms: Distributed Hash Table

Participants: Mathieu Feuillet, Philippe Robert.

The Distributed Hash Table (DHTs) consists of a large set of nodes connected through the Internet. Each file contained in the DHT is stored in a small subset of these nodes. Each node breaks down periodically and it is necessary to have back-up mechanisms in order to avoid data loss. A trade-off is necessary between the bandwidth and the memory used for this back-up mechanism and the data loss rate. Back-up mechanisms already exist and have been studied thanks to simulation. To our knowledge, no theoretical study exists on this topic. We modeled this problem thanks to standard queues in order to understand the behavior of a single file and the global dynamic of the system. With a very simple centralized model, we have been able to emphasize a trade-off between capacity and life-time with respect to the duplication rate. From a mathematical point of view, we have been able to study different time scales of the system with an averaging phenomenon. An article is in preparation on this subject. A more sophisticated distributed model with mean field techniques is under investigation.

On the side of this project, we notably studied the distribution of hitting times of the classical Ehrenfest and Engset models by using martingale techniques, furthermore their asymptotic behavior has been analyzed when the size of the system increases to infinity [5].

4.6. Stochastic Modeling of Biological Networks

Participants: Emanuele Leoncini, Philippe Robert.

This is a collaboration with Vincent Fromion from INRA Jouy en Josas, which started on October 2010.

The goal is to propose a mathematical model of the production of proteins in prokaryotes. Proteins are biochemical compounds that play a key role in almost all the cell functions and are crucial for cell survival and for life in general. In bacteria the protein production system has to be capable to produce about 2500 different types of proteins in different proportions (from few dozens for the replication machinery to more than 100000 for certain key metabolic enzymes). Bacteria uses more than the 85\% of their resources to the protein production, making it the most relevant process in these organisms. Moreover this production system must meet two opposing problems: on one side it must provide the minimal proteins quantities in order to ensure the smooth-running of the cell, on the other side it can not choose a “overproduction policy” for all the proteins, since this would impact the global performance of the system and of the bacterium itself.

Gene expression is intrinsically a stochastic process: gene activation/deactivation occurs by means the encounter of polymerase/repressor with the specific gene, moreover many molecules that take part in the protein production act at extremely low concentrations.

We have restated mathematically the classical model using Poisson point processes. This representation, well-known in the field of queueing networks but, as far as we know, new in the gene expression modeling, allowed us to weaken few hypothesis of the existing models, in particular the Poisson hypothesis, which is well-suited in some cases, but that, in some situations, is far from the biological reality as we consider for instance the protein assemblage.
The theoretical environment of Poisson point processes has lead us to propose a new model of gene expression which captures on one side the main mechanisms of the gene expression and on the other side it tries to consider hypothesis that are more significant from a biological viewpoint. In particular we have modeled: gene activation/deactivation, mRNA production and degradation, ribosome attachment on mRNA, protein production and degradation.

We have shown how the probability distribution of the protein production and the protein lifetime may have a significant impact on the fluctuations of the number of proteins. We have obtained analytic formulas when the duration of protein assemblage and degradation follows a general probability distribution, i.e. without the Poisson hypothesis. We have used our model to compare the variances resulting by choosing different hypotheses for the probability distribution of the protein production and degradation, in particular we have hypothesize the protein assembly and degradation to be deterministic. The model has showed how, under the previous hypothesis, the variance on the number of proteins is bigger than the classical model with the Poisson hypothesis.

4.7. Stochastic networks: large bike sharing systems

**Participant:** Christine Fricker.

This is a collaboration with Nicolas Gast (EPFL) started in December 2010. Bike sharing systems were launched by numerous cities as a serious alternative in urban transportation, for example Velib (20 000 bikes, 1 500 stations). One of the major issues is the availability of the resources: bikes or free slots to return the bikes. These systems have become a hot topic in Operation Research but there are few studies on these stochastic networks. To our knowledge, no theoretical study of such bike sharing systems exists taking into account the limited capacity of the stations.

We modeled this system in a symmetrical case. Mean field limit theorems give the dynamic of a large system and the stationary behavior of a single station. Analytical results are obtained and convergence proved in the standard model via Lyapunov functions. It allows to find the best ratio of bikes per station and to measure the improvement of incentive mechanisms, as choosing among two stations for example, or redistribution of bikes by trucks. It is under investigation. Further results deal with heterogeneous systems. Our goal is to propose via a theoretical study and tests simple algorithms to improve the system behavior.

4.8. Stochastic networks: heterogeneity

**Participants:** Christine Fricker, Hanène Mohamed.

Mean field techniques applied to non-symmetrical systems are explored. It appears as a promising way to obtain analytical results on systems with clusters.

4.9. Stochastic Networks: Jackson Networks

**Participant:** Danielle Tibi.

Lyapunov functions and essential spectral radius of Jackson networks, joint work with I. Ignatiouk-Robert (University of Cergy-Pontoise). A family of explicit multiplicative Lyapunov functions is constructed for any stable Jackson network. Optimizing the multiplicative factor over this family provides an upper bound for the essential spectral radius of the associated Markov process. For some particular classes of Jackson networks, this upper bound coincides with a lower bound derived from large deviations arguments, thus providing the exact value of the essential spectral radius. The main example is given by Jackson networks with routing matrix having a tree structure (in the sense that for any node i, at most one other node can route its customers to i). The result also holds for other types of routing matrices (e.g. completely symmetrical), under some conditions over the different arrival and service rates.

4.10. Scaling Methods: Interaction of TCP Flows

**Participant:** Philippe Robert.
This is a collaboration with Carl Graham (CMAP, École Polytechnique). Mathematical modeling of data transmission in communication networks has been the subject of intense activity for some time now. For data transmission, the Internet can be described as a very large distributed system with self-adaptive capabilities to the different congestion events that regularly occur at its numerous nodes. The coexistence of numerous connections in a network with a general number of nodes has been analyzed in a previous work through a mean-field limit of a Markovian model describing the interaction of several classes of permanent connections. In [6], this line of work has been generalized to the case when connections are not permanent but can be either active (ON) when it is transmitting data along its route, or idle (OFF). This year, the analysis of dynamic arrivals and departures has been investigated. The main technical problem is that mean-field asymptotics are not anymore usable. Instead, fluid limit schemes have to be considered in a quite delicate context, random measures.
6. New Results

6.1. Introduction

In 2011, we focused our research on the following areas:

- distributed algorithms for large and dynamic networks,
- Complex queries over peer-to-peer networks
- Trust and reputation management on P2P networks
- dynamic adaptation of virtual machines,
- services management in large scale environments,
- Formal and practical study of optimistic replication, incorporating application semantics.
- Decentralized commitment protocols for semantic optimistic replication.

6.2. Distributed algorithms

Participants: Luciana Arantes [correspondent], Franck Petit, Maria Potop-Butucaru [correspondent], Swan Dubois, Pierre Sens, Julien Sopena.

Our current research in the context of distributed algorithms focuses on two main axes. We are interested in providing fault-tolerant and self-* (self-organizing, self-healing and self-stabilizing) solutions for fundamental problems in distributed computing. More precisely, we target the following basic blocks: mutual exclusion, resources allocation, agreement and communication primitives.

In dynamic systems we are interested in designing building blocks for distributed applications such as: failure detectors, adequate communication primitives (publish/subscribe) and overlays. Moreover, we are interested in solving fundamental problems such as synchronization, leader election, membership and naming, and diffusion of information.

6.2.1. Failure Detectors for Dynamic Systems

Since 2009, we explore a distributed computing model of dynamic networks such as (MANET or Wireless sensor networks). The temporal variations in the network topology implies that these networks can not be viewed as a static connected graph over which paths between nodes are established beforehand. Path between two nodes is in fact built over the time. Furthermore, lack of connectivity between nodes (temporal or not) makes of dynamic networks a partitionable system, i.e., a system in which nodes that do not crash or leave the system might not be capable to communicate between themselves. In 2011 we propose a new failure detector protocol which implements an eventually strong failure detectors (⋄S) on a dynamic network with an unknown membership. Failure detector is a fundamental service, able to help in the development of fault-tolerant distributed systems. Our failure detector has the interesting feature to be time-free, so that it does not rely on timers to detect failures; moreover, it tolerates mobility of nodes and message losses [41].

6.2.2. Self-* Distributed Algorithms

The main challenges of our research activity over 2011 year were to develop self-* (self-stabilizing, self-organizing and self-healing) distributed algorithms for various type of networks. Self-stabilization is a general technique to design distributed systems that can tolerate arbitrary transient faults. Since topology changes can be considered as a transient failures, self-stabilization turns out to be a good approach to deal with dynamic networks. This is particularly relevant when the distributed (self-stabilizing) protocol does not require any global parameters, like the number of nodes or the diameter of the network. With such a self-stabilizing protocol, it is not required to change global parameters in the program when nodes join or leave the system. Therefore, self-stabilization is very desirable to achieve scalability and dynamicity.

- Snap-stabilizing Committee Coordination. The classic committee coordination problem characterizes a general type of synchronization called n-ary rendezvous as follows:2:
“Professors in a certain university have organized themselves into committees. Each committee has an unchanging membership roster of one or more professors. From time to time a professor may decide to attend a committee meeting; it starts waiting and remains waiting until a meeting of a committee of which it is a member is started. All meetings terminate in finite time. The restrictions on convening a meeting are as follows: (1) meeting of a committee may be started only if all members of that committee are waiting, and (2) no two committees may convene simultaneously, if they have a common member. The problem is to ensure that (3) if all members of a committee are waiting, then a meeting involving some member of this committee is convened.”

In [31], we propose two snap-stabilizing distributed algorithms for the committee coordination problem. Snap-stabilization is a versatile technique allowing to design algorithms that efficiently tolerate transient faults. Indeed, after a finite number of such faults (e.g. memory corruptions, message losses, etc), a snap-stabilizing algorithm immediately operates correctly, without any external intervention. The first algorithm maximizes the concurrency, whereas the latter maximizes the fairness.

- **Snap-stabilizing Message Forwarding.** We focus on end-to-end request and response delivery of messages that are carried over the network. This problem is also known as the message forwarding problem. It consists in the management of network resources in order to forward messages, i.e., protocols allowing messages to move from a sender to receiver over the network. Combined with a self-stabilizing routing protocol, achieving snap-stabilization for the message forwarding problem is a very desirable property because every message sent by the sender is delivered in finite time to the receiver. In other words, no message that was actually sent after the system started is lost. In [46], we present a snap-stabilizing algorithm for the message forwarding that works on a tree topology. It uses a constant number of buffers per link, mainly $2\delta + 1$ buffers by node, where $\delta$ is the degree of the node. Therefore, it is particularly well suited for large-scale and dynamic systems, e.g., overlays used in peer-to-peer systems.

- **Self-Organizing Swarms of Robots.**

Consider a distributed system where the computing units are weak mobile robots, i.e., devices equipped with sensors and are designed to move. By weak, we mean that the robots are anonymous, autonomous, disoriented, and oblivious, i.e., devoid of (1) any local parameter (such that an identity) allowing to differentiate any of them, (2) any central coordination mechanism or scheduler, (3) any common coordinate mechanism or common sense of direction, and (4) any way to remember any previous observation nor computation performed in any previous step. Furthermore, all the robots follow the same program (uniform or homogeneous), and there is no kind of explicit communication medium. The robots implicitly “communicate” by observing the position of the others robots. Two different environments are considered: (i) the continuous two-dimensional Euclidian space wherein robot can observe, compute and move with infinite decimal precision, and (ii) the discrete model in which the space is partitioned into a finite number of locations, conveniently represented by a graph where nodes represent locations that can be sensed, and where edges represent the possibility for a robot to move from one location to the other.

During 2011, we mainly investigated the following problems: the gathering onto the plane, and the exploration of a finite discrete environment.

1. **Gathering.** This problem can be stated as follows: Robots, initially located at various positions, gather at the same position in finite time and remain at this position thereafter. In [20], we investigate the self-stabilizing gathering problem in the plane, that is gathering the robots deterministically with no kind of restriction on the initial configuration. In particular, robots are allowed to share same positions in the initial configuration. Strong multiplicity detection is the ability for the robots to count the exact number of robots located at a given position. We show that assuming strong multiplicity detection, it is

---

possible to solve the self-stabilizing gathering problem with \( n \) weak robots in the semi-synchronous model if, and only if, \( n \) is odd. By contrast, we show that with an odd number of robots, the problem becomes solvable. Our proof is constructive, as we present and prove a deterministic self-stabilizing algorithm for the gathering problem.

In [19], we address the gathering in the discrete environment. We prove some asymptotical time and space complexity lower bounds to solve the problem. We propose an algorithm that is asymptotically optimal in both space and round complexities. Finally, we show that most of the assumptions we made are necessary to deterministically solve the rendezvous considering our initial scenario.

2. **Graph Exploration.**

The exploration problem is to visit a discrete and finite environment by a swarm of robots. We consider two types of explorations: The finite exploration and the perpetual exploration. The former requires that \( k \) robots, initially placed at different nodes, collectively explore a graph before stopping moving forever. By “collectively” explore we mean that every node is eventually visited by at least one robot. In [73], we propose optimal (w.r.t, the number of robots) solutions for the deterministic exploration of a grid shaped network by a team of \( k \) asynchronous oblivious robots in the asynchronous and non-atomic asynchronous model. In more details, we first show that no exploration protocol exists with less than three robots for any grid with more than three nodes, less than four robots for the \((2,2)\)-Grid, and less than five robots for the \((3,3)\)-Grid. Next, we show that the problem is solvable using only 3 robots for any \((i,j)\)-Grid, provided that \( j > 3 \). Our result is constructive as we present a deterministic algorithm that performs in the non-atomic asynchronous model. We also present specific deterministic protocols for the \((3,3)\)-Grid using five robots.

The second type of exploration is the perpetual exploration. It requires every possible location to be visited by each robot infinitely often. In [32], we investigate the exclusive perpetual exploration of grid shaped networks. We focus on the minimal number of robots that are necessary and sufficient to solve the problem in general grids. In more details, we prove that three deterministic robots are necessary and sufficient, provided that the size of the grid is \( n \times m \) with \( 3 \leq n \leq m \) or \( n = 2 \) and \( m \geq 4 \). Perhaps surprisingly, and unlike results for the exploration with stop problem (where grids are "easier" to explore and stop than rings with respect to the number of robots), exclusive perpetual exploration requires as many robots in the ring as in the grid. Furthermore, we propose a classification of configurations such that the space of configurations to be checked is drastically reduced. This pre-processing lays the bases for the automated verification of our algorithm for general grids as it permits to avoid combinatorial explosion.

6.2.3. **Combining Fault-Tolerance and Self-stabilization in Dynamic Systems**

Recently, we started to investigate complex faults scenarios. Distributed fault-tolerance can mask the effect of a limited number of permanent faults, while self-stabilization provides forward recovery after an arbitrary number of transient faults hit the system. FTSS (Fault-Tolerant Self-Stabilizing) protocols combine the best of both worlds since they tolerate simultaneously transient and (permanent) crash faults. To date, deterministic FTSS solutions either consider static (i.e., fixed point) tasks, or assume synchronous scheduling of the system components. We proposed in [30] a fault-tolerant and stabilizing simulation of an atomic register. The simulation works in asynchronous message-passing systems, and allows a minority of processes to crash. The simulation stabilizes in a pragmatic manner, by reaching a long execution in which it runs correctly. A key element in the simulation is a new combinatorial construction of a bounded labeling scheme accommodating arbitrary labels, including those not generated by the scheme itself. Our simulation uses a self-stabilizing implementation of a data-link over non-FIFO channels [61]. In [23] we present the first study of deterministic FTSS solutions for dynamic tasks in asynchronous systems, considering the unison problem as a benchmark.
Unison can be seen as a local clock synchronization problem as neighbors must maintain digital clocks at most one time unit away from each other, and increment their own clock value infinitely often. We present several impossibility results for this difficult problem and propose a FTSS solution (when the problem is solvable) for the state model that exhibits optimal fault containment.

6.3. Peer-to-peer systems

Participants: Pierre Sens [correspondent], Nicolas Hidalgo, Sergey Legtchenko, Olivier Marin, Sébastien Monnet, Gilles Muller, Maria Potop-Butucaru, Mathieu Valero.

6.3.1. Peer-to-peer storage

Distributed Hash Table (DHTs) provide a means to build a completely decentralized, large-scale persistent storage service from the individual storage capacities contributed by each node of the peer-to-peer overlay. However, persistence can only be achieved if nodes are highly available, that is, if they stay most of the time connected to the overlay. Churn (i.e., nodes connecting and disconnecting from the overlay) in peer-to-peer networks is mainly due to the fact that users have total control on theirs computers, and thus may not see any benefit in keeping its peer-to-peer client running all the time.

When connection/disconnection frequency is too high in the system, data-blocks may be lost. This is true for most current DHT-based system’s implementations. To avoid this problem, it is necessary to build really efficient replication and maintenance mechanisms. Since 2008 we study the effect of churn on an existing DHT-based P2P system namely PAST/Pastry. We have proposed RelaxDHT [25], a churn-resilient peer-to-peer DHT. RelaxDHT proposes an enhanced replication strategy with relaxed placement constraints, avoiding useless data transfers and improving transfer parallelization. This new replication strategy is able to cut down by 2 the number of data-block losses compared to PAST DHT. We are now starting to study the use of erasure coding mechanisms along with replication within DHTs. Our goal is to propose hybrid mechanisms to find a good tradeoff among 1) churn-resilience, 2) maintenance cost, and 3) storage space.

6.3.2. Overlays

Large-scale distributed systems gather thousands of peers spread all over the world. Such systems need to offer good routing performances regardless of their size and despite high churn rates. To achieve that requirement, the system must add appropriate shortcuts to its logical graph (overlay). However, to choose efficient shortcuts, peers need to obtain information about the overlay topology. In case of heterogeneous peer distributions, retrieving such information is not straightforward. Moreover, due to churn, the topology rapidly evolves, making gathered information obsolete. State of-the-art systems either avoid the problem by enforcing peers to adopt a uniform distribution or only partially fulfill these requirements. To cope with this problem, we propose DONUT [47], a mechanism to build a local map that approximates the peer distribution, allowing the peer to accurately estimate graph distance to other peers with a local algorithm. The evaluation performed with real latency and churn traces shows that our map increases the routing process efficiency by at least 20% compared to the state-of-the-art techniques. It points out that each map is lightweight and can be efficiently propagated through the network by consuming less than 10 bps on each peer.

6.3.3. Distributed trees

Publish/Subscribe implemented on top of distributed R-trees (DR-trees) overlays offer efficient DHT-free communication primitives. We have then extend the distributed R-trees (DR-trees) in order to reduce event delivery latency in order to meet the requirements of massively distributed video games such that pertinent information is quickly distributed to all the interested parties without degrading the load of nodes neither increasing the number of noisy events. In 2011, we explore how to improve robustness of distributed trees. Since each single crash can potentially break the tree structure connectivity, DR-trees are crash-sensitivite. We this have proposed a fault tolerant approach which exploits replication of non leaf nodes in order to ensure the tree connectivity in presence of crashes. This work will be published in [56].
In [85], we consider a complete binary tree and construct a random pairing between leaf nodes and internal nodes. We prove that the graph obtained by contracting all pairs (leaf-internal nodes) achieves a constant node expansion with high probability. In the context of P2P overlays our result can be interpreted as follows: if each physical node participating to the tree overlay manages a random pair that couples one virtual internal node and one virtual leaf node then the physical-node layer exhibits a constant expansion with high probability which improves the robustness of the overlay.

6.3.4. Complex query over peer-to-peer networks

A major limitation of DHTs is that they only support exact-match queries. In order to offer range queries over a DHT it is necessary to build additional indexing structures. Prefix-based indexes, such as Prefix Hash Tree (PHT), are interesting approaches for building distributed indexes on top of DHTs. Nevertheless, the lookup operation of these indexes usually generates a high amount of unnecessary traffic overhead which degrades system performance by increasing response time. In [42], we propose a novel distributed cache system called Tabu Prefix Table Cache (TPT-C), aiming at improving the performance of the Prefix-trees. We have implemented our solution over PHT, and the results confirm that our searching approach reduces up to a 70% the search latency and traffic overhead. In [44], we propose DRing an efficient layered solution that directly supports range queries over a ring-like DHT structure. We improve load balancing by using only the nodes that store data, and by updating neighbour information through an optimistic approach.

6.3.5. Trust management in peer-to-peer networks

An ongoing research work focuses on trust assessment in dynamic systems. Even if it is near impossible to fully trust a node in a P2P system, managing a set of the most trusted nodes in the system can help to implement more trusted and reliable services. Using these nodes can reduce the probability of introducing malicious nodes in distributed computations. Our work aims at the following objectives: 1. To design a distributed membership algorithm for structured Peer to Peer networks in order to build a group of trusted nodes. 2. To design a maintenance algorithm to periodically clean the trusted group so as to avoid nodes whose reputation has decreased under the minimum value. 3. To provide a way for a given node X to find at least one trusted node. 4. To design a prototype of an information system, such as a news dissemination system, that relies on the trusted group. In 2011, we propose the CORPS system for building a community of reputable peers in Distributed Hash Tables [26].

6.4. Virtual machine (VM)

Participants: Harris Bakiras, Bertil Folliot, Gaël Thomas [correspondent], Gilles Muller [correspondent], Julia Lawall, Arie Middlekoop, Thomas Preud’homme, Suman Saha.

Our research interests are in improving the way systems software is developed. One theme of our research is the development of virtual machines with a specific focus on resource management, isolation and concurrency management. Another theme of our research is related to bug finding in systems software.

6.4.1. Virtual machines

Isolation in OSGi: The OSGi framework is a Java-based, centralized, component oriented platform. It is being widely adopted as an execution environment for the development of extensible applications. However, current Java Virtual Machines are unable to isolate components from each other’s. By modifying shared variables or allocating too much memory, a malicious component can freeze the complete platform. We work on I-JVM, a Java Virtual Machines that provides a lightweight approach to isolation while preserving the compatibility with legacy OSGi applications. Our evaluation of I-JVM shows that it solves the 15 known OSGi vulnerabilities due to the Java Virtual Machine with an overhead below 20%. I-JVM has been presented in DSN 2009.
VMKit: Managed Runtime Environments (MREs), such as the JVM and the CLI, form an attractive environment for program execution, by providing portability and safety, via the use of a bytecode language and automatic memory management, as well as good performance, via just-in-time (JIT) compilation. Nevertheless, developing such a fully featured MRE, including features such as a garbage collector and JIT compiler, is a herculean task. As a result, new languages cannot easily take advantage of the benefits of MREs, and it is difficult to experiment with extensions of existing MRE based languages. VMKit is a first attempt to build a common substrate that eases the development of high-level MREs. We have successfully used VMKit to build two MREs: a Java Virtual Machine (J3) and a Common Language Runtime (N3). VMKit has performance comparable to the well established open source MREs Cacao, Apache Harmony and Mono. VMKit is freely distributed under the LLVM licence with the LLVM framework developed by the University of Illinois at Urbana-Champaign and now maintained by Apple.

A third MRE is being build in cooperation with the "Algorithms, Programmes and Resolution” team in LIP6. This integrates a functional machine (the Zinc Abstract Machine) in VMKit and show that the adataptation at the language level of our virtual machine. This project has been funded by the LIP6 in 2009-10 and 2010-11.

6.4.2. Semantic patches

Open source infrastructure software, such as the Linux operating system, Web browsers and n-tier servers, has become a well-recognized solution for implementing critical functions of modern life. Furthermore, companies and local governments are finding that the use of open source software reduces costs and allows them to pool their resources to build and maintain infrastructure software in critical niche areas. Nevertheless, the increasing reliance on open source infrastructure software introduces new demands in terms of security and safety. In principle, infrastructure software contains security features that protect against data loss, data corruption, and inadvertent transmission of data to third parties. In practice, however, these security features are compromised by a simple fact: software contains bugs.

We are developing a comprehensive solution to the problem of finding bugs in API usage in open source infrastructure software based on our experience in using the Coccinelle code matching and transformation tool, and our interactions with the Linux community.

Coccinelle has been successfully used for finding and fixing bugs in systems code. One of our main recent results is an extensive study of bugs in Linux 2.6 [51] that has permitted us to demonstrate that the quality of code has been improving over the last six years, even though the code size has more than doubled.

We have used Coccinelle to generate traditional patches for improving the safety of Linux. Some Linux developers have also begun to use the tool. Over 800 patches developed using Coccinelle have been integrated into the mainline Linux kernel. As part of the ABL ANR project, we are building on the results of Coccinelle by designing semantic patches to identify API protocols and detect violations in their usage [24].

Another work done as part of the ANR ABL project, and as the topic of Suman Saha’s PhD thesis, is the improvement of error handling code in Linux. We developed a program analysis for identifying the code structures used to represent error handling code and a transformation to convert existing error handling code to use gotos to shared cleanup code, which is the style preferred by the Linux community [53]. We subsequently worked on finding bugs in error handling code, following an approach that focuses on local patterns, i.e., within the current function, rather than patterns occurring across the entire code base. This approach has a low rate of false positives and can find bugs in the use of rarely called functions [39].

6.5. Hosted database replication service

Participant: Mesoac Makpangou [correspondent].

Today, the vast majority of content distributed on the web are produced by web 2.0 applications. Examples of such applications include social networks, virtual universities, multi-players games, e-commerce web sites, and search engines. These applications rely on databases to serve end-users’ requests. Hence, the success of these applications/services depends mainly on the scalability and the performance of the database backend.
The objective of our research is to provide a hosted database replication service. With respect to end-users applications, this service offers an interface to create, to register, and to access databases. Internally, each hosted database is fragmented and its fragments are replicated towards a peer-to-peer network. We anticipate that such a service may improve the performance and the availability of popular web applications, thanks to partial replications of backend databases. Partial database replication on top of a peer-to-peer network raises a number of difficult issues: (i) enforcing replica consistency in presence of update transactions, without jeopardizing the scalability and the performance of the system? (ii) accommodating the dynamic and the heterogeneity of a peer-to-peer network with the database requirements?

We designed a database access protocol, capable to spread out a transaction’s accesses over multiple database fragments replicas while guaranteeing that each transaction observes a consistent distributed snapshot of a partially replicated database. We have also proposed a replica control substrate that permits to enforce 1-Copy SI for database fragments replicated over a wide area network. For that, unlike most database replication, we separate the synchronisation from the certification concerns.

A small-scale group of schedulers that do not hold database replicas, cooperate with one another to certify update transactions. Only certified transactions are notified to replicas. Furthermore, each replica will be notified only the transactions that impact the that it stores. Thanks to this separation, we avoid waste of computation resource at replicas that will be used to decide whether to abort or commit an update transaction; Our design choices also permit to reduce bandwidth consumption.

In 2010, we focus on the development of a prototype implementation of the complete system. The current prototype includes: a tool that helps fragment a database into fragments; the support to deploy dynamically, for each fragment, the suitable number of replicas towards the network of hosting peers; the implementation of our proposal (i.e. our database access protocol and our replica control substrate); and the JDBC API extension for accessing replicated databases.

### 6.6. Commitment protocols for WAN replication

**Participants:** Marc Shapiro [correspondent], Pierre Sutra, Masoud Saeida Ardekani.

In a large-scale distributed system, replication is an essential technique for improving availability and read performance. However, writes raise the issue of consistency, especially in the presence of concurrent updates, network failures, and hardware or software crashes. So-called consensus constitutes a major primitive to solving these issues. The performance of large-scale systems depends crucially on the latency of consensus, especially in wide-area networks; to decrease it, we focus on generalised consensus algorithms, i.e., ones that leverage the commutativity of operations and/or the spontaneous ordering of messages by the network. One such algorithms is Generalized Paxos, which does not order concurrent commutative operations. However, when a collision occurs (i.e., two replicas receive non-commuting operations in a different order) Generalized Paxos requires a very high latency to recover, completely negating the gain. We designed FGGC, a new generalised consensus algorithm that minimises the cost of recovering from a collision, without decreasing resilience to faults. FGGC achieves the optimal latency (two communication steps when processes receive non-commutative operations in the same order, and three otherwise) when there are no faults. FGGC remains optimally fault-tolerant, as it tolerates $f < n/2$ crash faults and requires only $f + 1$ processes to make progress. Our experimental evaluation of FGGC shows that it is more efficient than the competing protocols. Another topic of relevance in WANs is partial replication, i.e., where any given server holds only a fraction of all shared objects. This decreases the workload per server and improves access times. However, this makes transactional concurrency control more difficult; indeed most existing algorithms assume full replication. We designed and implemented two genuine consensus protocols for partial replication, i.e., ones in which only relevant replicas need participate in the commit of a transaction. They were evaluated experimentally above the BerkeleyDB database engine. This work is the topic of Pierre Sutra’s PhD thesis.

### 6.7. Optimistic approaches in collaborative editing

**Participants:** Marc Shapiro [correspondent], Marek Zawirski, Pierpaolo Cincilla.
In recent years, the Web has seen an explosive growth of massive collaboration tools, such as wiki and weblog systems. By the billions, users may share knowledge and collectively advance innovation, in various fields of science and art. Existing tools, such as the MediaWiki system for wikis, are popular in part because they do not require any specific skills. However, they are based on a centralised architecture and hence do not scale well. Moreover, they provide limited functionality for collaborative authoring of shared documents.

A natural research direction is to use P2P techniques to distribute collaborative documents. This raises the issue of supporting collaborative edits, and of maintaining consistency, over a massive population of users, shared documents, and sites.

In order to avoid complex and unnatural concurrency control and synchronisation, and to enable different styles of collaboration (from online “what you see is what I see” to fully asynchronous disconnected work) we invented the concept of a Commutative Replicated Data Type (CRDT). A CRDT is one where all concurrent operations commute. The replicas of a CRDT converge automatically, without complex concurrency control.

In the context of collaborative editing, we propose, a novel CRDT design called Treedoc. An essential property is that the identifiers of Treedoc atoms are selected from a dense space. We study practical alternatives for implementing the identifier space based on an extended binary tree. We also focus storage alternatives for data and meta-data, and mechanisms for compacting the tree. In the best case, Treedoc incurs no overhead with respect to a linear text buffer. We validate the results with traces from existing edit histories.

Treedoc will be used in ANR projects PROSE (Section 7.1.5 ) and STREAMS, and will be further studied and developed in ANR project ConcoRDanT (Section 7.1.3 ) and under a Google European Doctoral Fellowship.

6.8. CRDTs, a principled approach to eventual consistency

Participants: Marc Shapiro [correspondent], Marek Zawirski.

Most well-studied approaches to replica consistency maintain a global total order of operations. This serialisation constitutes a performance and scalability bottleneck, while the CAP theorem imposes a trade-off between consistency and partition-tolerance. An alternative approach, eventual consistency or optimistic replication, is attractive. A replica may execute an operation without synchronising a priori with other replicas. The operation is sent asynchronously to other replicas; every replica eventually applies all updates, but possibly in different orders. This approach ensures that data remains available despite network partitions, and is perceived to scale well and to provide acceptable quality of service. The consensus bottleneck remains but is off the critical path. However, reconciliation is generally complex. There is little theoretical guidance on how to design a correct optimistic system, and ad-hoc approaches have proven brittle and error-prone. We propose a simple, theoretically sound approach to eventual consistency, the concept of a convergent or commutative replicated data type (CRDT), for which some simple mathematical properties ensure eventual consistency. Provably, any CRDT converges to a common state that is equivalent to some sequential execution. A CRDT requires no synchronisation, thus every update can execute immediately, unaffected by network latency, faults, or disconnection. It is extremely scalable and is fault-tolerant, and does not require much mechanism. Previously, only a handful of CRDTs were known. Our current research aims to push the CRDT envelope, to study the principles of CRDTs, and to design a library of useful CRDTs. So far we have designed variations on registers, counters, sets, maps (key-value stores), graphs, and sequences. Potential application areas include computation in delay-tolerant networks, latency tolerance in wide-area networks, disconnected operation, churn-tolerant peer-to-peer computing, and partition-tolerant cloud computing. CRDTs are the main topic of ANR project ConcoRDanT (Section 7.1.3 ). This research is also funded in part by a Google European Doctoral Fellowship.
6. New Results

6.1. Optimized protocols implementation and networking equipments

6.1.1. Locating Virtual Infrastructures: Users and InP Perspectives

Participants: Paulo Gonçalves, Guilherme koslovski.

This is a joint work with Pascale Vicat-Blanc (Lyatiss) and Sébastien Soudan (Lyatiss).

The Cloud Computing wave consolidates the on-demand provisioning of configurable virtual machines. Recent projects have proposed the extension of the original IaaS paradigm to provide dynamic virtual networks to interconnect virtual IT resources, composing Virtual Infrastructures (VIs). In this new scenario, users with different objectives and expectations can rent dynamically provisioned virtual infrastructures to execute their applications during a given time slot. VIs can be allocated anywhere on top of a distributed and virtualized substrate. This decoupling from the geographical location introduces concerns such as a latency increase in network communications (user’s perspective), and the fragmentation of physical resources (Infrastructure Provider’s - InP - perspective). This context motivates efforts to investigate and deploy new models and tools which consider the geographical location of virtual infrastructures. Our work concentrates on the allocation of VIs guided by both the user’s and the InP’s constraints. We propose a formulation of the allocation problem considering the user’s expectations as well as the physical-substrate provider’s goals. Our initial experiments demonstrate that it is possible to improve the quality of the virtual-infrastructure allocation (user perspective) while simultaneously decreasing the physical substrate’s fragmentation and the substrate’s cost.

6.1.2. Energy-efficient reservation infrastructure for large-scale distributed systems

Participants: Anne-Cécile Orgerie, Laurent Lefèvre, Guérin-Lassous Isabelle.

Over the past few years, the energy consumption of Information and Communication Technologies (ICT) has become a major issue. Nowadays, ICT accounts for 2% of the global CO2 emissions, an amount similar to that produced by the aviation industry. Large-scale distributed systems (e.g. Grids, Clouds and high-performance networks) are often heavy electricity consumers because – for high-availability requirements – their resources are always powered on even when they are not in use. Reservation-based systems guarantee quality of service, allow for respect of user constraints and enable fine-grained resource management. For these reasons, in the context of Anne-Cecile Orgerie Phd (defended in September 2011), we proposed an energy-efficient reservation framework to reduce the electric consumption of distributed systems and dedicated networks. The framework, called ERIDIS, is adapted to three different systems: data centers and grids, cloud environments and dedicated wired networks. By validating each derived infrastructure, we show that significant amounts of energy can be saved using ERIDIS in current and future large-scale distributed systems [54].

6.1.3. Energy efficiency in exascale infrastructures

Participants: Mehdi Diouri, Olivier Gluck, Laurent Lefèvre.

Joint work with F. Cappello (JLPC, joint laboratory between INRIA and NCSA).

In Diouri’s PhD, we address the issue of energy efficiency for exascale supercomputers. We first proposed a green architecture for exascale systems gathering some new solutions to “consume less” energy and to “consume better”. This architecture involves interactions with the different actors interfering directly or indirectly with the supercomputer: its user, its administrator, its resource manager and the energy supplier. Then we were interested into leaning on this green architecture in order to propose some green services that will be offered for applications that will run on exascale systems. Our approach consists in evaluating the power overhead induced by some existing services, and by proposing a green version of these services that takes into account the constraints imposed by the different actors involved. In 2011, we specifically aimed to apply our approach for fault tolerance protocols in their normal functioning stage and in case of failure. [38]
6.1.4. Energy profiling and green leverages for high performance computing applications

Participants: Ghislain Landry Tsafack, Jean-Patrick Gelas, Laurent Lefevre.

Ghislain Landry TSAFACK CHETSA has started his PhD in January 2011, within the framework of INRIA HEMERA project, on: “Energy profiling and green leverages for high performance computing applications” (co-advisement with Jean-Marc PIERSON and Patritia STOLF from IRIT). During the course of this first year, we have investigated the possibility of characterizing distributed applications considering their energy / power profile. We first carried out a set of experiments for a better understanding of the application’s behavior and the impact that this behavior may have on its power consumption. Results led us to the assumption that any individual application run can be represented as a sequence of basic operations including computation, memory accesses, disk and network accesses over a given time period. We next rely on that assumption to define application’s energy profile. Application’s energy profile helps to prevent the fallout of any action that may be taken to reduce its power usage. To guarantee reasonable results, i.e., reduce energy with less performance degradation, we designed an energy prediction model capable of predicting power usage of a wide range of high performance computing (HPC) applications.

6.1.5. Towards virtualized home gateways

Participants: Jean-Patrick Gelas, Laurent Lefevre, Anne-Cécile Orgerie.

Joint work with Dino Lopez Pacheco (Univeristy of Nice) and Referi Assefa (Addis Abeba University, Ethiopia).

Virtualizing services located on end to end parts of the networks and making them available for a large number of applications and users is now becoming a real challenge. Within the scope of the GreenTouch project, we are exploring models, simulations tools (ECOFEN) and software prototypes able to demonstrate the impact of such approach in terms of energy reduction.

6.2. Quality of service and transport layer for future networks

6.2.1. On the Impact of the Flow-Size Distribution’s Tail Index on Network Performance with TCP Connections

Participant: Paulo Gonçalves.

This is a joint work with Oana Goga (UPMC, Lip6) and Patrick Loiseau (Eurecom).

In this work, we studied the impact of the flow-size distribution on network performance in the case of a single bottleneck with finite buffer. To tackle the case where flows are transmitted with the TCP protocol, we use real experiments and ns-2 simulations. Our preliminary results show that the distribution’s tail index impacts the performance in a more complex way than what is reported in existing literature. In particular, we exhibit situations where a heavier tail gives better performance for certain metrics. We argue that a main cause of our observed results is the transient behavior at the beginning of each flow.

6.2.2. Available Bandwidth Estimation for Multi-hop Wireless Networks

Participants: Isabelle Guériin Lassous, Van Nam Nguyen.

Estimating the available bandwidth in IEEE 802.11-based multi-hop wireless networks is a very difficult task due to the medium sharing among contending nodes and collisions between hidden stations. Several methods have been proposed so far for these networks to compute the available bandwidth on wireless links. If some recent solutions such as ABE and IAB now take into account collisions and their impact on the mean backoff, none of them considers the packet retransmissions due to collisions, although these retransmissions have an impact on the available bandwidth. In this work, we have proposed a new available bandwidth estimation for multi-hop wireless networks called RABE (Retransmission-based Available Bandwidth). This method integrates the average number of retransmission attempts in the available bandwidth estimation, in addition to other relevant parameters like the idle periods durations and the collision probability. RABE has been evaluated
by simulation and the obtained results show that RABE can achieve a mean error ratio of 17% in comparison with the real measurement. Furthermore RABE is at least two times more accurate than ABE and ten times more accurate than IAB.

6.2.3. **On The Recovery Performance of Single- and Multipath OLSR in Wireless Multi-Hop Networks**

Participants: Inès Doghri, Isabelle Guérin Lassous.

In this work, we study and improve the recovery properties of single- and multipath routing strategies when facing network failure situations. In particular, we focus our study on two MANET routing protocols: OLSR and its multipath extension MP-OLSR. In various wireless multi-hop network environments, especially in multiple chain topologies, we define and evaluate the latency introduced by these protocols to find a new path after a link failure. Theoretical estimations and simulation results show that, under dual chain-topologies, this latency can be too long and incompatible with the needs of loss and delay constrained applications. As the source nodes cannot detect link failures immediately because of the delay incurred by the well-known nature of link state protocols in general, and of OLSR Topology Control (TC) messages in particular, these nodes keep sending packets along broken paths. We thus study the inconsistencies between the actual network topology and the nodes’ own representation. After analyzing the consequences of this long latency, we seek to alleviate these problems with the introduction of adapted mechanisms. We propose three new different schemes and accordingly extend the original OLSR and MP-OLSR protocols in order to decrease the expected latency and improve the protocol performance. Simulation results show a steep decrease of the latency when using these new schemes in dual chain-topologies. We also discuss these results in terms of packet loss, end-to-end delay and overhead.

6.3. High Speed Network’s traffic metrology and statistical analysis

6.3.1. **A long-range dependent model for network traffic with flow-scale correlations**

Participant: Paulo Gonçalves.

This is a joint work with Patrick Loiseau (Eurecom) and Pascale Vicat-Blanc (Lyatiss).

For more than a decade, it has been observed that network traffic exhibits long-range dependence and many models have been proposed relating this property to heavy-tailed flow durations. However, none of these models consider correlations at flow scale. Such correlations exist and will become more prominent in the future Internet with the emergence of flow-aware control mechanisms correlating a flow’s transmission to its characteristics (size, duration, etc.). In our present work, we study the impact of the correlation between flow rates and durations on the long-range dependence of aggregate traffic. Our results extend those of existing models by showing that two possible regimes of long-range dependence exist at different time scales. The long-range dependence in each regime can be stronger or weaker than standard predictions, depending on the conditional statistics between the flow rates and durations. In the independent case, our proposed model consistently reduces to former approaches. The pertinence of our model is validated on real web traffic traces, and its ability to accurately explain the Hurst parameter is validated on both web traces and numerical simulations.

6.3.2. **A reccurent solution of Ph/M/c/N-like and Ph/M/c-like queues**

Participant: Thomas Begin.

This work has been accepted for publication by the Journal of Applied Probability [50] and was performed in collaboration with Pr. Brandwajn (UCSC).
We propose an efficient semi-numerical approach to compute the steady-state probability distribution for the number of requests at arbitrary and at arrival time instants in Ph/M/c-like systems in which the inter-arrival time distribution is represented by an acyclic set of memoryless phases. Our method is based on conditional probabilities and results in a simple computationally stable recurrence. It avoids the explicit manipulation of potentially large matrices and involves no iteration. Due to the use of conditional probabilities, it delays the onset of numerical issues related to floating-point underflow as the number of servers and/or phases increases. For generalized Coxian distributions, the computational complexity of the proposed approach grows linearly with the number of phases in the distribution.

6.3.3. A Markovian model based on SIR epidemic classification to reproduce the workload dynamics of a VoD server

Participants: Shubhabrata Roy, Thomas Begin, Paulo Gonçalves.

We have devised a Markovian model, based on the SIR epidemic classification, to reproduce the workload dynamics that can be observed on a VoD (Video on Demand) server. This model basically relies on the dynamic between three distinct populations (i.e., current watchers, past watchers and potential watchers). It also embeds events with very low probability but high impact on its overall behavior corresponding to the occurrence of a flash crowd or the the buzz effect on a VoD server. The steady-state solution to this model has shown that it exhibits a behavior qualitatively close to what can be expected from a real-life VoD server. We have also shown that the workload process as delivered this model satisfies a large deviation principle. Our future work aims at taking advantage of this information to devise a new scheme for allocating available resources in a VoD server.

6.3.4. A comparative study of existing MBAC using real network traces

Participants: Doreid Ammar, Thomas Begin, Isabelle Guérin-Lassous.

We have evaluated the respective performance of several MBACs (Measurement-based admission control) using a realistic framework in which the pattern of the background traffic follows experimental traces collected on real-life networks. This study has allowed to highlight the respective discrepancies between MBACs in terms of easiness to implement and attained performance. This work will now focus on the design of a new MBAC based on a iteratively learned model.

6.3.5. Graph Based Classification of Content and Users in BitTorrent

Participants: Paulo Gonçalves, Marina Sokol.

This is a joint work with Konstantin Avrachenkov (INRIA Maestro) and Arnaud Legout (INRIA Planete).

P2P downloads still represent a large portion of today’s Internet traffic. More than 100 million users operate BitTorrent and generate more than 30% of the total Internet traffic. Recently, a significant research effort has been done to develop tools for automatic classification of Internet traffic by application. The purpose of our present work is to provide a framework for sub-classification of P2P traffic generated by the BitTorrent protocol. Unlike previous works, we cannot rely on packet level characteristics and on the standard supervised machine learning methods. The application of the standard supervised machine learning methods is based on the availability of a large set of parameters (packet size, packet inter-arrival time, etc.). Since P2P transfers are based on the same BitTorrent protocol we cannot use this set of parameters to classify P2P content and users. Instead we can make use of the bipartite user-content graph. This is a graph formed by two sets of nodes: the set of users (peers) and the set of contents (downloaded files). From this basic bipartite graph we also construct the user graph, where two users are connected if they download the same content, and the content graph, where two files are connected if they are both downloaded by at least one same user. The general intuition is that the users with similar interests download similar contents. This intuition can be rigorously formalized with the help of graph based semi-supervised learning approach.
6.3.6. Generalized Optimization Framework for Graph-based Semi-supervised Learning

Participants: Paulo Gonçalves, Marina Sokol.

This is a joint work with Konstantin Avrachenkov (INRIA Maestro).

We develop a generalized optimization framework for graph-based semi-supervised learning. The framework gives as particular cases the Standard Laplacian, Normalized Laplacian and PageRank based methods. We have also provided new probabilistic interpretation based on random walks and characterized the limiting behavior of the methods. The random walk based interpretation allows us to explain differences between the performances of methods with different smoothing kernels. It appears that the PageRank based method is robust with respect to the choice of the regularization parameter and the labelled data. We illustrate our theoretical results with two realistic datasets, characterizing different challenges: Les Miserables characters social network and Wikipedia hyper-link graph. The graph-based semi-supervised learning classifies the Wikipedia articles with very good precision and perfect recall employing only the information about the hyper-text links.

6.3.7. On the estimation of the large deviations spectrum

Participant: Paulo Gonçalves.

This is a joint work with Julien Barral (Univ. Paris 13)

We propose an estimation algorithm for large deviations spectra of measures and functions. The algorithm converges for natural examples of multifractals.

6.3.8. Adaptive Multiscale Complexity Analysis of Fetal Heart Rate

Participant: Paulo Gonçalves.

This is a joint work with Patrice Abry (ENS Lyon, CNRS) and Muriel Doret (Hospice civils de Lyon, Univ. Lyon 1)

Per partum fetal asphyxia is a major cause of neonatal morbidity and mortality. Fetal heart rate monitoring plays an important role in early detection of acidosis, an indicator for asphyxia. This problem is addressed in this paper by introducing a novel complexity analysis of fetal heart rate data, based on producing a collection of piecewise linear approximations of varying dimensions from which a measure of complexity is extracted. This procedure specifically accounts for the highly non-stationary context of labor by being adaptive and multiscale. Using a reference dataset, made of real per partum fetal heart rate data, collected in situ and carefully constituted by obstetricians, the behavior of the proposed approach is analyzed and illustrated. Its performance is evaluated in terms of the rate of correct acidosis detection versus the rate of false detection, as well as how early the detection is made. Computational cost is also discussed. The results are shown to be extremely promising and further potential uses of the tool are discussed.
5. New Results

5.1. Package understanding and Assessing

Participants: Stéphane Ducasse, Nicolas Anquetil, Usman Bhatti, Jannik Laval.

To support the understanding of large systems is to offer ways to understand and fix dependencies between software elements. We worked on how to semi-automatically reorganize packages to minimize coupling.

**Efficient Retrieval and Ranking of Undesired Package Cycles in Large Software Systems.** Many design guidelines state that a software system architecture should avoid cycles between its packages. Yet such cycles appear again and again in many programs. We believe that the existing approaches for cycle detection are too coarse to assist the developers to remove cycles from their programs. We describe an efficient algorithm that performs a fine-grained analysis of the cycles among the packages of an application. In addition, we define a metric to rank cycles by their level of undesirability, prioritizing the cycles that seem the least desirable to the developers. Our approach is validated on two large and mature software systems in Java and Smalltalk. [19]

**Legacy Software Restructuring: Analyzing a Concrete Case.** Software re-modularization is an old preoccupation of reverse engineering research. The advantages of a well structured or modularized system are well known. Yet after so much time and efforts, the field seems unable to come up with solutions that make a clear difference in practice. Recently, some researchers started to question whether some basic assumptions of the field were not overrated. The main one consists in evaluating the high-cohesion/low-coupling dogma with metrics of unknown relevance. In this paper, we study a real restructuring case (on the Eclipse platform) to try to better understand if (some) existing metrics would have helped the software engineers in the task. Results show that the cohesion and coupling metrics used in the experiment did not behave as expected and would probably not have helped the maintainers reach their goal. We also measured another possible restructuring which is to decrease the number of cyclic dependencies between modules. Again, the results did not meet expectations. [14]

**An empirical model for continuous and weighted metric aggregation.** It is now understood that software metrics alone are not enough to characterize software quality. To cope with this problem, most of advanced and/or industrially validated quality models aggregate software metrics: for example, cyclomatic complexity is combined with test coverage to stress the fact that it is more important to cover complex methods than accessors. Yet, aggregating and weighting metrics to produce quality indexes is a difficult task. Indeed certain weighting approaches may lead to abnormal situations where a developer increasing the quality of a software component sees the overall quality degrade. Finally, mapping combinations of metric values to quality indexes may be a problem when using thresholds. In this paper [20], we present the problems we faced when designing the Squale quality model, then we present an empirical solution based on weighted aggregations and on continuous functions. The solution has been termed the Squale quality model and validated over 4 years with two large multinational companies: Air France-KLM and PSA Peugeot-Citroen.

**Modularization Metrics: Assessing Package Organization in Legacy Large Object-Oriented Software.** In systems consisting of several thousands of classes, classes cannot be considered as units for software modularization. In such context, packages are not simply classes containers, but they also play the role of modules: a package should focus on providing well identified services to the rest of the software system. Therefore, understanding and assessing package organization is primordial for software maintenance tasks. Although there exist a lot of works proposing metrics for the quality of a single class and/or the quality of inter-class relationships, there exist few works dealing with some aspects for the quality of package organization and relationship. We believe that additional investigations are required for assessing package modularity aspects. The goal of these papers [13], [28] is to provide a complementary set of metrics that assess some modularity principles for packages in large legacy object-oriented software: Information-Hiding, Changeability and Reusability principles. Our metrics are defined with respect to object-oriented dependencies.
that are caused by inheritance and method call. We validate our metrics theoretically through a careful study of the mathematical properties of each metric.

5.2. Tools and Tool Infrastructure

**Participants:** Stéphane Ducasse, Veronica Uquillas-Gomez, Jannik Laval.

Reengineering large applications implies an underlying tool infrastructure that can scale and also be extended.

**Ring: a Unifying Meta-Model and Infrastructure for Smalltalk Source Code Analysis Tools.** Source code management systems record different versions of code. Tool support can then compute deltas between versions. To ease version history analysis we need adequate models to represent source code entities. As a first step to provide an infrastructure to support history analysis, this article [12] presents Ring, a unifying source code meta-model that can be used to support several activities and proposes a unified and layered approach to be the foundation for building an infrastructure for version and stream of change analyses. We re-implemented three tools based on Ring to show that it can be used as the underlying meta-model for remote and off-image browsing, scoping refactoring, and visualizing and analyzing changes. As a future work and based on Ring we will build a new generation of history analysis tools.

**AspectMaps: A Scalable Visualization of Join Point Shadows.** When using Aspect-Oriented Programming, it is sometimes difficult to determine at which join point an aspect executes. Similarly, when considering one join point, knowing which aspects will execute there and in what order is non-trivial. This makes it difficult to understand how the application will behave. A number of visualizations have been proposed that attempt to provide support for such program understanding. However, they neither scale up to large code bases nor scale down to understanding what happens at a single join point. In this paper [18], we present AspectMaps - a visualization that scales in both directions, thanks to a multi-level selective structural zoom. We show how the use of AspectMaps allows for program understanding of code with aspects, revealing both a wealth of information of what can happen at one particular join point as well as allowing to see the “big picture” on a larger code base. We demonstrate the usefulness of AspectMaps on an example and present the results of a small user study that shows that AspectMaps outperforms other aspect visualization tools.

**Challenges to support automated random testing for dynamically typed languages.** Automated random testing is a proved way to identify bugs and precondition violations, and this even in well tested libraries. In the context of statically typed languages, current automated random testing tools heavily take advantage of static method declaration (argument types, thrown exceptions) to constrain input domains while testing and to identify errors. For such reason, automated random testing has not been investigated in the context of dynamically typed languages. We present the key challenges that have to be addressed to support automated testing in dynamic languages. [17]

**SmartGroups, Focusing on Task-Relevant Source Artifacts in IDEs.** Navigating large software systems, even when using a modern IDE (Integrated Development Environment) is difficult, since conceptually related software artifacts are distributed in a huge software space. For most software maintenance tasks, only a small fraction of the entire software space is actually relevant. The IDE, however, does not reveal the task relevancy of source artifacts, thus developers cannot easily focus on the artifacts required to accomplish their tasks. Smart Groups help developers to perform software maintenance tasks by representing groups of source artifacts that are relevant for the current task. Relevancy is determined by analyzing historical navigation and modification activities, evolutionary information, and runtime information. The prediction quality of Smart Groups is validated with a benchmark evaluation using recorded development activities and evolutionary information from versioning systems. [24]

5.3. Constructs for Dynamic Languages

**Participants:** Stéphane Ducasse, Marcus Denker, Veronica Uquillas-Gomez, Gwenael Casaccio, Camillo Bruni, Jean-Baptiste Arnaud, Damien Pollet.

To support our research on secure dynamic languages, we focused on improving language infrastructure.
**Efficient Proxies in Smalltalk.** A proxy object is a surrogate or placeholder that controls access to another target object. Proxy objects are a widely used solution for different scenarios such as remote method invocation, future objects, behavioral reflection, object databases, inter-languages communications and bindings, access control, lazy or parallel evaluation, security, among others. Most proxy implementations support proxies for regular objects but they are unable to create proxies for classes or methods. Proxies can be complex to install, have a significant overhead, be limited to certain type of classes, etc. Moreover, most proxy implementations are not stratified at all and there is no separation between proxies and handlers. We present Ghost, a uniform, light-weight and stratified general purpose proxy model and its Smalltalk implementation. Ghost supports proxies for classes or methods. When a proxy takes the place of a class it intercepts both, messages received by the class and lookup of methods for messages received by instances. Similarly, if a proxy takes the place of a method, then the method execution is intercepted too. [22]

**Bootstrapping a Smalltalk.** Smalltalk is a reflective system. It means that it is defined in itself in a causally connected way. Traditionally, Smalltalk systems evolved by modifying and cloning what is called an image (a chunk of memory containing all the objects at a given point in time). During the evolution of the system, objects representing it are modified. However, such an image modification and cloning poses several problems: (1) There is no operational machine-executable algorithm that allows one to build a system from scratch. A system object may be modified but it may be difficult to reproduce its exact state before the changes. Therefore it is difficult to get a reproducible process. (2) As a consequence, certain classes may not have been initialized since years. (3) Finally, since the system acts as a living system, it is not simple to evolve the kernel for introducing new abstractions without performing some kind of brain surgery on oneself. There is a need to have a step by step process to build Smalltalk kernels from scratch. After an analysis of past and current practices to mutate or generate kernels, we describe a kernel bootstrap process step-by-step. First the illusion of the existence of a kernel is created via stubs objects. Second the classes and meta-classes hierarchy are generated. Code is compiled and finally information needed by the virtual machine and execution are generated and installed. [15]

**Flexible Object Layouts: Enabling Lightweight Language Extensions by Intercepting Slot Access.** Programming idioms, design patterns and application libraries often introduce cumbersome and repetitive boilerplate code to a software system. Language extensions and external DSLs (domain specific languages) are sometimes introduced to reduce the need for boilerplate code, but they also complicate the system by introducing the need for language dialects and inter-language mediation. To address this, we propose to extend the structural reflective model of the language with object layouts, layout scopes and slots. Based on the new reflective language model we can 1) provide behavioral hooks to object layouts that are triggered when the fields of an object are accessed and 2) simplify the implementation of state-related language extensions such as stateful traits. By doing this we show how many idiomatic use cases that normally require boilerplate code can be more effectively supported. We present an implementation in Smalltalk, and illustrate its usage through a series of extended examples. [25]

### 5.4. Resources

**Participants:** Stéphane Ducasse, Marcus Denker, Mariano Martinez-Peck, Nick Papoylias.

Resource management is important in the context of resource constrained devices as well as situations were a large amount of data is modeled but not accessed often. One example for resource constrained devices are autonomous robots. An example for large models that are accessed infrequently are typical models of systems for software re-engineering.

With Ecole des Mines de Douai we explore how to analyze and improve memory in the case of unused data.

**Problems and Challenges when Building a Manager for Unused Objects.** Large object-oriented applications may occupy hundreds of megabytes or even gigabytes of memory. During program execution, a large graph of objects is created and constantly changed. Most object runtimes support some kind of automatic memory management based on garbage collectors (GC) whose idea is the automatic destruction of unreferenced objects. However, there are referenced objects which are not used for a long period of time or that are
used just once. These are not garbage-collected because they are still reachable and might be used in the future. Due to these unused objects, applications use much more resources than they actually need. We present the challenges and possible approaches towards an unused object manager for Pharo. The goal is to use less memory by swapping out the unused objects to secondary memory and leaving in primary memory only those objects that are needed and used. When one of the unused objects is needed, it is brought back into primary memory. [23]

Clustered Serialization with Fuel. Serializing object graphs is an important activity since objects should be stored and reloaded on different environments. There is a plethora of frameworks to serialize objects based on recursive parsing of the object graphs. However such approaches are often too slow. Most approaches are limited in their provided features. For example, several serializers do not support class shape changes, global references, transient references or hooks to execute something before or after being stored or loaded. Moreover, to be faster, some serializers are not written taking into account the object-oriented paradigm and they are sometimes even implemented in the Virtual Machine hampering code portability. VM-based serializers such as ImageSegment are difficult to understand, maintain, and fix. For the final user, it means a serializer which is difficult to customize, adapt or extend to his own needs. We present a general purpose object graph serializer based on a pickling format and algorithm. We implement and validate this approach in the Pharo Smalltalk environment. We demonstrate that we can build a really fast serializer without specific VM support, with a clean object-oriented design, and providing most possible required features for a serializer. We show that our approach is faster that traditional serializers and compare favorably with ImageSegment as soon as serialized objects are not in isolation. [16]

Towards Structural Decomposition of Reflection with Mirrors Mirrors are meta-level entities introduced to decouple reflection from the base-level system. Current mirror-based systems focus on functional decomposition of reflection. We advocate that mirrors should also address structural decomposition. Mirrors should not only be the entry points of reflective behavior but also be the storage entities of meta-information. This decomposition can help resolve issues in terms of resource constraints (e.g. embedded systems and robotics) or security. Indeed, structural decomposition enables discarding meta-information. [21]

5.5. Empirical Studies in Software Product Line Engineering

Participant: Nicolas Anquetil.

Software Product Line Engineering (SPLE) is a new development paradigm that promises to offer faster development, with better quality. The idea is to develop a generic application (the Software Product Line) with pre-defined variation points. From this, new applications are derived from the generic application and the options are chosen for the possible variation points. Because it is still a new paradigm, Software Product Line development is still an active research domain where empirical research is useful to check the validity of the results.

These publications are the results of an earlier research project to which N. Anquetil participated.

Managing information flow in the SPL development processes. Traceability is a quality attribute in software engineering that establishes the ability to describe and follow the life of a requirement in both the forward and backward directions (i.e. from its origins throughout its specification, implementation, deployment, use and maintenance, and vice-versa). The IEEE Standard Glossary of Software Engineering Terminology defines traceability as “the degree to which a relationship can be established between two or more products of the development process, especially products having a predecessor-successor or master-subordinate relationship to one another”. According to (Palmer, 1997) “traceability gives essential assistance in understanding the relationships that exist within and across software requirements, design, and implementation”. Thus, trace relationships help in identifying the origin and rationale for artefacts generated during development lifecycle and the links between these artefacts. Identification of sources helps understanding requirements evolution and validating implementation of stakeholders’ requirements. The main advantages of traceability are: (i) to relate software artefacts and design decisions taken during the software development cycle; (ii) to give feedback to architects and designers about the current state of the development, allowing them to reconsider alternative
design decisions, and to track and understand bugs; and (iii) to ease communication between stakeholders. [26]

Empirical research in software product line engineering. Empirical evaluation has for many years been utilized to validate theories in other science disciplines. One of the first well-known reported examples of empirical evaluation occurred when Galileo wanted to prove that the rate of descent of objects was independent of their mass. This would disprove a theory put forward by Aristotle that the rate of descent is directly proportional to their weight. To prove his theory Galileo dropped two balls made from the same material but different masses from the top of the Tower of Pisa. When the experiment was performed Galileo’s theory was proved correct through the empirical evidence collected. What this story demonstrates is the importance of empirical validation to verify or disprove theories and hypotheses. The purpose of this publication [27] is to emphasize the importance and difficulties of empirical evaluation in the domain of SPLE.
6. New Results

6.1. High-Performance Intra-node Collective Operations

**Participants:** Brice Goglin, Stéphanie Moreaud.

- KNEM is known to improve the performance of point-to-point intra-node MPI communication significantly [60], [18].
- We designed an extended RMA interface in KNEM that suits the needs of point-to-point, collective and RMA operations.
- We showed that the native use of KNEM in MPI collective implementations enabled further optimization by combining the knowledge of collective algorithms with the mastering of KNEM region management and copies [35].
- This work was initiated in the context of our collaboration with the MPICH2 team and is now also pursued within the OPEN MPI project in collaboration with the University of Tennessee in Knoxville.

6.2. I/O-Affinity-aware MPI Communications

**Participants:** Brice Goglin, Stéphanie Moreaud.

- We demonstrated in the past that the locality of I/O devices within modern computing nodes has the significant impact of the MPI communication performance [11] (Non-Uniform I/O Access, NUIOA).
- A first way to deal with such affinities would be to privilege I/O-intensive processes by placing them near the network interfaces. However, determining the communication-intensiveness may be tricky. Also, some applications have uniform communication patterns. The other way to deal with I/O affinities is to modify the implementation of communication operations given a predetermined task placement.
- We demonstrated that the implementation of collective operations should take I/O affinities into account. Deciding which steps and leaders should be involved in the algorithms based on NUIOA effects led us to improve broadcast performance significantly [34], [18].

6.3. High-Performance Point-to-Point Communications

**Participants:** Alexandre Denis, Raymond Namyst.

- NEWMADELEINE is our communication library designed for high performance networks in clusters. We have worked on optimizations on low-level protocols so as to improve point-to-point performance.
- We have proposed [29] auto-tuning mechanisms for most parameters of a communication library: rendez-vous threshold, multi-rail ratio, optimization strategies.
- We have proposed a communication protocol [33] for InfiniBand that completely amortizes the cost of memory registration, through the use of a superpipeline that overlaps communication and memory copies. We have modeled the behavior of the network and proposed auto-tuning mechanism to adapt the protocol to the hardware properties.
6.4. Improve code-coupling performance in the SALOME platform

Participants: Alexandre Denis, Sébastien Barascou.

- SALOME platform is an open source software developed by EDF, CEA, and OpenCascade. It is an open simulation platform with pre-processing, post-processing, interoperability with CAD models, integration with computation kernels.
- YACS is the workflow engine used for code coupling applications in SALOME. It leverages CORBA for communications between kernels. We have ported [50] YACS atop PadicoTM, our communication platform for grids. It enables CORBA connections to use InfiniBand networks. Benchmarks show a significant improvement in code coupling performance.

6.5. Hardware topology-aware MPI applications

Participants: Emmanuel Jeannot, Guillaume Mercier, François Tessier.

- We have expanded our previous work dealing with MPI process placement. Indeed, our approach relied on tools and techniques which were outside the scope of the MPI standard itself. In order to allow the users to utilize our work in a portable way, we enhanced some routines of the MPI standard. We worked mainly with the MPICH2 implementation but we are also working on an OPEN MPI version as well.
- Instead of modifying the binding of the MPI processes onto the physical cores on the underlying architecture, we chose to create a new communicator for which the logical topology organization is optimized for the hardware. This work has been published in [37] and show interesting performance improvements for some class of MPI applications.
- The problem of process placement, which can be reduced to an NP-hard graph partitioning problem, can be dealt with several famous applications like Scotch or ParMETIS. To evaluate these solutions with TREEMATCH, we ran several benchmarks using NAS Parallel Benchmarks and a real CFD application. On the one hand we study the quality of processes permutation (which will impact the execution time) and on the other hand the computation time of the permutation. These results will allow us to conclude about the pertinence of what graph partitioner can be used to bind processes on process units or to do a dynamic processes reordering.

6.6. Mastering Heterogeneous Platforms

Participants: Cédric Augonnet, Olivier Aumage, Ludovic Courtès, Nathalie Furmento, Andra Hugo, Raymond Namyst, Samuel Thibault, Pierre-André Wacrenier.

- We continued our work on extending STARPU to master exploitation of Heterogeneous Platforms.
- We have extended the STARPU scheduler into managing parallel tasks which permit a better exploitation of CPUs and load balancing with GPUs.
- We have designed over STARPU a lightweight DSM over MPI, which permits to seamlessly execute STARPU applications over an MPI cluster of GPU-enhanced nodes.
- We have been developing a GCC plug-in which extends the C language with pragmas and attributes that make writing STARPU applications a lot easier.
- We have brought to STARPU support for automatically converting data between CPU and GPU formats (typically arrays of structures vs. structures of arrays). We are now optimizing it.
- We have added an OpenCL interface to STARPU, SOCL [42], which permits to execute unmodified OpenCL applications over STARPU.
- We have introduced in STARPU theoretical bound support [27]: from a record of the set of tasks submitted by the application, STARPU uses linear programming to give the execution time of an ideal scheduling, which can then be compared with the actual results.
- We have continued collaboration with the University of Tennessee, Knoxville for STARPU support in the state-of-the-art dense linear algebra library, Magma, in particular LU [26] and QR [27] factorizations. We have also collaborated with the University of Mons [41] and Linköping [32].
- Cédric Augonnet defended his PhD on STARPU [17].
6.7. Development of a flexible heterogeneous system-on-chip platform using a mix of programmable processing elements and hardware accelerators

Participants: Paul-Antoine Arras, Emmanuel Jeannot, Samuel Thibault.

- Today’s embedded applications are increasingly demanding in terms of computational power, especially in real-time digital signal processing (DSP) where tight timing requirements are to be fulfilled. More specifically, when it comes to video decoding (e.g. H.264/AVC) not only has it been almost impossible for some time to run such codecs on a stand-alone embedded processor, but it now also becomes quite impractical to execute them on homogeneous multicore platforms. In this context, STMicroelectronics is developing a scalable heterogeneous system-on-chip template called P2012 and aimed at meeting the latest codecs’ requirements.

- This year, the privileged axis of research was directed towards dataflow-based models, which benefit from such strong, well-known properties as analyzability, schedulability and expressivity. Furthermore, dataflow programming has already been used extensively in DSP, yielding a number of dedicated software synthesis tools. We have proposed a first version of the programming model that will be evaluated later.

6.8. Sparse GMRES on heterogeneous platforms in oil extraction simulation

Participants: Olivier Aumage, Corentin Rossignon, Samuel Thibault.

- We started a study on sparse matrix factorization and system resolution on heterogeneous platforms in collaboration with Pascal Hénon from company Total, in the context of oil extraction simulation. Sparse matrix computations are notoriously difficult to efficiently run on heterogeneous platforms in the general case due to the irregular memory access patterns they generate.

- However, in the specific context of this study, Corentin Rossignon showed as part of his Master Thesis [56] that the sparsity layout of matrices generated by such oil extraction simulation problems can lead to a much higher level of efficiency on heterogeneous platforms thanks when using a suitable sparse internal representation together with carefully written operators such as the sparse matrix-vector product together with the StarPU heterogeneous scheduler.

- Corentin Rossignon is now starting a PhD. Thesis in partnership with Total to build on these promising results.

6.9. Programming models for heterogeneous platforms

Participants: Olivier Aumage, Cyril Roelandt, Samuel Thibault, Ludovic Courtès.

- As part of Project FP3C with Japan, we started a study on to explore the use of StarPU as possible target runtime system for the XcalableMP language and compiler developed by Prof. Sato’s team from University of Tsukuba. XcalableMP is a pragma-based language designed for parallelising application on clusters of multicore processors. The compiler is responsible to expand XcalableMP pragma into complex work mapping, communication and data redistribution commands.

- The study of porting XcalableMP on top of StarPU was conducted by Cyril Roelandt during his Master Thesis [55], starting from the idea that computing node with one or more attached accelerating expansion cards can be seen as a distributed platform. The results of the study showed that on the one side, the power of the XcalableMP language itself is very interesting for the goal of simplifying the port of applications on heterogeneous platforms. However, a current assumption of the XcalableMP model is that the compiler does not insert implicit commands and behaviour except at the exact location of pragma annotations, which limit the range of optimizations available to the dynamic scheduler and memory manager of StarPU. We will thus continue to collaborate with Prof. Sato’s team within the FP3C to see how these limitations could be reduced or lifted when using XcalableMP with StarPU.
• In an effort to make it easier for C programmers to benefit from StarPU, the team-project has been working on extensions to the C language allowing important StarPU concepts to be expressed concisely. These C extensions are provided as a plug-in for the GNU Compiler Collection (GCC\(^1\)), and is now distributed as part of StarPU.

The GCC plug-in extends the syntax and semantics of C and related languages (C++, Objective-C) using attributes and pragmas. Attributes are used, for instance, to declare StarPU tasks and their implementations for the available targets (CPU, OpenCL, CUDA, etc.) Pragmas are used notably to provide programmers a way to describe data buffers that are passed to tasks, which in turn allows the StarPU run-time support to manage data transfers between main memory and GPUs as it sees fit. Finally, tasks are invoked like regular C functions.

In addition to easing application development, the GCC plug-in, thanks to its higher-level view of the program structure, is able to report certain classes of errors at compile-time, which would otherwise lead to run-time errors.

This project has been led by Ludovic Courtès of Inria’s Development and Experimentation Department (SED) at Bordeaux, as part of a joint development action with the SED.

6.10. Parallel Concha

Participants: Olivier Aumage, Marie-Christine Counilh.

• Within the ADT Ampli project, we contributed to the Concha CFD library developed by R. Becker’s Inria Team Concha in Pau. Together with R. Becker, E. Bergounioux and D. Trujillo from Concha Team, and François Rue from SED Bordeaux we designed and experimented with the MPI parallelization and the hybrid MPI+OpenMP parallelization of the library.

• The MPI parallelization is now finalized. The OpenMP level has been successfully tested on the Vanka smoother and is now being spread in the library. We will thus continue to contribute to this parallelization work, in particular with respect to the support of 3D simulation cases.

6.11. Scientific Application Analysis and Experiments

Participants: Olivier Aumage, Denis Barthou, Andres Charif-Rubial, François Tessier, Ludovic Stordeur.

• Within the context of the ANR ProHMPT project, we contributed a thorough analysis of hot spots, data structure usages and locality issues in memory accesses of an aerodynamics application from partner CEA CESTA.

• In accordance with these results, a new version of this application has been written by the CESTA Team with redesigned, locality-friendly data structures and simplified loop scheme. This new version performs much better than the previous one on both 2D and 3D cases.

• We also conducted tests about the port of selected kernels of this application on accelerated heterogeneous platforms. The results of these tests were disappoiting with the first version of the application due to the layout of the main data structures that led to a lot of memory transfers between the central memory and the accelerated memory.

• We are now working on conducting these experiments with the redesigned version of the application whose new data structures should dramatically reduce the amount of data transfers.

\(^1\)See [http://gcc.gnu.org/](http://gcc.gnu.org/), for more information on GCC.
6.12. Virtualization of GPUs for OpenCL

**Participants:** Sylvain Henry, Alexandre Denis, Denis Barthou.

- We propose a new approach for OpenCL programming, using a unique virtual accelerator instead of using the physical accelerator. Placement on the real hardware is handled by the runtime instead of the user, improving productivity and performance scalability. This proposition relies on OpenCL standard but changes the way its API is used.
- We have shown on some simple examples how this approach, using StarPU as a runtime, enables executions with a better load balance and performance. We are working on how to generalize this to more complex benchmarks. This work has been presented in Renpar[ 42 ] workshop.

6.13. Automatically Adapting Task Grain for Hybrid Architectures

**Participants:** Sylvain Henry, Alexandre Denis, Denis Barthou.

- Given a parallel task graph, a runtime such as StarPU can place each task on different hardware. However, there is still the need to adapt the number of tasks, the granularity of these tasks, according to the target hardware. Due to architectures with CPUs and GPUs, it is potentially interesting to have tasks of different granularities. We explore transformations that enable to either automatically split tasks into small ones, or given some user knowledge on the tasks, decide how and when to split a large task into small ones.
- This work starts from a high-level representation of the code, using an explicit data-flow graph.

6.14. Performance modeling for power consumption reduction on the SCC

**Participants:** Bertrand Putigny, Brice Goglin, Denis Barthou.

- We build a model to predict performance of HPC code on the SCC ship. This model can predict runtime of regular code as well as power consumption for different frequency.
- This allows users to choose either to optimize power consumption, power efficiency or raw performance.
- This work has been published in an Intel Symposium [ 38 ].

6.15. Modeling cache coherence protocol overhead

**Participants:** Bertrand Putigny, Denis Barthou, Brice Goglin.

- We are building a fine grained cache model to understand common cache coherence issue.
- This model is built on a set of micro-benchmarks and can also be used to improve find some bottlenecks in memory bound code. Our set of micro-benchmarks can also be used as a test bed for new architectures [ 54 ].

6.16. Memory Performance Analysis and Tool for OpenMP codes

**Participants:** Andres Charif-Rubial, Denis Barthou.

- We propose a performance analysis of OpenMP codes, based on memory accesses and cache hierarchies.
- This analysis relies on memory traces for multi-threaded codes and on static analysis of binary code. Memory traces are obtained through MAQAO by static binary rewriting and are compressed online, building polyhedral iteration domains. The static analysis, mostly induction variable detection on binary code, provides the same analysis whenever possible, removing the need in some cases for dynamic instrumentation.
- The analysis focuses on a number of issues in multi-threaded executions: thread affinity issues, false sharing, cache pollution.
- This work is in collaboration with Exascale Computing Lab.
6.17. Data-layout Optimization for Stencil codes on multi-cores and GPUs

Participants: Julien Jaeger, Denis Barthou.

- We develop a new approach for stencil code generation, optimizing data-layout for multi-threaded, SIMD code on multicores and CUDA code on GPU. The transformation handles different stencil parameters, and memory constraints.
- The code generated reaches high levels of performance, outperforming related works for multicores and with similar performance on GPUs. This work is submitted to publication and was first presented in a workshop [52].
5. New Results

5.1. Languages and Foundations: Process algebra

Participants: Damien Pous, Alan Schmitt, Jean-Bernard Stefani, Claudio Mezzina, Cinzia di Giusto.

The goal of this work is to study process algebraic foundations for component-based distributed programming. Most of this work takes place in the context of the ANR PiCoq project.

To develop composable abstractions for programming dependable systems, we investigate concurrent reversible models of computation, where arbitrary executions can be reversed, step by step, in a causally consistent way. This year we have continued the study of the reversible higher-order pi-calculus and obtained a new encoding of it in the higher-order pi-calculus which improves on the result we published in Concur 2010 by proving the faithfulness of the encoding with a much finer equivalence relation. We also developed a reversible variant of the higher-order pi-calculus where reversibility can be controlled by means of an explicit rollback primitive [37]. We have proved that this rollback primitive is sound and complete in that it provides a causally consistent and complete reversal of concurrent computations, and we have developed a low-level semantics for this primitive, closer to an actual distributed implementation, which we have proved equivalent to the high-level one. All these results are presented in detail in Claudio Mezzina’s forthcoming PhD thesis, and have been developed in cooperation with the INRIA Focus team at the University of Bologna.

An interesting and expressive component model for embedded systems is the BIP component model [58], developed by J. Sifakis’ team at the Verimag Laboratory, which features hierarchical software architectures, explicit constructs for specifying component compositions (glues), and multipoint synchronization under priority constraints. We have begun a process calculus analysis of BIP, with a view to combine the reactive features of BIP with the dynamic reconfiguration features of Fractal. Our first result takes the form of new process calculus, called CAB, which we have proved to be a conservative extension of BIP. CAB also enabled us to study the intrinsic expressivity of the BIP model and to prove that priority constraints are essential to BIP expressivity [34].

We have made significant progress on the formalization in the Coq proof assistant of a core higher-order π-calculus, called HOcore [20]. We have in particular adapted a canonical locally nameless representation of binding to handle alpha-conversion in our formalization. Several major theorems of HOcore, in particular the fact that IO-bisimulation is correct in relation to barbed congruence and is decidable. This work has been submitted for publication.

A longer version of our work on untyping theorems and cyclic linear logic has been accepted for publication in LMCS [24], and a book chapter on up-to techniques for bisimulations, written with Davide Sangiorgi from the INRIA Focus team in Bologna, has been published by Cambridge University Press [46].

Together with Filippo Bonchi (LIP, ENS Lyon), we have worked on a new algorithm for checking the language equivalence of non-deterministic finite automata (NFA). This algorithm improves on the standard Hopcroft and Karp’ algorithm, by using up-to techniques. The first empirical tests look really promising [47].

Together with Tom Hirschowitz (LAMA, U. de Chambéry), we have worked on a categorical model of CCS, where innocent strategies are pre-sheaves. This work has been presented at the ICE workshop [36], and a long version has been submitted to SACS.

5.2. Languages and Foundations: Proof tactics

Participants: Damien Pous, Thomas Braibant.
The goal of this work is to develop proof-assistant-based tools for verifying distributed systems and distributed abstract machines. In particular, we aim to support the derivation of fully formal proofs of correctness for abstract machines supporting the component-based languages and programming models we develop.

We have presented our work about tools for rewriting modulo AC in Coq at CPP’11 [32]. An extended version of our work on Kleene algebra (ATBR, first published at ITP’10), was accepted for publication in LMCS [19]. Also on the Coq side, we have developed a library for verifying hardware circuits, which was also presented at CPP’11 [31].

5.3. Control for adaptive systems: Discrete control for adaptive and reconfigurable systems

Participants: Eric Rutten, Noël de Palma, Olivier Gruber, Fabienne Boyer, Tayeb Bouhadiba, Xin An.

The goal of this work is to apply control techniques based on the behavioral model of reactive automata and the algorithmic techniques of discrete controller synthesis. We adopt the synchronous approach to reactive systems, and use an associated effective controller synthesis tool, Sigali, developed at INRIA Rennes. Both are integrated into a programming language, called BZR, and its compiler, as an extension of the Heptagon language. We thus have a complete tool-supported method from control modeling down to concrete execution, considering different execution models, and targetting either software or hardware. We explore control theory for computer science, as an original alternative to computer science for control (as more usually in embedded systems), and to classical discrete control systems (as more usually applied to manufacturing). We are exploring several target application domains, where we expect to find commonalities in the control problems, and variations in the definitions of configurations, and in the criteria motivating adaptation.

We have obtained this year the following results:

- At the programming language level, we are continuing the development of a modelling and controller generation language called BZR, which involves DCS in its compilation. BZR is designed and developed in cooperation with the Pop Art and VerTecs (INRIA Rennes) teams [40].
- We have developed a technique for designing reconfiguration controllers in the Fractal component-based framework, where discrete control loops automatically enforce safety properties on the interactions between components, concerning, e.g., mutual exclusions, forbidden or imposed sequences [29] [48].
- We have integrated BZR with Orcad, a programming environment for real-time control systems, in cooperation with the NeCS and SED teams at INRIA Grenoble [28].
- We are investigating administration loops in virtual machine-based distributed systems [44], and the coordination of such loops, especially in relation with green computing problems. We are starting a new ANR project called Ctrl-Green on this topic in 2012.
- We work on the formal modelling and control of dynamic reconfiguration in FPGA circuits, in cooperation with the DaRT team (INRIA Lille) [43] and the Lab-STICC laboratory in Lorient [42], building upon earlier work related to the MARTE framework.
- In cooperation with GIPSA-Lab and ENSI Tunis, we have adapted the discrete controller synthesis technique to the control of decentralized systems that are composed of several subsystems spread across remote sites [17].
- In cooperation with Orange labs and GIPSA-Lab we are beginning to explore the application of discrete event systems and supervisory control to the domain of Machine to Machine and Internet of Things, with the objective to manage energy aspects; this will start with the CIFRE PhD (U. Grenoble) of Mengxuan Zhao (co-advised with H. Alla, G. Privat).

5.4. System configuration and deployment

Participants: Loris Bouzonnet, Fabienne Boyer, Willy Malvault, Noël de Palma, Vivien Quéma, Jean-Bernard Stefani.
The goal of this work is to study system configuration and software deployment issues in large distributed systems.

System configuration and software deployment in a distributed environment can be greatly aided by the use of a uniform component model to support software assembly, software configuration and deployment, as well as runtime system configuration. We have developed a specialization of the Fractal component model that provides a reference model for heterogeneous software assembly and configuration. In particular, we have shown how this reference model can be used to assemble and configure software architectures built from heterogeneous software packages (e.g. OSGI bundles for Java packages, Debian or RPM packages for Linux modules and applications). The definition of this model, a description of its implementation and its evaluation are documented in Loris Bouzonnet’s PhD thesis [13].

As an alternative to current public cloud infrastructures, which rely on large data centers, we have started the study of a cloud infrastructure based on a peer-to-peer (P2P) overlay network built on gossip-based protocols. More precisely, we have studied how to implement a distributed resource allocation service in a P2P environment maintained by a gossip-based peer-sampling protocol [81]. The resulting system, called Salute, provides for the allocation of application-specific overlays out of an underlying P2P network. By combining several P2P services (including peer-sampling, topology maintenance, and node synchronization), and by partitioning available nodes into free nodes (available for the allocation of new application overlays) and reserve nodes (nodes dedicated to the maintenance of allocated overlays), Salute provides a churn-resilient, completely decentralized cloud infrastructure. In addition, we have shown that Salute can provide its allocation service while maintaining fairness and avoiding starvation. The Salute architecture has been validated through simulations using network traces from different real-world P2P environments. The Salute architecture, algorithms and their validation are documented in Willy Malvault’s PhD thesis [14].

In a cloud computing context the complexity of deploying and configuring non-trivial software architectures is exacerbated. In line with our previous work on architecture-based distributed system management, we have proposed a novel algorithm for configuring component-based distributed applications deployed within several virtual machines in an IaaS environment. The algorithm is completely decentralized, relies on a message queuing middleware and exploits the software architecture descriptions of the applications to deploy and configure, written in an extension of the Fractal Architecture Description Language. A first version of this algorithm, that does not take into account potential failures during the configuration process, has been formally specified in collaboration with Gwen Salaün from the INRIA Vasy team in Grenoble, and presented at IEEE Cloud 2011 [35].

5.5. System support: System support for multicore machines

Participants: Vivien Quéma, Renaud Lachaize, Fabien Gaud, Baptiste Lepers, Sylvain Genevès, Fabien Mottet.

Multicore machines with Non-Uniform Memory Accesses (NUMA) are becoming commodity platforms. Efficiently exploiting their resources remains an open research problem. Most of the body of existing work focuses on increasing locality between computations and memory or I/O resources. This is achieved by allocating data items preferably in local memory nodes, by moving computations close to I/O devices or by moving already allocated memory pages close to the applications which use them most. In all these works, researchers always assume that all processors have equal memory performance. Nevertheless, this assumption is not always valid. In 2011, we have studied the performance achieved by a 16-core NUMA exhibiting an irregular connectivity between processors. Some processors are directly connected to all other processors and access memory nodes with a low latency. Other processors have a lower degree of connectivity and need more hops to access some memory nodes and access memory with a higher latency.

Current operating systems are not aware of such performance characteristics. We have shown that the completion time of applications taken from the PARSEC benchmark suite can vary by up to 15% depending on the processor they are scheduled on. We have thus proposed a new OS scheduler that takes this asymmetry into account in order to make efficient decisions. This scheduler relies on a new metric, called MAPI (number
of main Memory Accesses Per retired Instruction), to predict the impact of processor interconnect asymmetry on the performance of applications. We have empirically evaluated the relevance of this metric on applications taken from the PARSEC benchmark suite. We have shown that this metric helps estimating the performance gap between running an application on a "well-interconnected" processor and on a "weakly-interconnected" one. Using this metric, the scheduler we proposed makes efficient decisions. More precisely, we have observed that the scheduler always performed within 3% of the best possible scheduling decision. This work is currently under submission.

5.6. System support: Protocols for resilient systems

Participants: Vivien Quéma, Alessio Pace.

We have worked on replication protocols for P2P systems. In particular, we have worked on replication in so called Distributed Hash-Tables (DHTs). DHTs provide a simple high-level put/get abstraction that can be used to build efficient distributed storage systems. DHTs gained wide popularity in the last decade, fostering a large amount of interest in the academia, and inspiring the design of key/value distributed storage systems deployed in production.

DHTs provide a way to deterministically map objects to nodes and allow efficiently retrieving objects in a distributed fashion. Nodes and objects are logically arranged in a large numeric key-space, according to a given variant of consistent hashing. Typically, the node in charge of an object is the one whose position immediately follows the object in the key-space.

To guarantee that objects are reliably stored, DHTs rely on replication. A replication protocol is in charge of ensuring that, at any time, each object is replicated on a sufficiently large number of replicas. Several replication strategies have been proposed in the last years. The most efficient ones use predictions about the availability of nodes to reduce the number of object migrations that need to be performed: objects are preferably stored on highly available nodes.

We have proposed an alternative replication strategy. Rather than exploiting highly available nodes, we have designed a protocol that leverages nodes that exhibit regularity in their connection pattern. Roughly speaking, the strategy consists in replicating each object on a set of nodes that is built in such a way that, with high probability, at any time, there are always at least k nodes in the set that are available. We have evaluated this new replication strategy using traces of two real-world systems: eDonkey and Skype. Our evaluation showed that our regularity-based replication strategy induces a systematically lower network usage than existing state of the art replication strategies. This work has been published at the International Symposium on Reliable Distributed Systems, in October 2011.

5.7. System support: End-to-end caching

Participants: Sara Bouchenak, Dàmian Serrano.

Cloud Computing is a paradigm for enabling remote, on-demand access to a virtually infinite set of configurable computing resources. This model aims to provide hardware and software services to customers, while minimizing human efforts in terms of service installation, configuration and maintenance, for both cloud provider and cloud customer. A cloud may have the form of an Infrastructure as a Service (IaaS), a Platform as a Service (PaaS) or a Software as a Service (SaaS). Clouds pose significant challenges to the full elasticity of clouds, their scalability and their dependability in large scale data management and large scale computing resources. Caching is a means for high performance and scalability of distributed systems. Although caching solutions have been successfully studied for individual systems such as database systems or web servers, if collectively applied, these solutions violate the coherence of cached data. We precisely studied this issue in e-Caching, a novel end-to-end caching system.

The contribution of this work is twofold: guaranteeing the coherence of cached data at multiple locations of a distributed system, while improving the overall performance of the system. In collaboration with Marta Patino and Ricardo Jimenez from Universidad Politecnica de Madrid, we proposed a novel distributed caching protocol, implemented it and evaluated it with real online services. The experiments showed that e-Caching was successfully able to improve service performance by two orders of magnitude.
This work has been presented at CFSE, the French Chapter of ACM-SIGOPS in May 2011. An extended version has been submitted for publication in a journal.

5.8. System support: Performance and dependability benchmarking

Participants: Amit Sangroya, Dàmian Serrano, Sara Bouchenak [correspondant].

MapReduce has become a popular programming model and runtime environment for developing and executing distributed data-intensive and compute-intensive applications. It offers developers a means to transparently handle data partitioning, replication, task scheduling and fault tolerance on a cluster of commodity computers. MapReduce allows a wide range of applications such as log analysis, data mining, Web search engines, scientific computing, bioinformatics, decision support and business intelligence.

There has been a large amount of work on MapReduce towards improving its performance and reliability. Several efforts have explored task scheduling policies in MapReduce, cost-based optimization techniques, replication and partitioning policies. There has also been a considerable interest in extending MapReduce with other fault tolerance models, or with techniques from database systems. However, there has been very little in the way of empiric evaluation for the comparison of the different systems. Most evaluations of these systems have relied on microbenchmarks based on simple MapReduce programs. While microbenchmarks may be useful in targeting specific system features, they are not representative of full distributed applications, and they do not provide multi-user realistic workloads. Furthermore, as far as we know, no studies have investigated dependability benchmarking of MapReduce.

Thus, we provide MapReduce Benchmarking (MRB), a novel MapReduce benchmark suite to enable a thorough analysis of a wide range of features of MapReduce systems. MRB has the following features. First, it enables the empirical evaluation of the performance and dependability of MapReduce systems. This provides a means to analyze the effectiveness of scalability and fault tolerance, two key features of MapReduce. Second, it covers a variety of application domains, workload and faultload characteristics, ranging from compute-oriented to data-oriented applications, batch applications to online real-time applications. While MapReduce frameworks were originally limited to offline batch processing, recent works are exploring the extension of MapReduce beyond batch processing. Moreover, in order to stress MapReduce dependability and performance, the benchmark suite enables different fault injection rates, workloads and concurrency levels. Finally, the benchmark suite is portable and easy to use on a wide range of platforms, covering different MapReduce frameworks and cloud infrastructures. This work has been submitted for publication.

5.9. System support: Self-adaptive Internet services

Participant: Sara Bouchenak.

Although distributed services provide a means for supporting scalable Internet applications, their ad-hoc provisioning and configuration pose a difficult tradeoff between service performance and availability. This is made harder as Internet service workloads tend to be heterogeneous, and vary over time in amount of concurrent clients and in mixture of client interactions. This work proposes an approach for building self-adaptive Internet services through utility-aware capacity planning and provisioning. First, an analytic model is presented to predict Internet service performance, availability and cost. Second, a utility function is defined and a utility-aware capacity planning method is proposed to calculate the optimal service configuration which guarantees SLA performance and availability objectives while minimizing functioning costs. Third, an adaptive control method is proposed to automatically apply the optimal configuration to the Internet service. Finally, the proposed model, capacity planning and control methods are implemented and applied to an online bookstore. Experimental evaluations show that the service successfully self-adapts to both workload mix and workload amount variations, and present significant benefits in terms of performance and availability, with a saving of resources underlying the Internet service.

This work is part of the MyCloud ANR project. It has been described in a chapter of the book titled Performance and Dependability in Service Computing, 2011. There has been an industrial transfer of the MOkA software prototype.
5.10. Self-Configuration of distributed system in the Cloud

Participants: Fabienne Boyer, Noël de Palma.

Cloud computing environments fall under three main kinds of offers according to the resources they provide. The Infrastructure as a Service (IaaS) level enables the access to virtualized hardware resources (processing, storage and network). The Software as a Service (SaaS) layer aims at providing the end-users with software applications. The intermediary layer, called Platform as a Service (PaaS), offers a set of tools and runtime environments that allow managing the applications life-cycle. This life-cycle includes the phases related to the design, the development, the deployment of applications, and generally speaking all their management stages (workload, fault tolerance, security). This article focuses on the deployment of distributed applications in virtualized environments such as cloud computing. Such deployments require to generate the virtual images that will be instantiated as virtual machines, thus ensuring the execution of the application on an IaaS platform. Each image embeds technical elements (operating system, middleware pieces) and functional ones (data and applicative software entities). Once it has been instantiated, each virtual machine is subjected to a stage of dynamic settings, which finalizes the global configuration of the distributed application.

On the whole, the deployment solutions currently available do not take into account these different configuration parameters, which are mostly managed by dedicated scripts. Moreover these solutions are not able to automate the images generation, their instantiation as virtual machines and their configuration independently from the kind of distributed application to be deployed. For instance, Google App Engine solution only deals with Web services organized into precisely defined tiers. In our opinion, the absence of general solutions results essentially from a lack of formalism for describing the distributed application architecture with its configuration constraints in a virtualized infrastructure such as cloud computing. Our work focused on a general solution, for Virtual Applications Management Platform, that automates the deployment of any distributed applications in the cloud. The suggested approach is architectural, meaning that it is based upon an explicit representation of the applications’ distributed architecture. We offer, on the one hand, a formalism for describing an application as a set of interconnected virtual machines and, on the other hand, an engine for interpreting this formalism and automating the application deployment on an IaaS platform. Specifically, we study three contributions:

- A formalism that offers a global view of the application to be deployed in terms of components with the associated configuration- and interconnection constraints and with their distribution within virtual machines. This formalism extends OVF language, dedicated to virtual machines description, with an architecture description language (ADL) that allows describing a distributed application software architecture;
- A deployment engine, i.e. a runtime support able to deploy automatically an application described with this formalism. This engine is based on a decentralized protocol for self-configuring the instantiated virtual machines. In our opinion it can ease the scalability of the dynamic configuration stage;
- A performance evaluation of the proposed solution on an industrial IaaS platform.

We published in this context two journal articles (TPDS [26] and TAAS [18]) and three conference papers (Cloud11 [35], UCC11 [39] and SAC12).

5.11. Virtual Machine

Participants: Olivier Gruber, Fabienne Boyer, Damien Pous, Ludovic Demontes, Clément Deschamps.

A core aspect of the Synergy virtual machine is its ability to reconfigure component-based applications at execution time. We have focused on the reconfiguration protocol with the intent of verifying and proving its robustness.
In a first step, we have formalized and verified that any correct and complex reconfiguration through our reconfiguration protocol can be processed as a sequence of elementary reconfiguration operations and always results in a component assembly that is architecturally consistent. This aspect has been verified using model-checking techniques. This work has been done in collaboration with Gwen Salaün from the VASY team (Inria Rhône-Alpes). It lead to a publication in the Formal Method (FM’11) conference [30].

In a second step, we have considered software failures that may occur during a reconfiguration. Although the protocol is trusted code, it invokes components to reconfigure them, thereby executing unsafe code that may fail. This work with Damien Pous produced a high-level formalisation of our reconfiguration protocol and a completely certified modelisation of these algorithms in Coq [50]. This work resulted in a submitted publication.

Finally, we have also investigated the control of complex reconfiguration through using discrete synchronous control techniques with Eric Rutten and Gwenael Delaval [44].
6. New Results


Processes have received a lot of attention in the last decade and proposed workflow solutions for office automation. The topic is subject today to a lot of interests carried by the expansion of business on the Web. However it is required need to satisfy new application requirements and execution contexts. We are interested in different aspects of process engineering: the management of change in business process; modeling and implementing Quality of Services properties (time, security, constraints...); composing existing process fragments of different nature and models; decentralizing a global process for a distributed execution with organizational constraints; process governance. Most of these aspects are considered within the frame of Web services and/or peer to peer architectures, and we are also interested in proposing new models of process for new applications domains.

6.1.1. Optimized decentralization and synchronization of cross-organizational business processes

Participants: Claude Godart, Walid Fdhila.

Globalization and the increase of competitive pressures created the need for agility in business processes, including the ability to outsource, offshore, or otherwise distribute its once-centralized business processes or parts thereof. While hampered thus far by limited infrastructure capabilities, the increase in bandwidth and connectivity and decrease in communication cost have removed these limits. An organization that aims for such fragmentation of its business processes needs to be able to separate the process into different parts. Therefore, there is a growing need for the ability to fragment one’s business processes in an agile manner, and be able to distribute and wire these fragments so that their combined execution recreates the function of the original process. Additionally, this needs to be done in a networked environment, which is where Service Oriented Architecture plays a vital role.

Our work is focused on solving some of the core challenges resulting from the need to dynamically restructure enterprise interactions. Restructuring such interactions corresponds to the fragmentation of intra and inter enterprise business process models. It describes how to identify, create, and execute process fragments without loosing the operational semantics of the original process models. It also proposes methods to optimize the fragmentation process in terms of QoS properties and communication overhead [21], [10]. Moreover, it presents a framework to model web service choreographies in Event Calculus formal language.

Walid Fdhila has successfully defended his thesis on October, 7th [1].

6.1.2. A Declarative Approach to Web Services Computing

Participants: Olivier Perrin, Ehtesham Zahoor, Claude Godart.

Web services composition and monitoring are still highly active and widely studied research directions. Little work however has been done in integrating these two dimensions using an unified framework and formalism. Classical approaches introduce an additional layer for handling the composition monitoring and thus do not provide the important execution time violations feedback to the composition process. This year, we proposed the DISC framework which aims to provide a highly declarative event-oriented model to accommodate various aspects such as composition design and exceptions, data relationships and constraints, business calculations and decisions, compliance regulations, security or temporal requirements. Then, the same model is used for combining the control of the composition definition, its execution and the composition monitoring. We proposed a service oriented architecture with a flexible logic, including complex event patterns and choreographies, business oriented rules, and dynamic control of compositions. Advantages of this unified framework are the higher level of abstraction to design, execute, and reason upon a composition, the flexibility
of the approach, and the ability to easily include non-functional requirements such as temporal or security issues and we implement the DISC framework using the Discrete Event Calculus reasoner. Ehtesham Zahoor defended his thesis in November [3], and had presented the DISC framework at ICWS 2011 [20].

We also continued the previous work initiated within the Associate Team INRIA VanaWeb about the provisioning of Web services composition using constraints solvers. The approach consists in instantiating this abstract representation of a composite Web service by selecting the most appropriate concrete Web services. This instantiation is based on constraint programming techniques which allow matching Web services according to a given request. The proposal performs this instantiation in a distributed manner, i.e., the solvers for each service type are solving some constraints at one level, and they are forwarding the rest of the request (modified by the local solution) to the next services. When a service cannot provision part of the composition, a distributed backtrack mechanism enables to change previous solutions (i.e., provisions). A major interest of this approach is to preserve privacy: solutions are not sent to the whole composition, services know only the services to which they are connected, and parts of the request that are already solved are removed from the next requests.

6.1.3. Alignement between Business Process and Service Architecture

Participants: François Charoy, Karim Dahman, Claude Godart.

In the continuation of work done previously on change management during process execution, we are conducting work on the governance of change at the business level and on its implications at the architecture and infrastructure level of an information system. Last year was devoted to the definition of the transformation rules that allowed to go from a business model to an IT model, i.e. a transformation between model based on different paradigms. During this year, a great deal of effort has been done in order to extend our work on Business to IT alignment management. Our goal is still to maintain this alignment at the lowest possible cost when the business process are changing [9]. Further than that we are trying to describe and validated an engineering method to help designer to maintain this alignment [8].

6.1.4. Distributed Processes with Security Constraints

Participants: Olivier Perrin, Aymen Baouab, Claude Godart.

Distributed processes governance is a very important challenge. In the past, we proposed various approaches for dealing with distribution, particularly for computing a set of sub-processes that can be distributed and that are equivalent to a given process. However, we did not deal with non-functional requirements as the focus was only on control and data flows. In this work, we try to deal not only with functional requirements, but also with non-functional requirements, in particular the security aspects. With Aymen Baouab, we already proposed an event-based approach that is able to verify that choreographies are valid with respect to given constraints (security constraints for instance) [7].

6.1.5. A Crisis Management Process Model

Participants: François Charoy, Joern Franke.

As said before, crisis management has been a very fruitful domain to investigate new approaches in the domain of high value, human driven activity coordination in a multi organisational setting. Our work benefits from a large amount of use cases and detailed accounts of previous dramatic events to analyze requirements and confront our proposals. 2011 has been devoted to terminate the evaluation and the validation of the model that we have defined during the previous years. It has also been devoted to complete the work done in the previous years on the inter organisational dimension of the coordination management [11]. The entire contribution on crisis management, i.e. the model, the system and the evaluation are described in Joern Franke PhD document [2].

In order to try to leverage this work in a more information system oriented way, we have started some collaboration to confront our view on coordination with existing reference model for humanitarian operations[12]. We are currently looking for alternative financing vehicle in order to continue this work.
This work was conducted as a cooperation with SAP Research Sophia Antipolis and partially funded by a CIFRE Grant.

6.2. Distributed Collaborative Systems

Starting with Web2.0 era, the web became easily writable and changeable, and nowadays, it is getting more real-time. Rather than requiring that users or their software check a source periodically for updates, real-time web is a paradigm based on the principle of pushing information to users as soon as it is available. We are faced with an explosion of real-time social software (Twitter, Facebook, etc.). Even if many social software are currently available, most of them rely on collaborative systems with a centralized architecture or authority and consequently suffer of intrinsic problems of centralization: lack of fault tolerance, poor scalability, costly infrastructure, problems of privacy.

Distributed collaborative systems (DCS) ease and coordinate collaboration among multiple users who jointly fulfill common tasks over computer networks without the need of a central architecture or authority.

We continued our work on migrating DCS to pure peer-to-peer architectures. This year we focussed on the real-time aspect of the collaboration. We evaluated replication mechanisms suitable for real-time collaboration over peer-to-peer architectures.

Moreover, peer-to-peer collaborative systems need revisiting traditional security models that prevent users from accessing to data and granted rights are checked before access is allowed. These access control mechanisms are too strict and they do not scale well in a peer-to-peer architecture. We make the assumption that an effective collaboration should rely on a flexible optimistic security model based on trust. This year we proposed a new collaboration model based on contracts where we rely on an optimistic security model. Rather than adopting an a priori strict enforcement of security rules, in this optimistic solution, access is given first to data without control but with restrictions that are verified a posteriori.

6.2.1. Evaluation of algorithms for Peer-to-Peer Real-time collaboration

Participants: Mehdi Ahmed-Nacer, Claudia Ignat, Gérald Oster, Hyun-Gul Roh, Pascal Urso.

Nowadays, real-time collaborative editing systems such as Etherpad or Google Docs became very popular. The operational transformation (OT) approach is a traditional optimistic replication mechanism that was used for real-time collaboration. Recently, Commutative Replicated Data Types (CRDTs) were introduced as a new class of replication mechanisms whose concurrent operations are designed to be natively commutative. CRDTs, such as WOOT, Logoot, Treedoc, and RGAs, are expected to be substitutes of replication mechanisms in collaborative editing systems.

We investigated the suitability of CRDTs for real-time collaborative editing [6]. To reflect the tendency of decentralized collaboration, which can resist censorship, tolerate failures, and let users have control over documents, we collect editing logs from real-time peer-to-peer collaborations. We provided a theoretical evaluation as well as an experimental one by replaying the editing logs on various CRDTs and OT algorithms implemented in the same environment. We found out that CRDT algorithms initially designed for peer-to-peer asynchronous collaboration are suitable for real-time collaboration. Moreover, they outperform some representative operational transformation approaches that were well established for real-time collaboration in terms of generation time, remote integration time and space complexity.

6.2.2. Contract-based collaboration

Participants: Claudia Ignat, Hien Thi Thu Truong.

In the push-pull-clone collaborative editing model widely used in distributed version control systems users replicate shared data, modify it and redistribute modified versions of this data without the need of a central authority. However, in this model no usage restriction mechanism is proposed to control what users can do with the data after it has been released to them. We extended the push-pull-clone model with contracts that express usage restrictions and that are checked a posteriori by users when they receive the modified data [18], [25]. We proposed a merging algorithm that deals not only with modifications on data but also with contracts. A
log-auditing protocol [19] was used to detect users who do not respect contracts after they received data and to adjust user trust levels. The associated trust values can be computed by using any existing decentralised trust model. Our proposed contract-based model has been implemented and evaluated by using PeerSim simulator.

6.3. Interoperability and Enterprise Modeling

Participants: Nacer Boudjlida [contact], Khalid Benali.

In the continuation of our previous work on semantic-based and model-based solutions for interoperability, we applied and experienced a variety of semantic annotation types (structural, terminological and behavioral) in the frame of dynamic web services discovery and for competence management systems. In addition, we explored semantic graphs as a formal framework for competence description and management. Further, in order to ease semantic interoperability of heterogeneous competence management systems, we are defining a generic representation model that could serve as a shared ontology for these types of systems.
5. New Results

5.1. Flexible Radio Node

Participants: Florin Hutu, Tanguy Risset, Jacques Verdier, Guillaume Villemaud, Cédric Levy-Bencheton.

This section summarizes the early results obtained from the research axis flexible radio nodes.

In [41], [75], a candidate architecture for LTE-Advanced receiver is proposed. Based on the combination of MIMO techniques and flexible spectrum access, LTE-Advanced terminals will require the increasing of the analog front-end complexity. To reduce the complexity of the analog front-end, an innovative architecture based on the merge between the double IQ and the code multiplexing structures is proposed. Simulation and measurement results show that, in a Gaussian case, the bit error rate is similar when using the proposed architecture and the state of the art front-end stack-up structure. A complexity evaluation study reveals significantly reduced power consumption of the proposed single front-end architecture.

The current generation of mobile terminals can communicate on multiple modes using several antennas. However, their energy consumption remains a critical parameter. In [58], [74], we explore the combination of multiple communication modes and MIMO as a possible way to reduce the energy consumption of both the terminals and the network. We propose a realistic energy model for the PHY layer of a MIMO and multi-mode terminal, taking into account the MAC layer behaviour. We show that the combination of MIMO and multi-mode provides a solution to reduce global energy consumption.

Software means programmable. Hence software defined radio means that the radio should now be programmable. We know what computer programming means, and we agree, up to a certain level, on how it should be done. But do we know what programming a radio means? Several questions are still open: what will an SDR platform look like in ten years? Will there exist software radio code? What will be the technical challenges and commercial issues behind this code? Programming is more precise than configuring or tuning, it implies a much greater level of freedom for the programmer. But it also means much cheaper implementations in many cases and in particular a re-use of the same hardware for different protocols (i.e. with different programs). This is, to our point of view [76], the main difficulty of software radio programming: reconfiguration and in particular dynamic reconfiguration. Dynamic (i.e. very fast) reconfiguration is now mandatory because some protocols, 3GPP-LTE (Third Generation Partnership Program Long Term Evolution) for instance, propose channel adapting for each frame, requiring a setting of the channel estimation parameter in a few milliseconds.

5.2. Agile radio resource sharing

Participants: Jean-Marie Gorce, Claire Goursaud, Katia Jaffrès-Runser, Nikolaï Lebedev, Guillaume Villemaud, Paul Ferrand, Philippe Mary.

This section presents our recent results concerning the realistic modeling of wireless links to develop realistic models and efficient simulations. This work include theoretical developments like symbol error outage modeling, but also some applications in the context of LTE multi-cells association, or opportunistic relaying in the context of wireless sensor networks. Other contributions about resource sharing are presented in next sections below, in the section ‘network optimization’ and the section ‘network coding’.

In [28], we addressed the problem of finding a tractable expression for the symbol error outage (SEO) in flat Nakagami-m fading and shadowing channels. We deal with M-ary phase shift keying (M-PSK) and quadrature amplitude modulation (M-QAM) which extends our previous results on BPSK signaling. We propose a new tight approximation of the symbol error probability (SEP) holding for M-PSK and M-QAM signals which is accurate over all signal to noise ratios (SNRs) of interest. We derive a new generic expression for the inverse SEP which facilitates the derivation of a tight approximation of the SEO in a lognormal shadowing environment.
In [44], we consider on-body BAN nodes transmitting information towards a common sink, in a star topology (Body Area Networks (BAN) offer amazing perspectives to instrument and support humans in many aspects of their lives). While this setup is usual in wireless networks, the high instability of the BAN radio channel and the proximity of the body make classical communication protocols inefficient. These networks are further constrained by the low transmission power required for both battery life and health concerns. Opportunistic cooperation techniques are of great interest in such environment to ensure reliable communications. In previous works, we studied simple opportunistic relaying schemes under independent BAN links, using a packet error outage criterion. In this paper, we introduce a more realistic case where shadowing variations around the body are now assumed strongly correlated. Generally speaking, there is a lack of definitive measurements and models for the shadowing correlation in multi-hop networks, while it can play a crucial role at the higher layers. Based on the measurement and simulation results of the French BANET project, we use the BAN context as an illustrative example to exhibit how shadowing correlations have a strong impact on relaying approaches performance.

Opportunistic networking aims at exploiting sporadic radio links to improve the connectivity of a multi-hop network and to foster data transmissions. Broadcast nature of the wireless channel is an important feature that can be exploited to improve transmissions by using several potential receivers. Opportunistic relaying is thus the first brick for opportunistic networking. However, the advantage of opportunistic relaying may be balanced by energy increase related to having simultaneous active receivers. In [32], we proposed a thorough analysis of opportunistic relaying efficiency under different realistic radio channel conditions. The study aims at finding the best trade-off between two objectives: energy and latency minimizations, under a hard reliability constraint. We derive an optimal bound, namely, the Pareto front of the related optimization problem, which offers a good insight into the benefits of opportunistic routings compared with classical multi-hop routing schemes. Meanwhile, the lower bound provides a framework to optimize the parameters in physical layer, MAC layer and routing layer from the viewpoint of cross layer during the design or planning phase of a network.

This work has been extended in In [70] for relay channels. The gain induced by using relay channels in a linear network under both a capacity constraint and a realistic energy model is evaluated. We express a general model based on a convex optimization problem, allowing us to use numerical tools to obtain similar results for outer and inner bounds to the capacity of the full and half duplex relay channel. We then further the study with more complex networks based on relay channels, especially networks formed by a linear chain of nodes. We describe the Pareto optimal solutions of the minimization problem for with respect to the consumed energy and latency in such a linear network. From the simple case of the linear multi-hop network, we study the gains when implementing a linear chain of relay channels and compare these results to the simpler multi-hop transmission. This work will be published in 2012 in IEEE WCNC.

In [82] we extended this formalism derived for a linear network to a more general case: the problem of deriving fundamental trade-off bounds for wireless ad hoc networks when multiple performance criteria are of interest. It proposes a MultiObjective (MO) performance evaluation framework composed of a broadcast and interference-limited network model, a steady state performance metric derivation inspired by a discrete Markov chain formalism and formulates the associated MO optimization problem. Pareto optimal performance bounds between end-to-end delay and energy for a capacity-achieving network are given for the 1-relay and 2-relay networks and assessed through simulations.

5.3. Autonomous wireless networking

Participants: Isabelle Augé-Blum, Bernard Tourancheau, Fabrice Valois, Ibrahim Amadou, Cédric Chauvenet, Quentin Lampin, Alexandre Mouradian, Bilel Romdhani.

Designing protocols for large scale wireless sensors networks is a challenging issue, if realistic environments are considered. Finding a trade-off between energy consumption and delay, or capacity, is difficult. The most promising ideas rely on zero-protocol approaches and on virtual coordinates use. the special case of VANETs is presented in the next section.
In [64], we focus on Wireless Sensor Networks (WSNs) in a more realistic case than classical studies and previous works: we consider wireless sensor nodes having different transmission ranges according to the environment and/or to the wireless chipset. The main consequence of this heterogeneity is the existence of asymmetric links. Such links in a WSN degrade the performance of most protocols which have not been designed to support this heterogeneity and to deal with asymmetric links: so, mainly, these links are pruned. Under this assumption, we propose a routing protocol for data collection from sensors nodes to the sink node in heterogeneous WSNs. Our proposal detects and takes benefit from asymmetric links caused by this heterogeneity. Our proposal, denoted MURA, (1) provides a high delivery ratio, (2) reduces the number of duplicated packets and (3) reduces the number of hop counts by exploiting the asymmetric links.

Due to the efficiency and scalability of greedy routing in WSNs and the financial cost of GPS chips, Virtual Coordinate Systems (VCSs) for WSNs have been proposed. A category of VCSs is based on the hop-count from the sink, this scheme leads to many nodes having the same coordinate. The main advantage of this system is that the hops number of a packet from a source to the sink is known. Nevertheless, it does not allow to differentiate the nodes with the same hop-count. We propose in [87] a novel hop-count-based VCS which aims at classifying the nodes having the same hop-count depending on their connectivity and at differentiating nodes in a 2-hop neighborhood. Those properties make the coordinates, which also can be viewed as a local identifier, a very powerful metric which can be used in WSNs mechanisms.

Duty-cycled medium access protocols allow for long lasting autonomous networks by periodically putting nodes to sleep. However, this life expectancy improvement comes at the cost of a lesser network capacity and a poor adaptability to bursty traffic loads. Indeed, existing contention algorithms do not provide efficient algorithms to dynamically elect multiple senders per wake-up periods. In [84], the medium is divided in several logical channels (e.g., obtained by a time/frequency division of the communication medium) and we propose to allocate them dynamically among senders. For this purpose, we propose a joint contention/scheduling algorithm, named Extended Slot Selection (ESS), that schedules multiple sender/receiver pairs to available logical channels.

Energy-efficient communication protocol is a primary design goal for Wireless Sensor Networks (WSNs). Many efforts have been done to save energy: MAC with duty cycle, energy-aware routing protocols, data aggregation schemes, etc. Recently, beacon-less strategies have emerged as new direction to improve considerably the WSN lifetime. However, the main contributions are not suitable to real radio environments because of hole avoiding strategies based on either planarization or explicit neighbor solicitations. We propose in [34] PFMAC (Pizza-Forwarding Medium Access Control), which combines beacon-less geo-routing and energy efficient MAC protocol via a cross-layer design to save more energy with higher reliability. PFMAC supports radio interferences, asymmetric radio links, etc. PFMAC supports a greedy forwarding strategy and, a reactive and optimized neighborhood discovery at 2-hop to deal with holes. Intensive simulations are proposed to highlight the behavior and the performance of PFMAC compared to BOSS over BMAC.

To provide for reliability in Wireless Sensor Networks (WSNs), Medium Access Control (MAC) protocols must be adapted by mechanisms taking cross-layer approaches into account. In [51], [52], we describe AreaCast which is designed for enhancing reliability in WSNs. AreaCast is a MAC layer mechanism independent of the routing layer, but uses only local topological and routing information to provide a communication by area instead of a traditional, node-to-node communication (i.e., unicast). In AreaCast, a source node addresses a set of nodes: an explicit relay node chosen as the next hop by a given routing protocol, and k other implicit relay nodes. The neighboring nodes select themselves as implicit relays according to their location from the explicit relay node. This mechanism uses overhearing to take advantage of the inherent broadcast nature of wireless communications. Without changing the routing protocol, AreaCast is able to dynamically avoid a byzantine node or an unstable link, allowing to benefit from the inherent topological redundancy of densely deployed sensor networks. Simulation results show that AreaCast significantly improves the packet delivery rate while having a good reliability-energy consumption trade-off.

Improving the network lifetime is an important design criterion for wireless sensor networks especially if we want to use standard solution like IPv6. In [38], we propose a novel approach which applies source-coding on addresses in heterogeneous IPv6 Cluster-based wireless sensor network. We formulate the problem...
of maximizing the network lifetime when Slepian-wolf coding is applied on addresses in network composed of line-powered and battery-powered sensors. The numerical results show that a significant network lifetime improvement can be achieved (about 25% in typical scenario). In [36], we investigate the sinks mobility in IPv6-based wireless sensors networks and specially in the new IETF proposed protocol RPL (Routing Protocol for Low power and Lossy Networks). We also show that even the mobility of sinks is not an explicit design criteria, the use of mobile sinks improves the network lifetime.

5.4. Wireless networking in VANETs

Participants: Marco Fiore, Sandesh Uppoor.

VANETS (Vehicular Ad hoc Networks) represents a challenging context for designing new protocols as it offers new challenges related to the high dynamicity of the network. In cooperation with external researchers, we derived recent results on mobility modeling and data dissemination in VANETS. This work is a part of the work on 'Autonomous wireless networking', but dedicated specially for VANETS.

Simulation is the tool of choice for the large-scale performance evaluation of upcoming telecommunication networking paradigms that involve users aboard vehicles, such as next-generation cellular networks for vehicular access, pure vehicular ad hoc networks, and opportunistic disruption-tolerant networks. The single most distinguishing feature of vehicular networks simulation lies in the mobility of users, which is the result of the interaction of complex macroscopic and microscopic dynamics. Notwithstanding the improvements that vehicular mobility modeling has undergone during the past few years, no car traffic trace is available today that captures both macroscopic and microscopic behaviors of drivers over a large urban region, and does so with the level of detail required for networking research. In [66], we present a realistic synthetic dataset of the car traffic over a typical 24 hours in a 400-km² region around the city of Koln, in Germany. We outline how our mobility description improves today’s existing traces and show the potential impact that a comprehensive representation of vehicular mobility can have on the evaluation of networking technologies.

In [21], [30], we investigate data dissemination in vehicular networks. Content downloading in vehicular networks is a topic of increasing interest: services based upon it are expected to be hugely popular and investments are planned for wireless roadside infrastructure to support it. We focus on a content downloading system leveraging both infrastructure-to-vehicle and vehicle-to-vehicle communication. With the goal to maximize the system throughput, we formulate a max-flow problem that accounts for several practical aspects, including channel contention and the data transfer paradigm. Through our study, we identify the factors that have the largest impact on the performance and derive guidelines for the design of the vehicular network and the roadside infrastructure supporting it.

In [45] we address cooperative caching in wireless networks, where the nodes may be mobile and exchange information in a peer-to-peer fashion. We consider both cases of nodes with large- and small-sized caches. For large-sized caches, we devise a strategy where nodes, independent of each other, decide whether to cache some content and for how long. In the case of small-sized caches, we aim to design a content replacement strategy that allows nodes to successfully store newly received information while maintaining the good performance of the content distribution system. Under both conditions, each node takes decisions according to its perception of what nearby users may store in their caches and with the aim of differentiating its own cache content from the other nodes. The result is the creation of content diversity within the nodes neighborhood so that a requesting user likely finds the desired information nearby. We simulate our caching algorithms in different ad hoc network scenarios and compare them with other caching schemes, showing that our solution succeeds in creating the desired content diversity, thus leading to a resource-efficient information access.

Performance and reliability of content access in mobile networks is conditioned by the number and location of content replicas deployed at the network nodes. In [27], we design a practical, distributed solution to content replication that is suitable for dynamic environments and achieves load balancing. Simulation results show that our mechanism, which uses local measurements only, approximates well an optimal solution while being robust against network and demand dynamics. Also, our scheme outperforms alternative approaches in terms of both content access delay and access congestion.
5.5. Optimization in wireless networks

**Participants:** Jean-Marie Gorce, Nikolaï Lebedev, Hervé Rivano, Fabrice Valois, Anis Ouni, Virgile Garcia, Cengis Hasan.

In the context of the common lab between Inria and Alcatel Lucent Bell Labs and the ANR Ecoscells project, we work on optimizing wireless networks performance. In one side, we work on distributed algorithms for optimal resource allocation and/or mobile-BS association. On the other side, we work on mesh wireless networks optimization.

Multi-cell processing, also called Coordinated Multiple Point (CoMP), is a promising distributed technique that uses neighbor cells’ antennas [48]. It is expected to be the part of next generation cellular standards such as LTE-A. Small cell networks in dense urban environments are limited by interferences and CoMP can strongly take advantage of this fact to improve cell-edge users’ throughput. The present study introduces a distributed criterion for mobiles to select their optimal set of Base Stations (BS) to perform CoMP, and evaluates the impact of this association on the fairness and the total cell throughput. For that, we use a known theoretical expression for the capacity outage probability of CoMP under Rayleigh fading and evaluate the goodputs of antennas associations. The proposed criterion is used in combination with fair resource allocation to perform a joint double-objective optimization of fairness and efficiency. In [48], [91], we provide the analysis of the downlink Coordinated Multiple Point (CoMP) used in conjunction with the basic MIMO. The CoMP is the joint multi-cell transmission from several BS to mobiles, coupled here to an open-loop MIMO technique that does not require the perfect channel state knowledge. We show by simulation, that even for $4 \times 4$ MIMO transmission, the CoMP can improve the spectral efficiency for some mobiles, depending on capacity outage requirements.

In [33], we considered downlink transmission in cellular networks where we target to reduce the energy consumption by switching off some base stations by such a way that the distribution of SINR remains unchanged. This is a mean of green networking in cellular networks in downlink consideration. This paper analyzes for line and plane cases, the gain in power consumption obtained after switching off base stations. By computations we observe that the more the operational cost the more the gain in power consumption.

In [47], we propose an autonomous radio resource allocation and optimization scheme that chooses the transmit power and precoding vector among codebooks for multiple antennas transmitters to improve spectral and power efficiency and provide user fairness. Network self-optimization is an essential feature for supporting the cell densification in future wireless cellular systems. The proposed self-optimization is inspired by Gibbs sampler. We show that it can be implemented in a distributed manner and nevertheless achieves system-wide optimization which improves network throughput, power utilization efficiency, and overall service fairness. In addition, we extend the work and include power pricing to parametrize and enhance energy efficiency further. Simulation results show that the proposed scheme can outperform today’s default modes of operation in network throughput, energy efficiency, and user fairness.

In [55], we focused on broadband wireless networks based on OFDMA resource management, such as LTE systems. We have investigated two optimization problems, one concerning a backhauling mesh infrastructure while the other is the allocation of modulation and coding, subcarriers and power to users in LTE. Considering a realistic SINR model of the physical layer with a fine tuned power control at each node, a linear programing model using column generation has been developed for computing power efficient schedules with high network capacity for wireless mesh backhauling networks. Correlation between capacity and energy consumption have been analyzed as well as the impact of physical layer parameters - SINR threshold and path-loss exponent. With these models, we highlight that there is no significant tradeoff between capacity and energy when the power consumption of idle nodes is important. We also show that both energy consumption and network capacity are very sensitive to the SINR threshold variation. Finally, simulation results show that compared to classic reuse schemes the proposed approach is able to pack more users into the same bandwidth, decreasing the probability of user outage.

In [62], we focus on broadband wireless mesh networks like 3GPP LTE-Advanced. This technology is a key enabler for next generation cellular networks which are about to increase by an order of magnitude
the capacity provided to users. Such an objective needs a significant densification of cells which requires an efficient backhauling infrastructure. In many urban areas as well as under-developed countries, wireless mesh networking is the only available solution. Besides, economical and environmental concerns require that the energy expenditure of such infrastructure is optimized. We propose a multi-objective analysis of the correlation between capacity and energy consumption of LTE-like wireless mesh networks. We provide a linear programming model using column generation for an efficient computation of the Pareto front between these objectives. Based on this model, we observe that there is actually no significant capacity against energy trade-off.

In [63], broadband wireless mesh networks based on OFDMA resource management are studied considering a realistic SINR model of the physical layer with a fine tuned power control at each node. A linear programming model using column generation leads to compute power efficient schedules with high network capacity. Correlation between capacity and energy consumption is analyzed as well as the impact of physical layer parameters - SINR threshold and path-loss exponent. We highlight that there is no significant tradeoff between capacity and energy when the power consumption of idle nodes is important. We also show that both energy consumption and network capacity are very sensitive to the SINR threshold variation.

5.6. Network coding in WSN

Participants: Jean-Marie Gorce, Cédric Lauradoux, Marco Fiore, Claire Goursaud, Marine Minier, Anya Apavatjurt, Yuanyuan Zhang, Wassim Znaidi.

Network coding associated with Fountain codes is a very efficient approach to increase the throughput of multi-hop networks. However, severe outcomes are still expected, especially to develop robust and energy efficient approaches for transmitting data over a large scale networks. Network coding is also very promising for security issues as presented below.

Diversity is a powerful means to increase the transmission performance of wireless communications. For the case of fountain codes relaying, it has been shown previously that introducing diversity is also beneficial since it counteracts transmission losses on the channel. Instead of simply hop-by-hop forwarding information, each sensor node diversifies the information flow using XOR combinations of stored packets. This approach has shown to be efficient for random linear fountain codes. However, random linear codes exhibit high decoding complexity. In [19], we propose diversity increased relaying strategies for the more realistic and lower complexity Luby Transform code in a linear network. Results are provided herein for a linear network assuming uniform imperfect channel states.

In [29], the exact probability that a receiver obtains N linearly independent packets among K over N received packets is computed, when the sender/s use/s random linear network coding over a Galois Field of size q. Such condition maps to the receiver’s capability to decode the original information, and its mathematical characterization helps to design the coding so to guarantee the correctness of the transmission. The proposed formulation represents an improvement over the current upper bound for the decoding probability, and provides theoretical grounding to simulative results in the literature.

In [35], we focus on the proper use of fountain codes for the transmission of sporadic data in a wireless sensor network (WSN). Fountain codes offer great perspectives for the self-organization of WSNs: they self adapt to the channel error rate without any control data. When deploying fountain codes on a WSN, two problems arise. First, the size of the data transmitted by a sensor is small in comparison to the size considered traditionally with fountain codes. Second, the communications are done in an hop-by-hop fashion. It implies that the destination of the data can not acknowledge instantaneously its reception to the source. Therefore, the transmissions of useless packets for the destination can not be prevented. The flooding traffic has been evaluated as well through realistic simulations for three different relaying strategies where packets are lost due to both small scale fading and collisions for an unslotted IEEE 802.15.4 medium access layer.

Network coding has attracted the attention of many researchers in security and cryptography. We have investigated several aspects of network coding security. In [20], we propose efficient solutions to thwart pollution attacks in which an adversary injects false information into data flow. This work was further expanded
in [54] to find rational strategy to minimize the energy cost and the impact of the attack. We also came to the conclusion that dealing with pollution attacks was not enough as long as the acknowledgment messages are not also protected. The risk is to suffer from a flooding attack. This goes beyond the capabilities of cryptographic solutions and we investigate the security capabilities of multipath acknowledgment in [67].

5.7. Security


Security is an important issue for wireless networks, especially for wireless sensor networks facing an amizing increase of the number of nodes. We review in this section all contributions related to the security issue, some of them being strongly related with the PHY layer or the networking protocols. As it can be seen below, some results are strongly connected to the models and protocols derived in the other sections.

In [59], we provide the first independent analysis of the (2nd-round tweaked) 256-bit version of the SHA3 candidate SHAvite-3. By leveraging recently introduced cryptanalysis tools such as rebound attack or Super-Sbox cryptanalysis, we are able to derive chosen-related-salt distinguishing attacks on the compression function on up to 8 rounds (12 rounds in total) and free-start collisions on up to 7 rounds. In particular, our best results are obtained by carefully controlling the differences in the key schedule of the internal cipher. Most of our results have been implemented and verified experimentally.

In [50], we study a class of insider attacks called the terrorist fraud. This is a relay attack against distance bounding protocols where the prover conspires with an adversary to misrepresent the distance between himself and the verifier. In ideal situations, the adversary does not gain any knowledge about the prover’s long-term secret. This makes designing a distance bounding protocol resistant to such fraud tricky: the secrets of an honest prover must be protected, while those of a dishonest one should be disclosed as an incentive not to cheat. We demonstrate that using a secret-sharing scheme, possibly based on threshold cryptography, is well suited for thwarting the terrorist fraud. Although such an idea has been around since the work of Bussard and Bagga, this is the first time that secret-sharing and terrorist fraud have been systematically studied altogether.

In [40], we deal with the problem of radio jamming. Jamming is a major threat against wireless communications. In this paper, we evaluate the effect of jamming on an UWB link employing a PPM non-coherent receiver. We optimize the jammer parameters that are the central frequency and the bandwidth based on the metric of the signal-to-jamming ratio. The optimization depends on different system parameters such as the channel model and the integration time of the receiver.

In [23], we focus on the resiliency of wireless sensor network routing protocols against selective forwarding attacks by compromised nodes. Informally, resiliency should be understood as the capacity of the routing protocol to endure and mitigate the presence of a certain number of compromised nodes seeking to disturb the routing process. To provide for security when nodes may be compromised, cryptographic solutions must be completed by algorithmic solutions considering "beyond cryptography" approaches. After discussing the shortcomings of existing routing protocols against packet-dropping malicious nodes we describe some protocol behaviors enhancing routing resiliency under several combined routing attacks. We propose in this paper the behaviors enhancing the resiliency of routing protocols under several combined routing attacks.

5.8. Network simulation tools

Several works in 2011 have been using simulation results. Nevertheless, Swing members are strongly working on improving network simulation frameworks to provide realistic simulations. Several contributions to the simulation tools wiplan ans wsnets have been proposed.

Some contributions to WSnet concern BAN environments implementation [44] and network coding features [19], [81]. Different protocols have been also implemented for wireless sensor networks [34], [84], specifically in the context of our collaboration with Orange Labs, Grenoble.
The wiplan simulator has been developed at CITI for several years. It is based on a frequency domain ParFlow (MR-FDPF) implementation that represents a unique finite elements based method for estimating the radio propagation in complex environments. In the context of heterogeneous networks, femtocells are very promising. In order to properly simulate their behavior and their impact on the macrocell layer, it is necessary to be able to simulate the radio coverage of femtocells. Hence ParFlow is a possible deterministic model that can be used for such simulation. In [42], two implementations of ParFlow are presented: time domain and frequency domain. The performance are compared and the advantages/drawbacks of each model are investigated.

In [56], we propose to use finite difference propagation methods to evaluate the wide band properties of the fast fading. For this purpose we adapted the MR-FDPF propagation model to simulate large bandwidth by combining numerous narrow band simulations. The results are compared with a channel sounder measurement campaign covering a bandwidth of up to 70 MHz. It is verified that fading characteristics in wireless channels varies with frequency and the MR-FDPF method is capable for simulating this variation of fadings for wide band systems.

In [56], a new approach is proposed allowing extracting the fading statistics for indoor radio channels based on the electric field strength predicted with the MR-FDPF method. The performance of the proposed approach is verified both by simulations and measurements.

In [65], we propose a new hybrid modeling method for indoor-to-outdoor radio coverage prediction. The proposed method is a combination of a ray-optical channel modeling approach and the frequency domain ParFlow method. While the former is widely used for modeling outdoor propagation environments, the latter is computationally efficient and accurate for modeling indoor environments.

In [90], we propose to use finite difference propagation methods to evaluate the wide band properties of the fast fading. For this purpose we adapted the MR-FDPF propagation model to simulate large bandwidth by combining numerous narrow band simulations. The results are compared with a channel sounder measurement campaign covering a bandwidth of up to 70 MHz. It is verified that fading characteristics in wireless channels varies with frequency and the MR-FDPF method is capable for simulating this variation of fading for wide band systems.
6. New Results

6.1. Design and Performance Analysis of Wireless Networks

Participants: François Baccelli, Florence Bénézit, Bartłomiej Blaszczyszyn, Chung Shue Chen, Mir Omid Haji Mirsadeghi, Frédéric Morlot, Tien Viet Nguyen, Van Minh Nguyen.

This axis bears on the analysis and the design of wireless access communication networks. Our contributions are organized in terms of network classes: cellular networks, wireless LANs and MANETs, VANETs. We also have a section on generic results that regard more general wireless networks. We are interested both in macroscopic models, which are particularly important for economic planning and in models allowing the definition and the optimization of protocols. Our approach combines several tools, queueing theory, point processes, stochastic geometry, random graphs, distributed control algorithms, self organization protocols.

6.1.1. Cellular Networks

The activity on cellular networks has several complementary facets ranging from performance evaluation to protocol design. The work is mainly based on strong collaborations with Alcatel-Lucent and Orange Labs.

6.1.1.1. Effect of Opportunistic Scheduling on the Quality of Service Perceived by the Users in OFDMA Cellular Networks

Our objective in [20] is to analyze the impact of fading and opportunistic scheduling on the quality of service perceived by the users in an Orthogonal Frequency Division Multiple Access (OFDMA) cellular network. To this end, assuming Markovian arrivals and departures of customers that transmit some given data volumes, as well as some temporal channel variability (fading), we study the mean throughput that the network offers to users in the long run of the system. Explicit formulas are obtained in the case of allocation policies, which may or may-not take advantage of the fading, called respectively opportunistic and non-opportunistic. The main practical results of the present work are the following. Firstly we evaluate for the non-opportunist allocation the degradation due to fading compared to Additive White Gaussian Noise (AWGN) (that is, a decrease of at least 13% of the throughput). Secondly, we evaluate the gain induced by the opportunistic allocation. In particular, when the traffic demand per cell exceeds some value (about 2 Mbits/s in our numerical example), the gain induced by opportunism compensates the degradation induced by fading compared to AWGN. Partial results were presented at ComNet in 2009 [62].

6.1.1.2. Impact of Shadowing on QoS in Cellular Networks

Shadowing is believed to degrade the quality of service (QoS) in wireless cellular networks. Assuming log-normal shadowing, and studying mobile’s path-loss with respect to the strongest (serving) base station (BS) and the corresponding interference factor (the ratio of the sum of the path-gains form interfering BS’s to the path-gain from the serving BS), which are two key ingredients of the analysis and design of the cellular networks, in [48] we discovered a more subtle reality. We observe, as commonly expected, that a strong variance of the shadowing increases the mean path-loss with respect to the serving BS, which in consequence, may compromise QoS. However, in some cases, an increase of the variance of the shadowing can significantly reduce the mean interference factor and, in consequence, improve some QoS metrics in interference limited systems, provided the handover policy selects the strongest BS as the serving one. We exemplify this phenomenon, similar to stochastic resonance, studying the blocking probability in regular, hexagonal networks in a semi-analytic manner, using a spatial version of the Erlang’s loss formula combined with Kaufman-Roberts algorithm. More detailed probabilistic analysis explains that increasing variance of the log-normal shadowing amplifies the ratio between the strongest signal and all other signals thus reducing the interference. The above observations might shed new light, in particular on the design of indoor communication scenarios. Partial results were presented at IFIP WMNC’2010 [63].
6.1.1.3. Self-Optimization of Radio Resources in Cellular Networks

In [65], we developed mathematical and algorithmic tools for the self-optimization of mobile cellular networks. Scalable algorithms which are based on local measurements and do not require heavy coordination among the wireless devices were proposed. We focused on the optimization of transmit power and of user association. The method is applicable to both joint and separate optimizations. The global utility minimized is linked to potential delay fairness. The distributed algorithm adaptively updates the system parameters and achieves global optimality by measuring SINR and interference. The algorithms are built on Gibbs’ sampler and offer a unified framework that can be easily used for different purposes.

In [32], we investigated the joint optimization of radio resources in heterogeneous cellular networks made of a juxtaposition of macro and small cells. We showed that within this context, it is essential to use algorithms able to simultaneously solve the problems of channel selection, user association and power control. In such networks, the unpredictability of the cell and user patterns also requires self-optimized schemes. We proposed a generalized solution which is based on Gibbs’ sampler. It can be implemented in a distributed way and nevertheless achieves minimal system-wide potential delay. Results show that it is effective in both throughput and energy efficiency.

In [35], we extended it to an autonomous radio resource allocation and optimization scheme that chooses the transmit power and precoding vector among codebooks for multiple antennas transmitters to improve spectral and power efficiency and provide user fairness. Network self-optimization is an essential feature for supporting the cell densification in future wireless cellular systems. Besides, we included power pricing to parametrize and to enhance the energy efficiency. Simulation results show that the proposed scheme can outperform today’s default modes of operation in network throughput, energy efficiency, and user fairness.

Three patents were filed under the INRIA/Alcatel–Lucent joint laboratory.

6.1.1.4. Best Signal Quality in a Wireless Network

In a wireless network composed of randomly scattered nodes, the characterization of the distribution of the best signal quality received from a group of nodes is of primary importance for many network design problems. The thesis of Van Minh Nguyen [7] developed a framework for analyzing this distributions using shot noise models for the interference field. The joint distribution of the interference and the maximum signal strength was identified. The best signal quality can be represented as a function of these two quantities. Particular practical scenarios were also analyzed where explicit expressions can be obtained.

6.1.1.5. Cellular Network Tomography

The Foschini-Miljanic’s [67] algorithm is used for power control in cellular networks when users require a fixed bit rate. It leads to an optimal choice of power by the users in a distributed way when such a solution exists. If the users are too greedy or too many, the network saturates, and it is not possible to provide the required bit rates. We have been working on the question of residual bandwidth estimation in [61]. The residual bandwidth of a user is defined as the rate that this user should have to saturate the network when all other users stick to their initial rate requirement and all users use power control. The aim is to determine the residual bandwidth of a given user by local measurements. We showed that by simply changing their SINR target slightly and by listening to the evolution of interference, users can locally inverse Foschini-Miljanic’s algorithm and compute their residual bandwidth.

6.1.1.6. Coverage in Cellular Networks

Cellular networks are usually modeled by placing the base stations according to a regular geometry such as a grid, with the mobile users scattered around the network either as a Poisson point process (i.e. uniform distribution) or deterministically. These models have been used extensively for cellular design and analysis but suffer from being both highly idealized and not very tractable. Thus, complex simulations are used to evaluate key metrics such as coverage probability for a specified target rate (equivalently, the outage probability) or average/sum rate. More tractable models have long been desirable. In a joint work with J. Andrews and R. Ganti [UT Austin, USA] [9] and [34], we developed general models for multi-cell signal-to-noise-plus-interference ratio (SINR) based on homogeneous Poisson point processes and derived the coverage probability.
and rate. Under very general assumptions, the resulting expressions for the SINR cumulative distribution function involve quickly computable integrals, and in some important special cases of practical interest these integrals can be simplified to common integrals (e.g., the Q-function) or even to exact and quite simple closed-form expressions. We also derived the mean rate, and then the coverage gain (and mean rate loss) from static frequency reuse. We compared the coverage predictions obtained by this approach to the standard grid model and an actual base station deployment. We observed that the proposed model is pessimistic (a lower bound on coverage) whereas the grid model is optimistic. In addition to being more tractable, the proposed model may better capture the increasingly opportunistic and dense placement of base stations in urban cellular networks with highly variable coverage radii.

Cellular networks are in a major transition from a carefully planned set of large tower-mounted base-stations (BSs) to an irregular deployment of heterogeneous infrastructure elements that often additionally includes micro, pico, and femtocells, as well as distributed antennas. In a collaboration with H. Dhillon, J. Andrews and R. Ganti [UT Austin, USA] [66], we extended the approach of we developed a model for a downlink heterogeneous cellular network (HCN) consisting of $K$ tiers of randomly located BSs, where each tier may differ in terms of average transmit power, supported data rate and BS density. Assuming a mobile user connects to the strongest candidate BS, the resulting Signal-to-Interference-plus-Noise-Ratio (SINR) is greater than 1 when in coverage, Rayleigh fading, we derived an expression for the probability of coverage (equivalently outage) over the entire network under both open and closed access. One interesting observation for interference-limited open access networks is that at a given SINR, adding more tiers and/or BSs neither increases nor decreases the probability of coverage or outage when all the tiers have the same SINR threshold.

### 6.1.2. Mobile Ad Hoc Networks

A MANET is made of mobile nodes which are at the same time terminals and routers, connected by wireless links, the union of which forms an arbitrary topology. The nodes are free to move randomly and organize themselves arbitrarily. Important issues in such a scenario are connectivity, medium access (MAC), routing and stability. This year, we worked on the analysis of MAC and routing protocols in multi-hop MANETS in collaboration with Paul Mühlethaler [INRIA HIPERCOM], and on a game theoretic view of Spatial Aloha in collaboration with E. Altman and M.K. Hanawal [INRIA MAESTRO] [68].

#### 6.1.2.1. Improvement of CSMA/CA’s Spatial Reuse

The most popular medium access mechanism for such ad hoc networks is CSMA/CA with RTS/CTS. In CSMA-like mechanisms, spatial reuse is achieved by implementing energy based guard zones. In a new collaboration with Qualcomm ([26] and [14]), we considered the problem of simultaneously scheduling the maximum number of links that can achieve a given signal to interference ratio (SIR). Using tools from stochastic geometry, we studied and maximized the medium access probability of a typical link. Our contributions are two-fold: (i) We showed that a simple modification to the RTS/CTS mechanism, viz., changing the receiver yield decision from an energy-level guard zone to an SIR guard zone, leads to performance gains; and (ii) We showed that this combined with a simple modification to the transmit power level – setting it to be inversely proportional to the square root of the link gain – leads to significant improvements in network throughput. Further, this simple power-level choice is no worse than a factor of two away from optimal over the class of all "local" power level selection strategies for fading channels, and further is optimal in the non-fading case. The analysis relies on an extension of the Matérn hard core point process which allows us to quantify both these SIR guard zones and this power control mechanism.

#### 6.1.2.2. Opportunistic versions of CSMA/CA

In collaboration with Gustavo de Veciana and Yuchul Kim [UT Austin, ECE] we studied the benefits of channel-aware (opportunistic) scheduling of transmissions in ad-hoc networks using CSMA/CA [36]. The key challenge in optimizing the performance of such systems is finding a good compromise among three interdependent quantities, the density and channel quality of the scheduled transmitters, and the resulting interference at receivers. We propose two new channel-aware slotted CSMA protocols: opportunistic CSMA (O-CSMA) and quantile-based CSMA (QT-CSMA) and develop stochastic geometric models allowing us to quantify their performance in terms of spatial reuse and spatial fairness. When properly optimized these
protocols offer substantial improvements in terms of both of these metrics relative to CSMA - particularly when the density of nodes is moderate to high. Moreover, we show that a simple version of QT-CSMA can achieve robust performance gains without requiring careful parameter optimization. The paper supports the case that the benefits associated with channel-aware scheduling in ad hoc networks, as in centralized base station scenarios, might far outweigh the associated overhead, and this can be done robustly using a QT-CSMA like protocol.

6.1.3. Cognitive Radio Networks

We wrote a survey [22] on the probabilistic framework which can be used to model and analyze cognitive radio networks using various classes of MAC protocols (including carrier sensing based multiple access schemes and Aloha schemes). For each model, analytical results were derived for important performance metrics. This leads to a quantification of the interplay between primary and secondary users in such networks.

6.1.4. Generic Wireless Networks

6.1.4.1. Power Control in Wireless Networks

In [10], in collaboration with N. Bambos, [Stanford] and N. Gast [EPFL], we formulated a delay-power control (DPC) scheme for wireless networking, which balances delay against transmitter power on each wireless link. The DPC scheme is scalable, as each link autonomously updates its power based on the interference observed at its receiver; no cross-link communication is required. It is shown that DPC converges to a unique equilibrium power and several key properties are established, concerning the nature of channel bandwidth sharing achieved by the links. The DPC scheme is contrasted to the well-known Foschini-Miljanic (FM) formulation for transmitter power control in wireless networks, and some key advantages are established. Based on the DPC and FM schemes, two protocols are developed, which leverage adaptive tuning of DPC parameters. One of them is inspired by TCP and exhibits analogous behavior.

In [21], we studied the weighted sum rate maximization problem in wireless networks consisting of multiple source-destination pairs. The optimization problem is to maximize a weighted sum of data rates by adjusting the power of each user. The problem is in general a non-convex optimization problem that will lead to multiple local maxima. A Gauss-Seidel type iterative power control algorithm was presented. We showed by simulation that the proposed algorithm converges to the global maximum with very high probability, if we initialize the initial power allocation uniformly at random. The proposed algorithm also has the favorable properties that only simple operations are needed in each iteration, and the convergence is fast. Performance comparison under different user densities has also indicated its effectiveness. Finally, we discussed some simple and optimal power allocation strategies under special cases of the problem if the network can be represented by a certain approximation.

6.1.4.2. Simultaneous Decoding

In [15], in collaboration with A. El Gamal [Stanford, USA] and D. Tse [UC Berkeley, USA], we analyzed a network made of a collection of transmitter-receiver links where each link is considered to be part of a Multiple Access Channel (MAC) together with a collection of co-transmitters, rather than treating the messages of the latter as noise. This MAC extension is meant to improve the rate of the link and not to decode the messages of the co-transmitters. The necessary and sufficient condition for the feasibility of some rate when using successive interference cancellation and simultaneous decoding were provided. The reasons why simultaneous decoding is preferable to successive interference cancellation and simultaneous decoding were also given. The gain obtained when using this type of simultaneous decoding rather than treating interference as noise was then quantified in a network made of a large random collection of such links. The gains in coverage and in rate were analyzed in terms of ensemble averages, evaluated using stochastic geometry. Closed form or integral expressions were obtained for the outage/coverage probability in networks where nodes are randomly distributed like a Poisson point process on an infinite plane. In the CDMA limit (large bandwidth, low SINR per hertz, high density), the ensemble average of the link rates tends to 0 when interference is treated as noise whereas it tends to a positive constant when simultaneous decoding of infinite order is used. The whole analysis was conducted in the AWGN case.
6.2. Network Dynamics


This traditional research topic of TREC has several new threads like perfect simulation, active probing or Markov decision.

6.2.1. Network Calculus

Network calculus is a theory that aims at computing deterministic performance guarantees in communication networks. This theory is based on the (min,plus) algebra. Flows are modeled by an arrival curve that upper-bounds the amount of data that can arrive during any interval, and network elements are modeled by a service curve that gives a lower bound on the amount of service offered to the flows crossing that element. Worst-case performances are then derived by combining these curves.

6.2.1.1. Performance bounds networks with static priorities

In cooperation with Aurore Junier [INRIA/IRISA], we present in [29] algorithms to compute worst-case performance upper bounds when the service policy is static priorities, using linear programming. Linear programming does not lead to tight bounds, but when combining this method with (min,plus) methods, we obtain bounds that outperform the already known bounds. Also, we prove that in tandem networks, the the worst-case performance bound under arbitrary multiplexing can be obtain by a policy with static priorities, the “shortest-destination first” policy.

6.2.1.2. Feed-forward networks with wormhole routing discipline

In collaboration with Bruno Gaujal [INRIA Rhone Alpes] and Nadir Farhi [IFFSTAR] we are working on a model of performance bound calculus on feed-forward networks where data packets are routed under wormhole routing discipline. We are interested in determining maximum end-to-end delays and backlogs for packets going from a source node to a destination node, through a given virtual path in the network. Our objective is to give a “network calculus” approach to calculate the performance bounds. For this, we propose a new concept of curves that we call packet curves. The curves permit to model constraints on packet lengths for data flows, when the lengths are allowed to be different. We used this new concept to propose an approach for calculating residual services for data flows served under non preemptive service disciplines. This notion also enabled us to differentiate different classes of service policies: those that are based on a packet count (like round-robin and its generalized version), where the packet curve will be useful to tighten the bounds computed, and those that are based on the amount of data served (FIFO, priorities), where it won’t be useful. These results can be found in [44] and have been presented in ILAS 2011.

6.2.1.3. Composition of service curves in Network Calculus

In envelope-based models for worst-case performance evaluation like Network Calculus or Real-Time Calculus, several types of service curves have been introduced to quantify some deterministic service guarantees. We compare those different classes of service curves regarding the composition (servers in tandem) and individual service curves (when several flows share a server, what service curve can be guaranteed to each of the flows?). In short, there are two main classes of service curves, simple and strict service curves. Individual service curve can not always be computed when simple service curves are considered, and strict service curves is not a stable class regarding the two operations described. We show that there can be no equivalence between the two main classes of service curves and that no notion of service curve in-between can be defined, that behaves well for the composition. We complete this study by studying other classes of service curves from this viewpoint. These results have been presented in [28]

6.2.1.4. Residuation in (max,plus) automata

With Éric Badouel, Philippe Darondeau [INRIA/IRISA] and Jan Komenda [Institute of Mathematics, Brno], we study in [27] the decidability of existence and the rationality of delay controllers for systems with time weights in the tropical and interval semirings. Depending on the (max,+)- or (min,+)-rationality of the series specifying the controlled system and the control objective, cases are identified where the controller series
defined by residuation is rational, and when it is positive (i.e., when delay control is feasible). When the control objective is specified by a tolerance, i.e. by two bounding rational series, a nice case is identified in which the controller series is of the same rational type as the system specification series.

6.2.2. Queueing Theory and Active Probing

6.2.2.1. Inverse Problems

Active probing began by measuring end-to-end path metrics, such as delay and loss, in a direct measurement process which did not require inference of internal network parameters. The field has since progressed to measuring network metrics, from link capacities to available bandwidth and cross traffic itself, which reach deeper and deeper into the network and require increasingly complex inversion methodologies. The thesis of B. Kauffmann [6] investigates this line of thought as a set of inverse problems in queueing theory. Queueing theory is typically concerned with the solution of direct problems, where the trajectory of the queueing system, and laws thereof, are derived based on a complete specification of the system, its inputs and initial conditions. Inverse problems aim to deduce unknown parameters of the system based on partially observed trajectories. A general definition of the inverse problems in this class was provided and the key variants were mapped out: the analytical methods, the statistical methods and the design of experiments. We also show how this inverse problem viewpoint translates to the design of concrete Internet probing applications.

Inverse problems in bandwidth sharing networks theory were also investigated. A bandwidth sharing networks allocates the bandwidth to each flow in order to maximize a given utility function (typically an $\alpha$-fairness), with the constraints given by the capacity of the different servers. In particular, it has been shown that the equilibrium distribution of the bandwidth allocated by TCP to many competing connections is oscillating around an $\alpha$-fair allocation. As such, the theory of bandwidth sharing network is a high-level viewpoint of networks. The meaning of inverse problems in this theory, and their relation to the active probing paradigm are analyzed. In two simple examples of network, the capacity of the different servers and the flow population can estimated, and an algorithm to perform this estimation was proposed.

6.2.2.2. Internet Tomography

Most active probing techniques suffer of the “Bottleneck” limitation: all characteristics of the path after the bottleneck link are erased and unreachable. we are currently investigating a new tomography technique, based on the measurement of the fluctuations of point-to-point end-to-end delays, and allowing one to get insight on the residual available bandwidth along the whole path. For this, we combined classical queueing theory models with statistical analysis to obtain estimators of residual bandwidth on all links of the path. These estimators were proved to be tractable, consistent and efficient. In [59] we evaluated their performance with simulation and trace-based experiments.

Lately this method has been generalized in [72] to a probing multicast tree instead of a single path. This work deals with the complexity of the combinatorials in trees, and gives an explicit formula for the iteration of the Expectation-Maximization (E-M) algorithm. The E-M algorithm is notoriously slow, and we provided three speed-up techniques which are effective in our case (up to a factor $10^3$ in the computation time). These techniques are general, and can be applied to other instances of E-M, or even several other iterative algorithms.

6.2.3. Perfect Sampling of Queueing Systems

Propp and Wilson introduced in 1996 a perfect sampling algorithm that uses coupling arguments to give an unbiased sample from the stationary distribution of a Markov chain on a finite state space $\mathcal{X}$. In the general case, the algorithm starts trajectories from all $x \in \mathcal{X}$ at some time in the past until time $t = 0$. If the final state is the same for all trajectories, then the chain has coupled and the final state has the stationary distribution of the Markov chain. Otherwise, the simulations are started further in the past. This technique is very efficient if all the events in the system have appropriate monotonicity properties. However, in the general (non-monotone) case, this technique requires that one consider the whole state space, which limits its application only to chains with a state space of small cardinality.
6.2.3.1. Piecewise Homogeneous Events

In collaboration with Bruno Gaujal [INRIA Grenoble - Rhone-Alpes], we proposed in [47] a new approach for the general case that only needs to consider two trajectories. Instead of the original chain, we used two bounding processes (envelopes) and we showed that, whenever they couple, one obtains a sample under the stationary distribution of the original chain. We showed that this new approach is particularly effective when the state space can be partitioned into pieces where envelopes can be easily computed. We further showed that most Markovian queueing networks have this property and we propose efficient algorithms for some of them.

The envelope technique has been implemented in a software tool PSI2 (see Section 5.2).

6.2.3.2. Acceleration of Perfect Sampling by Skipping Events

In collaboration with Bruno Gaujal [INRIA Grenoble - Rhone-Alpes], we proposed a new method to speed up perfect sampling of Markov chains by skipping passive events during the simulation [38]. We showed that this can be done without altering the distribution of the samples. This technique is particularly efficient for the simulation of Markov chains with different time scales such as queueing networks where certain servers are much faster than others. In such cases, the coupling time of the Markov chain can be arbitrarily large while the runtime of the skipping algorithm remains bounded. This was further illustrated by several experiments that also show the role played by the entropy of the system in the performance of our algorithm.

6.2.3.3. Aggregated Envelopes

When the cardinality of the state space is so huge that even storing the state of the Markov chain becomes challenging, we propose to combine the ideas of bounding processes and the aggregation of Markov chains [30]. We illustrate the proposed approach of aggregated envelope bounding chains on queueing models with joint arrivals and joint services, often referred to in the literature as assemble-to-order systems. Due to the finite capacity, and coupling in arrivals and services, the exact solving techniques are inefficient for larger problem instances. For instance, for the service tools model proposed by Vliegen and Van Houtum (2009), the aggregated envelope method reduces exponentially the dimension of the state space and allows effective perfect sampling algorithms. We also provide bounds for the coupling time, under the high service rate assumptions.

6.2.4. Markov Chains and Markov Decision Processes

Solving Markov chains is in general difficult if the state space of the chain is very large (or infinite) and lacking a simple repeating structure. One alternative to solving such chains is to construct models that are simple to analyze and provide bounds for a reward function of interest. The bounds can be established by using different qualitative properties, such as stochastic monotonicity, convexity, submodularity, etc. In the case of Markov decision processes, similar properties can be used to show that the optimal policy has some desired structure (e.g. the critical level policies).

6.2.4.1. Stochastic Monotonicity

In collaboration with Jean-Michel Fourneau [PRiSM, Université de Versailles Saint-Quentin] we consider two different applications of stochastic monotonicity in performance evaluation of networks [18]. In the first one, we assume that a Markov chain of the model depends on a parameter that can be estimated only up to a certain level and we have only an interval that contains the exact value of the parameter. Instead of taking an approximated value for the unknown parameter, we show how we can use the monotonicity properties of the Markov chain to take into account the error bound from the measurements. In the second application, we consider a well known approximation method: the decomposition into submodels. In such an approach, models of complex networks are decomposed into submodels whose results are then used as parameters for the next submodel in an iterative computation. One obtains a fixed point system which is solved numerically. In general, we have neither an existence proof of the solution of the fixed point system nor a convergence proof of the iterative algorithm. Here we show how stochastic monotonicity can be used to answer these questions. Furthermore, monotonicity properties can also help to derive more efficient algorithms to solve fixed point systems.
6.2.4.2. Componentwise Bounds

In collaboration with Jean-Michel Fourneau [PRISM, Université de Versailles Saint-Quentin] we proposed an iterative algorithm to compute component-wise bounds of the steady-state distribution of an irreducible and aperiodic Markov chain [17]. These bounds are based on very simple properties of $(\max, +)$ and $(\min, +)$ sequences. We showed that, under some assumptions on the Markov chain, these bounds converge to the exact solution. In that case we have a clear tradeoff between computation and the tightness of bounds. Furthermore, at each step we know that the exact solution is within an interval, which provides a more effective convergence test than usual iterative methods.

6.2.4.3. Markov Reward Processes and Aggregation

In a joint work with I.M. H. Vliegen [University of Twente, The Netherlands] and A. Scheller-Wolf [Carnegie Mellon University, USA] [19], we presented a new bounding method for Markov chains inspired by Markov reward theory: Our method constructs bounds by redirecting selected sets of transitions, facilitating an intuitive interpretation of the modifications of the original system. We show that our method is compatible with strong aggregation of Markov chains; thus we can obtain bounds for an initial chain by analyzing a much smaller chain. We illustrated our method by using it to prove monotonicity results and bounds for assemble-to-order systems.

6.2.4.4. Critical Level Policies in Controlled Queuing Systems

In a joint work with Emmanuel Hyon [University of Paris Ouest Nanterre La Defense and LIP6] [39], we consider a single-item lost sales inventory model with different classes of customers. Each customer class may have different lost sale penalty costs. We assume that the demands follow a Poisson process and we consider a single replenishment hypoexponential server. We give a Markov decision process associated with this optimal control problem and prove some structural properties of its dynamic programming operator. This allows us to show that the optimal policy is a critical level policy. We also discuss some possible extensions to other replenishment distributions and give some numerical results for the hyperexponential server case.

6.2.5. Dynamic Systems with Local Interactions

Dynamic systems with local interactions can be used to model problems in distributed computing: gathering a global information by exchanging only local information. The challenge is two-fold: first, it is impossible to centralize the information (cells are indistinguishable); second, the cells contain only a limited information (represented by a finite alphabet $\mathcal{A}; \mathcal{A} = \{0, 1\}$ in our case). Two natural instantiations of dynamical systems are considered, one with synchronous updates of the cells, and one with asynchronous updates. In the first case, time is discrete, all cells are updated at each time step, and the model is known as a Probabilistic Cellular Automaton (PCA) (e.g. Dobrushin, R., Kryukov, V., Toom, A.: Stochastic cellular systems: ergodicity, memory, morphogenesis, 1990). In the second case, time is continuous, cells are updated at random instants, at most one cell is updated at any given time, and the model is known as a (finite range) Interacting Particle System (IPS) (e.g. Liggett, T.M.: Interacting particle systems, 2005).

6.2.5.1. Density Classification on Infinite Lattices and Trees

In a joint work with N. Fatès [INRIA Nancy – Grand-Est], J. Mairesse and I. Marcovici [LIAFA, CNRS and Université Paris 7] [46] we consider an infinite graph with nodes initially labeled by independent Bernoulli random variables of parameter $p$. We address the density classification problem, that is, we want to design a (probabilistic or deterministic) cellular automaton or a finite-range interacting particle system that evolves on this graph and decides whether $p$ is smaller or larger than $1/2$. Precisely, the trajectories should converge (weakly) to the uniform configuration with only $0$'s if $p < 1/2$, and only $1$'s if $p > 1/2$. We present solutions to that problem on $\mathbb{Z}^d$, for any $d \geq 2$, and on the regular infinite trees. For $\mathbb{Z}$, we propose some candidates that we back up with numerical simulations.

6.2.6. Stochastic Stability

6.2.6.1. Ergodicity of Probabilistic Cellular Automata
In a joint work with J. Mairesse and I. Marcovici [LIAFA, CNRS and Université Paris 7] [31], we considered ergodicity properties of probabilistic cellular automata (PCA). A classical cellular automaton (CA) is a particular case of PCA. For a 1-dimensional CA, we proved that ergodicity is equivalent to nilpotency, and is therefore undecidable. We then proposed an efficient perfect sampling algorithm for the invariant measure of an ergodic PCA. Our algorithm does not assume any monotonicity properties of the local rule. It is based on a bounding process which is shown to be also a PCA. We then focused on the PCA Majority, whose asymptotic behavior is unknown, and performed numerical experiments using the perfect sampling procedure.

6.2.6.2. Spatial Queues

In a joint work with S. Foss [Heriot–Watt University, UK] [13], we considered a queue where the server is the Euclidean space, the customers are random closed sets of the Euclidean space arriving according to a Poisson rain and where the discipline is a hard exclusion rule: no two intersecting random closed sets can be served at the same time. We use the max plus algebra and Lyapunov exponents to show that under first come first serve assumptions, this queue is stable for a sufficiently small arrival intensity. We also discuss the percolation properties of the stationary regime of the random closed sets in the queue.

6.3. Economics of Networks

Participants: François Baccelli, Emilie Coupechoux, Marc Lelarge.

6.3.1. Diffusion and Cascading Behavior in Random Networks

The spread of new ideas, behaviors or technologies has been extensively studied using epidemic models. In [69], we considered a model of diffusion where the individuals’ behavior is the result of a strategic choice. We studied a simple coordination game with binary choice and give a condition for a new action to become widespread in a random network. We also analyze the possible equilibria of this game and identify conditions for the coexistence of both strategies in large connected sets. Finally we look at how can firms use social networks to promote their goals with limited information.

Our results differ strongly from the one derived with epidemic models. In particular, we showed that connectivity plays an ambiguous role: while it allows the diffusion to spread, when the network is highly connected, the diffusion is also limited by high-degree nodes which are very stable. In the case of a sparse random network of interacting agents, we computed the contagion threshold for a general diffusion model and showed the existence of (continuous and discontinuous) phase transitions. We also computed the minimal size of a seed of new adopters in order to trigger a global cascade if these new adopters can only be sampled without any information on the graph. We showed that this minimal size has a non-trivial behavior as a function of the connectivity. Our analysis extends methods developed in the random graphs literature based on the properties of empirical distributions of independent random variables, and leads to simple proofs.

6.3.2. Impact of Clustering on Diffusions and Contagions in Random Networks

In [33] we extend some results of the previous results to a model of random graphs having both a given degree distribution and a tunable clustering coefficient. This work shed new light on the impact of clustering on the spread of new ideas, technologies, viruses or worms. We consider two types of growth processes: the (classical SI) diffusion model, and the contagion model, which is inspired by a simple coordination game played on the network and is characterized by a threshold rule and a random seed. While clustering inhibits the diffusion process (on regular graphs), its impact for the contagion process is more subtle and depends on the connectivity of the graph: in a low connectivity regime, clustering also inhibits the contagion, while in a high connectivity regime, clustering favors the appearance of global cascades but reduces their size.

6.3.3. Economic Value of User Localization in Wireless Networks

The defining characteristic of wireless and mobile networking is user mobility, and related to it is the ability for the network to capture (at least partial) information on where users are located and how users change location over time. Information about location is becoming critical, and therefore valuable, for an increasingly larger number of location-based or location-aware services. A key open question, however, is how valuable exactly this information is. Our goal in this paper is to help understand and estimate the economics, or the value of location information.
In a joint work with J. Bolot [Sprint ATL, USA], [25], we addressed in particular the value of different granularities of location information, for example how much more valuable is it to know the GPS location of a mobile user compared to only knowing the access point, or the cell tower, that the user is associated with. We made three main contributions. First, we presented novel models, which capture the location-based economic activity of mobile users. Second, we derived closed-form analytic solutions for the economic value generated by those users. Third, we augmented the models to consider uncertainty about the users’ location, and derived expressions for the economic value generated with different granularities of location information.

6.4. Point Processes, Stochastic Geometry and Random Geometric Graphs

Participants: François Baccelli, Bartłomiej Blaszczyszyn, Pierre Brémaud, Yogeshwaran Dhandapani, Kumar Gaurav, Mir Omid Haji Mirsadeghi, Justin Salez.

6.4.1. Comparison of Clustering and Percolation of Point Processes and Random Graphs

Heuristics indicate that point processes exhibiting clustering of points have larger critical radius \( r_c \) for the percolation of their continuum percolation models than spatially homogeneous point processes. It has already been shown in [64], [73] that the directionally convex \((dcx)\) ordering of point processes is suitable to compare their clustering tendencies. Hence, it was tempting to conjecture that \( r_c \) is increasing in \( dcx \) order. Some numerical evidences support this conjecture for a special class of point processes, called perturbed lattices, which are "toy models" for determinantal and permanental point processes. However the conjecture is not true in full generality. In 2011 we have prepared three publications on this subject.

6.4.1.1. On comparison of clustering properties of point processes

In [52] we provide a large class of perturbed lattice point processes, monotone in \( dcx \) order and comparable to Poisson point processes that is commonly considered as the reference model in the comparative study of clustering phenomena. We also introduce a weaker order based on the comparison of only void probabilities and factorial moment measures. We prove that determinantal and permanental processes, as well as, more generally, negatively and positively associated point processes are comparable in this weaker sense to the Poisson point process of the same mean measure.

6.4.1.2. Clustering and percolation of point processes

In [49] we show that simple, stationary point processes of a given intensity on \( \mathbb{R}^d \), having void probabilities and factorial moment measures smaller than those of a homogeneous Poisson point process of the same intensity, admit uniformly non-degenerate lower and upper bounds on the critical radius \( r_c \) for the percolation of their continuum percolation models. Examples are negatively associated point processes and, more specifically, determinantal point processes. More generally, we show that point processes \( dcx \) smaller than a homogeneous Poisson point processes (for example perturbed lattices) exhibit phase transitions in certain percolation models based on the level-sets of additive shot-noise fields of these point processes. Examples of such models are \( k \)-percolation and SINR-percolation models. We also construct a Cox point process with degenerate critical radius \( r_c = 0 \), that is \( dcx \) larger than a given homogeneous Poisson point process. This is a counterexample for the aforementioned conjecture in the full generality.

6.4.1.3. Ordering of non-standard critical radii

As explained above, heuristically one expects finiteness of the critical radii for percolation of sub-Poisson point processes. However, in [49] we have show that it is non-zero as well. In a more elaborate paper [50] we present a reasoning as to why this non-triviality is to be expected. Specifically, we defined two (non-standard) critical radii for percolation of the Boolean model, called the lower and upper critical radii, and related, respectively, to the finiteness of the expected number of void circuits around the origin and asymptotic of the expected number of long occupied paths from the origin in suitable discrete approximations of the continuum model. These radii sandwich the usual critical radius \( r_c \) for percolation of the Boolean model. We show that \( dcx \) order preserves the upper critical radii and reverses the lower critical radii.
6.4.1.4. Local weak convergence and stochastic comparison

Many random models are parametrized by the size of the model, and the essential properties of the model are the asymptotic ones as the size of the graph tends to infinity. In the master thesis [57], we show that the theory of local weak convergence provides a natural setting to investigate stochastic (convex) ordering of such models. We consider both the geometric context of [71] and the discrete one of Galton-Watson branching process and Configuration Model, cf. [5]. In this latter case we define and study a convex order in the context of random trees and graphs which converge in the local weak sense. In particular, we’re interested in the effect of ordering on percolation. It turns out that while in the case of Galton-Watson trees, convex ordering leads to the ordering of percolation probabilities, we cannot conclude this in the case of configuration model. In this case, we could only obtain the ordering of percolation thresholds.

6.4.1.5. AB random geometric graphs

We investigated percolation in the AB Poisson-Boolean model in d-dimensional Euclidean space, and asymptotic properties of AB random geometric graphs on Poisson points in [0, 1]^d. The AB random geometric graph we studied is a generalization to the continuum of a bi-directed graph called the AB percolation model on discrete lattices. Such an extension is motivated by applications to secure communication networks and frequency division duplex networks. The AB Poisson Boolean model is defined as a bi-directed graph on two independent Poisson point processes of intensities \( \lambda \) and \( \mu \) in the d-dimensional Euclidean space in the same manner as the usual Boolean model with a radius \( r \). We showed existence of AB percolation for all \( d \geq 2 \), and derived bounds for a critical intensity. Further, in \( d = 2 \), we characterize a critical intensity. The set-up for AB random geometric graphs is to construct a bi-directed graph on two independent Poisson point process of intensities \( n \) and \( c n \) in the unit cube. We provided almost sure asymptotic bounds for the connectivity threshold for all \( c > 0 \) and a suitable choice of radius cut-off functions \( r_n(c) \). Further for \( c < c_0 \), we derived a weak law result for the largest nearest neighbor radius. This work, which was a part of the PhD thesis [73] will appear in [23].

6.4.2. Random Packing Models

Random packing models (RPM) are point processes (p.p.s) where points which "contend" with each other cannot be simultaneously present. These p.p.s play an important role in many studies in physics, chemistry, material science, forestry and geology. For example, in microscopic physics, chemistry and material science, RPMs can be used to describe systems with hard-core interactions. Applications of this type range from reactions on polymer chains, chemisorption on a single-crystal surface, to absorption in colloidal systems. In these models, each point (molecule, particle, \( \cdots \)) in the system occupies some space, and two points with overlapping occupied space contend with each other. Another example is the study of seismic and forestry data patterns, where RPMs are used as a reference model for the data set under consideration. In wireless communications, RPMs can be used to model the users simultaneously accessing the medium in a wireless network using Carrier Sensing Medium Access (CSMA). In this context, each point (node, user, transmitter, \( \cdots \)) does not occupy space but instead generates interference to other points in the network. Two points contend with each other if either of them generates too much interference to the other. Motivated by this kind of application, we studied in [70] the generating functionals of several models of random packing processes: the classical Matérn hard-core model; its extensions, the \( k \)-Matérn models and the \( \infty \)-Matérn model, which is an example of random sequential packing process. The main new results are: 1) A sufficient condition for the \( \infty \)-Matérn model to be well-defined (unlike the other two, the \( \infty \)-Matérn model may not be well-defined on unbounded space); 2) the generating functional of the resulting point process which is given for each of the three models as the solution of a differential equation; 3) series representation and bounds on the generating functional of the packing models; 4) moment measures and other useful properties of the considered packing models which are derived from their generating functionals.

6.4.3. Extremal and Additive Matérn Point Processes

In the simplest Matérn point processes, one retains certain points of a Poisson point process in such a way that no pairs of points are at distance less than a threshold. This condition can be reinterpreted as a threshold condition on an extremal shot–noise field associated with the Poisson point process. In a joint work with P.
Bermolen (Universidad de la República, Montevideo, Uruguay) [11], we studied extensions of Matérn point processes where one retains points that satisfy a threshold condition based on an additive shot–noise field of the Poisson point process. We provide an analytical characterization of the intensity of this class of point processes and we compare the packing obtained by the extremal and additive schemes and certain combinations thereof.

6.4.4. Spatial Birth and Death Point Processes

In collaboration with F. Mathieu [INRIA GANG] and Ilkka Norros [VTT, Finland], we started studying a new spatial birth and death point process model where the death rate is a shot noise of the point configuration [60]. We showed that the spatial point process describing the steady state exhibits repulsion. We studied two asymptotic regimes: the fluid regime and the hard–core regime. We derived closed form expressions for the mean (and in some cases the law) of the latency of points as well as for the spatial density of points in the steady state of each regime.

6.4.5. Information Theory and Stochastic Geometry

In a joint work with V. Anantharam [UC Berkeley], [58], we studied the Shannon regime for the random displacement of stationary point processes. Let each point of some initial stationary point process in $n$-dimensional Euclidean space give rise to one daughter point, the location of which is obtained by adding a random vector to the coordinates of the mother point, with all displacement vectors independently and identically distributed for all points. The decoding problem is then the following one: the whole mother point process is known as well as the coordinates of some daughter point; the displacements are only known through their law; can one find the mother of this daughter point? The Shannon regime is that where the dimension $n$ tends to infinity and where the logarithm of the intensity of the point process is proportional to $n$. We showed that this problem exhibits a sharp threshold: if the sum of the proportionality factor and of the differential entropy rate of the noise is positive, then the probability of finding the right mother point tends to 0 with $n$ for all point processes and decoding strategies. If this sum is negative, there exist mother point processes, for instance Poisson, and decoding strategies, for instance maximum likelihood, for which the probability of finding the right mother tends to 1 with $n$. We then used large deviations theory to show that in the latter case, if the entropy spectrum of the noise satisfies a large deviation principle, then the error probability goes exponentially fast to 0 with an exponent that is given in closed form in terms of the rate function of the noise entropy spectrum. This was done for two classes of mother point processes: Poisson and Matérn. The practical interest to information theory comes from the explicit connection that we also establish between this problem and the estimation of error exponents in Shannon’s additive noise channel with power constraints on the codewords.

We currently investigate extensions of this approach to network information theoretic channels.

6.4.6. Navigation on Point Processes and Graphs

In [12], we studied optimal navigations in wireless networks in terms of first passage percolation on some space-time SINR graph. We established both “positive” and “negative” results on the associated the percolation delay rate (delay per unit of Euclidean distance, also called time constant in the classical terminology of percolation). The latter determines the asymptotics of the minimum delay required by a packet to progress from a source node to a destination node when the Euclidean distance between the two tends to infinity. The main negative result states that the percolation delay rate is infinite on the random graph associated with a Poisson point process under natural assumptions on the wireless channels. The main positive result states that when adding a periodic node infrastructure of arbitrarily small intensity to the Poisson point process, the percolation delay rate is positive and finite.

A new direction of research was initiated aiming at defining a new class of measures on a point process which are invariant under the action of a navigation on this point process. This class of measures has properties similar to Palm measures of stationary point processes; but they cannot be defined in the classical framework of Palm measures.

6.5. Random Graphs and Combinatorial Optimization

Participants: Hamed Amini, Emilie Coupechoux, Mathieu Leconte, Marc Lelarge, Justin Salez.
6.5.1. Rank of Large Random Graphs

In [16], with Charles Bordenave [CNRS-Université de Toulouse], we investigated the rank of the adjacency matrix of large diluted random graphs: for a sequence of graphs converging locally to a Galton-Watson tree, we provided an explicit formula for the asymptotic multiplicity of the eigenvalue 0 in terms of the degree generating function. In the first part, we showed that the adjacency operator associated with a Galton-Watson tree is self-adjoint with probability one; we analyzed the associated spectral measure at the root and characterize the distribution of its atomic mass at 0. In the second part, we established a sufficient condition for the expectation of this atomic mass to be precisely the normalized limit of the dimension of the kernel of the adjacency matrices of the sequence of graphs. Our proofs borrow ideas from analysis of algorithms, functional analysis, random matrix theory, and statistical physics.

6.5.2. Matchings in infinite graphs

In [43], we proved that for any sequence of (deterministic or random) graphs converging locally, the corresponding sequence of normalized matching numbers converges, and this limit depends only on the limit of the graph sequence. In the particular case where this limit is a unimodular Galton Watson tree, we were able to compute explicitly the value for the limit of the sequence of (normalized) matching numbers. This leads to an explicit formula that considerably extends the well-known one by Karp and Sipser for Erdős-Rényi random graphs.

We considered a natural family of Gibbs distributions over matchings on a finite graph, parameterized by a single positive number called the temperature. The correlation decay technique can be applied for the analysis of matchings at positive temperature and allowed us to establish the weak convergence of the Gibbs marginal as the underlying graph converges locally. However for the zero temperature problem (i.e. maximum matchings), we showed that there is no correlation decay even in very simple cases. By using a complex temperature and a half-plane property due to Heilmann and Lieb, we were able to let the temperature tend to zero and obtained a limit theorem for the asymptotic size of a maximum matching in the graph sequence.

6.5.3. Counting spanning subgraphs subject to local constraints

In [53], we use negative association and local weak convergence to establish the validity of the cavity method for counting spanning subgraphs subject to local constraints. Specifically, the normalized logarithm of the associated generating polynomial (or partition function) is shown to converge along any sequence of graphs whose random weak limit is a tree, and the limit is directly expressed in terms of the unique solution to a limiting cavity equation. On a Galton-Watson tree, the latter simplifies into a recursive distributional equation which can be solved explicitly. As an illustration, we provide an asymptotic formula for the maximal size of a spanning subgraph with maximal degree $b$ in the Erdős-Rényi model with fixed average degree and diverging size, for any $b \in \mathbb{N}$.

6.5.4. Bipartite graph structures for efficient balancing of heterogeneous loads

With Laurent Massoulié [Technicolor], we extend the results obtained previously on the asymptotic size of maximum matchings in random graphs converging locally to Galton-Watson trees to so-called $b$-matchings (with non-unitary capacity at vertices as well as constraints on individual edges). Compared to the matching case, this involves studying the convergence of a message passing algorithms which transmits vectors instead of single real numbers. We also look further into an application of these results to large scale distributed content service platforms, such as peer-to-peer video-on-demand systems. In this context, the density of maximum $b$-matchings corresponds to the maximum fraction of simultaneously satisfiable requests, when the service resources are limited and each server can only handle requests for a predetermined subset of the contents which it has stored in memory. An important design aspect of such systems is the content placement strategy onto the servers depending on the estimated content popularities; the results obtained allow to characterize the efficiency of such placement strategies and the optimal strategies in the limit of large storage capacity at servers are determined.
6.5.5. Flooding in Weighted Random Graphs

In a joint work [24] with Moez Draief [Imperial College London], we studied the impact of the edge weights on distances in diluted random graphs. We interpret these weights as delays, and take them as i.i.d exponential random variables. We analyzed the edge flooding time defined as the minimum time needed to reach all nodes from one uniformly chosen node, and the edge diameter corresponding to the worst case edge flooding time. Under some regularity conditions on the degree sequence of the random graph, we showed that these quantities grow as the logarithm of $n$, when the size of the graph $n$ tends to infinity. We also derived the exact value for the prefactors.

These allowed us to analyze an asynchronous randomized broadcast algorithm for random regular graphs. Our results show that the asynchronous version of the algorithm performs better than its synchronized version: in the large size limit of the graph, it will reach the whole network faster even if the local dynamics are similar on average.
6. New Results

6.1. Model Driven and Aspect Oriented Design

6.1.1. Requirements Engineering

Participants: Olivier Barais, Benoît Baudry, Benoît Combemale, Maha Driss, Jean-Marc Jézéquel, Emmanuelle Rouillé, Nicolas Sannier, Didier Vojtisek.

Model-driven engineering can have a huge impact on the early design and analysis of complex systems. We have investigated modeling for requirements engineering in three areas:

- We use executable metamodeling techniques developed in the team to capture formal relationships between regulatory requirements and accepted practices in systems engineering [47], [42].
- We propose an approach for facilitating Web service selection according to user requirements. These requirements specify the needed functionality and expected QoS, as well as the composable between each pair of services. The originality of our approach is embodied in the use of Formal Concept Analysis (FCA) and its extension Relational Concept Analysis (RCA) [33] [25].
- We have analyzed a real industrial software process to illustrate the need for bridging the gap between software processes and software development tools to automate the development tools configuration, deployment, integration and adaptation [46].

6.1.2. Dynamically adaptive interactive systems

Participants: Arnaud Blouin, Jean-Marc Jézéquel, Grégory Nain.

Combining Aspect-Oriented Modeling with Property-Based Reasoning to Improve User Interface Adaptation: in this work we combined aspect-oriented modeling with an interactive system architecture to support dynamic adaptation of interactions and user interfaces [28].

6.1.3. Dynamically adaptive component-based systems

Participants: François Fouquet, Olivier Barais, Viet-Hoa Nguyen, Noël Plouzeau.

Continuous Design to Achieve Intelligent Reflection in Distributed Systems: we defined an intelligent reflection model to support fast adaptation of distributed systems by architecture modification without stopping the system. This adaptation mechanism is well suited to rapidly changing needs (continuous design of eternal systems) or fast paced modifications of the context of the running system (for instance for Internet of Things distributed systems) [34].

6.1.4. Architecture for Services-based applications

Participants: Olivier Barais, Johann Bourcier, Erwan Daubert, Jean-Marc Jézéquel.

The architecture of service-based applications can have a huge impact on their dynamic adaptability. We have investigated various framework for architecting service-based applications:

- Designing SAFDIS: a self adaptive framework for distributed applications based on services. SAFDIS includes facilities to support the coordination of distributed reconfigurations [24]. SAFDIS also takes benefit of the Infrastructure As A Service to dynamically reconfigure Software As A Service [44].
- Analyzing and improving consistency between functional and business view of telecom services architecture. This work is based on the definition of a strategic alignment of the target functional view with the target business view. Alignment is validated with a real case study implemented and deployed at Orange–France Telecom on their messaging service [22].
Designing AutoHome: a service oriented framework to simplify the development and runtime adaptive support of autonomic pervasive applications. This includes the amalgamation of the two computing areas of Autonomics and Service Orientation, to produce a Component-based platform providing facilities. This infrastructure uniquely blends the advantages of distributed autonomic control with global conflict management in a management hierarchy [17].

6.2. Model V&V and Testing

6.2.1. Formal MDE Foundations
**Participants:** Benoit Baudry, Benoit Combemale.

- Formally Tracing Executions From an Analysis Tool Back to a Domain Specific Modeling Language’s Operational Semantics: in this work, we propose a formal and operational framework for tracing results back (e.g., a program crash log, or a counterexample returned by a model checker) from execution and verification tools to an original DSML’s syntax and operational semantics [31].
- A Proof Assistant Based Formalization of components in MDE: using the Coq proof assistant we propose a formalization of some operators for model fragment extraction and composition, as defined in the ReuseWare toolset [39].
- We have developed a methodology to explicitly model the context in which a temporal property must be verified. This contextual information is expressed in the requirements, and an explicit model allows to reduce the complexity of automated verification [41].

6.2.2. Pairwise testing for highly variable systems
**Participants:** Benoit Baudry, Aymeric Hervieu.

Variability management is a major concern for the development of software intensive systems. In particular, the explosion of variants is an issue for testing and analysis. Feature models allow to explicitly capture the variability in a formal model and get a complete view on all possible variants of the system. We have investigated pair-wise generation from feature models in order to test software product lines [36], and to evaluate QoS contracts in variable web service compositions [38].

6.2.3. Testing aspect-oriented programs
**Participant:** Benoit Baudry.

Aspect-oriented mechanisms introduce new risks for reliability that must be tackled by specific testing techniques in order to fully benefit from the use of this paradigm. We have investigated a monitoring mechanism of advices in an aspect-oriented program and use this information to build test cases that target faults in pointcut descriptors [18].

6.2.4. Modeling model quality metrics
**Participants:** Benoit Baudry, Jean-Marc Jézéquel.

We have developed a model-driven measurement approach to measure models of a domain specific modeling language. The approach uses models as unique and consistent metric specifications for the automated generation of a metric tool. The benefit from applying the approach is evaluated by four case studies [20]. In particular, we have evaluated the ability of the approach to build a tool for the measurement of requirements documents [21].
6.3. Meta-Modeling

6.3.1. Model Driven Language Engineering

Participants: Benoit Baudry, Arnaud Blouin, Juan-Jose Cadavid Gomez, Benoit Combemale, Clément Guy, Jean-Marc Jézéquel, Didier Vojtisek.

- Model-Driven Engineering and Optimizing Compilers: A bridge too far? In this work, we report and analyze an experience about the use of MDE technologies to build and evolve compiler infrastructures in the optimizing compiler domain. From this study, we highlight challenges and propose a roadmap for the cross-fertilization of the MDE and compiler domains [35], [45].
- Modeling Model Slicers: model slicing is a model operation that consists in extracting a subset of a model. Because the creation of a new DSL implies the creation from scratch of a new model slicer, we proposed the Kompren language that models and generates model slicers for any DSL [27].
- Empirical Evaluation of the Conjunct Use of MOF and OCL: we evaluate in this work the conjunct usage of MOF (Meta-Object Facility) and OCL (Object Constraint Language) in the development of Domain-Specific Modeling Languages. We observe the state of practice to understand how experts use them and find patterns on its usage, in order to provide techniques to improve the experience [29].

6.3.2. Model Transformation and Composition

Participants: Olivier Barais, Benoit Baudry, Arnaud Blouin, Mickaël Clavreul, Benoit Combemale, Xavier Dolques, Jean-Marc Jézéquel.

- Model operations such as transformation and composition declare source metamodels that are usually larger than the set of concepts and relations actually used by the operation. We have proposed and validated a static operation analyzer to retrieve the metamodel footprint of the operation [37].
- Service-Oriented Architecture Modeling: Bridging the Gap between Structure and Behavior: In this approach, we propose to detect divergences among structural and behavioral models to support a semi-automatic process of synchronization between class diagrams and workflow models [30].
- The paper propose a technique for discovering matchings between two model elements modeling the same system, but being instances of different metamodels. This is achieved by using property names and models structure thanks to the adaptation of a schema matching technique named Anchor-PROMPT [32].
- Specifying and implementing UI Data Bindings with Active Operations: based on the concept of active operations, this work proposes a framework to bind models at runtime and more precisely to bind data and their possible representations [26].
- We propose a requirement-centric approach for Web service composition which allows: (i) modeling users’ requirements with the MAP formalism and specifying required services using an Intentional Service Model (ISM); (ii) discovering and selecting relevant Web services and high QoS services; and (iv) generating automatically BPEL coordination processes by applying the model transformation technique [19].
6. New Results

6.1. Geometry Processing

**Participants:** Laurent Alonso, Alejandro Galindo, Phuong Ho, Samuel Hornus, Bruno Lévy, David Lopez, Kun Liu, Vincent Nivoliers, Nicolas Ray, Dmitry Sokolov, Rhaleb Zayer.

- **Sampling:** In the frame of the Goodshape project, we continued developing new techniques for sampling shapes optimally, based on the notion of Centroidal Voronoi Tessellations. We developed a new technique for computing clipped 3D Voronoi diagrams, suitable to volumetric meshing [23], and an accelerated GPU-based centroidal Voronoi tessellation [20]. We developed an optimization technique to suppress obtuse triangles [21]. We also studied periodic boundary conditions [29], suitable to some numerical simulations.

- **Geometric modeling and computational geometry:** Dobrina Boltcheva joined the team, with her expertise on simplicial homology [10]. Vincent Nivoliers published a L-system based knot insertion rules [19] that he developed during his master in Gipsa lab. We also studied furthest polygon Voronoi diagrams [15].

- We gave an invited course on geometry processing at SIGGRAPH Asia [24]

6.2. Computer Graphics

**Participants:** Laurent Alonso, Anass Lasram, Samuel Hornus, Bruno Jobard, Anass Lasram, Sylvain Lefebvre, Bruno Lévy, Vincent Nivoliers, Nicolas Ray.

- **Foundations of Computer Graphics:** We developed new algorithms and data structures for spatial caching and hashing on the GPU [16]. We continued our work on noise generation based on Gabor convolution and proposed several improvements [17] (see Figure 5). We also studied the fundamental problem of interpolating functions and developed an algorithm based on optimal transport theory [11].

- **Applications:** We developed an algorithm to change the lighting in images that contain trees [12] (with REVES, see Figure 6).
Figure 6. Relighting photos of canopies.

Figure 7. Unified visualization of atomic bounds and molecular surfaces.
Figure 8. Left: flow lines between injectors (black squares) and well (center square). Right: Adapted Voronoi grid.

Figure 9. Fitting a spline surface (bottom) to a triangle mesh (top) starting from an initial template (center). Data courtesy of Distene.
6.3. Scientific Visualization and scientific computing

Participants: Samuel Hornus, Bruno Jobard, Bruno Lévy, Romain Merland, Vincent Nivoliers, Jeanne Pellerin, Nicolas Ray, Dmitry Sokolov, Raleb Zayer.

- **Molecular visualization:** We continued the development of our Micromegas software, and developed several improvements [13], [14] (see Figure 7).

- **Geo-Modeling and geo-visualization:** In the frame of our partnership with the Gocad consortium, we developed an evaluation of multi-valued data depiction techniques [22]. We also developed several meshing tools dedicated to the numerical simulation of oil exploitation [28], [26], [25] (see Figure 8).

- **Reverse engineering:** We developed methods to convert mesh surfaces and point sets into parametric splines, using either a global parameterization technique [18] or an approximation of surface-to-surface distance based on Voronoi diagrams [27] (see Figure 9).
ALPAGE Project-Team

6. New Results

6.1. Advances in symbolic parsing with DyALog/FRMG

Participant: Éric Villemonte de La Clergerie.

Within the team is developed a wide-coverage French meta-grammar (FRMG) and an efficient hybrid TAG/TIG parser based on the DyALog logic programming environment [127] and on the Lefff morphological and syntactic lexicon [118]. It relies on the notion of factorized grammar, themselves generated from a representation that lies at a higher level of abstraction, named Meta-Grammars [129]. At that level, linguistic generalizations can be expressed, which in turn makes it possible to transfer meta-grammars from one language to a closely related one. The hybrid TAG/TIG parser generator itself implements all kinds of parsing optimizations: lexicalization (in particular via hypertags), left-corner guiding, top/bottom feature analysis, TIG analysis (with multiple adjoining), and others.

Éric de La Clergerie has continued to improve the coverage, quality and efficiency of the French meta-grammar FRMG. On the EasyDev corpus (around 4000 sentences), parsing times have improved over 2011 from an average of 1.03s per sentence to 0.28s, coverage (in terms of sentences with full parses) has improved from 72.5% to 82.60%, and accuracy (in terms of f-measure over relations) from 64.54% to 68.28%.

A part of the accuracy gains comes from the addition of a new output format for FRMG, namely the CONLL format, allowing us to use the CONLL-based dependency version of the French Treebank (around 12K sentences) for training and evaluation. We also used new machine learning techniques to improve FRMG’s disambiguation algorithm, allowing us to combine heuristic based disambiguation rules (with manually provided weights) with more standard parsing features associated with automatically learned weights. More precisely, the idea was to study the efficiency of the disambiguation rules over the French treebank and to favor (resp. penalize) well-working (resp. bad working) rules by adjusting their weight, taking into account additional (and more standard) features. Using these techniques, on ftb6_3 test part, FRMG improved from a base accuracy of 82.31% (in terms of CONLL Labeled Attachment Score) to 84.54%. These gains resulting from a training over the French TreeBank have also been observed (with however a lesser impact) on the EasyDev corpus (using a different format and using a different evaluation metric).

6.2. Task-based evaluation of syntactic lexica: coupling FRMG with various resources

Participants: Éric Villemonte de La Clergerie, Benoît Sagot.

The FRMG symbolic parser was used for comparing the performances of various syntactic lexicons as sources of information for parsing. The idea is to convert syntactic lexica other than the Lefff into the Lefff’s format, i.e., turn them into Alexina lexicons, and then use the resulting lexica together with the FRMG grammar for producing several parsers. These parsers only differ by the lexical information they rely on. Preliminary results had already been obtained in 2009 [119], but were restricted to one external lexicon, namely Lexicon-Grammar tables, and only to verbal entries (other entries were gathered from the Lefff when using Lexicon-Grammar-based verbal entries). However, conversion tools for other resources, such as Dicovalence [136], had already been developed, in the context of the development and improvement of the Lefff. Moreover, the development of a new version of the Lefff verbal entries.
Task-based evaluation results have been obtained on parsing with FRMG, showing that the Lefff performs better than both Lexique-Grammaire and DICOVALENCE (after conversion to the Alexina formalism) [48], [49]. The new version of the Lefff, mentioned above, leads for now to lower results than the current version, but its results are better than with Lexique-Grammaire or DICOVALENCE data, despite a significant increase of the average amount of entries per lemma. These results are satisfying both because they show that the Lefff is a useful resource for symbolic parsing, but also because they illustrate the relevance of converting other resources into the Alexina formalism, in order to merge the valuable linguistic information they contain — as done in the last years for improving the Lefff [85], [84], [86], [111], [87], [112].

6.3. Information extraction from corpora parsed with FRMG

Participants: Yayoi Nakamura-Delloye, Rosa Stern, Éric Villemonte de La Clergerie, Benoît Sagot.

Following previous experiments, in particular in the context of the FUI-funded project Scribo that ended in 2010, work has been achieved for extracting information from corpora parsed with FRMG.

In the context of the EDyLex project, we have proposed two pattern-based named entity extraction methods for ontology enrichment [36], [35]. The proposed methods are characterized by the use of entity relation patterns obtained by our unsupervised extraction method. These patterns correspond to syntactic paths that connect two named entities in dependency trees produced by FRMG. This work aims to take advantage of parsing benefits and also offers solutions for parsing disadvantage. The proposed methods are characterized by the use of entity relation patterns obtained by our unsupervised extraction method. These patterns correspond to syntactic paths that connect two named entities in dependency trees. This work aims to take advantage of parsing benefits and also offers solutions for parsing disadvantage.

We also developed a mechanism for integrating the results into an domain ontology, namely the ontology under deployment at the Agence France-Presse [37].

6.4. Advances in statistical parsing

Participants: Marie Candito, Benoît Crabbé, Djamé Seddah, Enrique Henestroza Anguiano.

6.4.1. Improving statistical dependency parsing

Alpage has provided state-of-the art results for French statistical Parsing, adapting existing techniques for French, a richer morphological language than English, either for constituency parsing or dependency parsing. The Bonsai tool (see section 5.4) is available, that gathers preprocessing tools and models for dependency parsing French. We have innovated in the tuning of tagsets and the handling of unknown words. In the last years, Alpage has then contributed on four main points:

- conversion of the French Treebank [59] used as constituency training data into dependencies [72], the resulting treebank being used by several teams for dependency parsing;
- an original method to reduce lexical data sparseness and include coverage and robustness by replacing tokens by unsupervised word clusters or morphological clusters [69], [121], [73]; all of our morphological clustering approaches were integrated into our parsing chains; data driven lemmatization required the adaptation of a state-of-the-art part-of-speech tagger and lemmatizer (Morfette [77]) based on a data-driven joint model benefiting of the inclusion of external lexica such as the Lefff [121];
- a parser-agnostic postprocessing step, developed this year, which uses specialized models for dependency parse correction [30]; dependencies in an input parse tree are revised by selecting, for a given dependent, the best governor from within a small set of candidates, using a discriminative linear ranking model that includes a rich feature set that encodes syntactic structure in the input parse tree; the parse correction framework can correct attachments using either a generic model or specialized models tailored to difficult attachment types like coordination and pp-attachment; our
experiments have shown that parse correction, combining a generic model with specialized models for difficult attachment types, can successfully improve the quality of predicted parse trees output by several representative state-of-the-art dependency parsers for French.

- an adaptation of the above-mentioned technique of word clustering to the problem of adapting statistical parsers to different text domains [25]. We show that in order to parse texts from a different domain than the one a statistical parser is trained on (namely to parse target domain text using a parser trained on indomain treebank), word clusters computed over a bridge corpus that couples indomain and target domain raw texts do improve parsing performance on target domain, without degrade performance on indomain texts (contrary to previous domain adaptation techniques). To evaluate these experiments, we use as target domain biomedical texts. We have supervised the manual syntactic annotation of a test corpus from the biomedical domain (European Public Assessment Reports concerning the marketing authorization of medicinal products).

Besides this line of work, it should be noted that two parsing models built around Stochastic Tree Insertion Grammars are currently under investigation: experiments have been conducted on Spinal TIGs [122]. Moreover, we are still improving the TIG-based dependency parser MICA, developed in collaboration with University of Marseilles, Columbia university and AT&T [61] (see section 5.5).

6.4.2. Functional labelling

Alpage worked towards the improvement of a functional labeller to be used as a post-parsing tool on an unfolded parse forest (as outputted e.g. by the Berkeley parser in the Bonsai architecture) using CRF models of various orders thereby extending the previous maximum entropy labeller designed in the team. The use of CRFs for modelling triggered a collaboration with Isabelle Tellier and JP Prost (LIFO, Orleans). The labeller implementation has been considerably improved and the accuracy of the labeller has improved as well on correct treebank trees. However we found out that the feature engineering work outweighs the formal improvements since we were able to show that the use of higher order graphical models were not contributing significantly to improve an unstructured model. Our modest gains come mostly from feature engineering. Moreover we notice that combined with a constituent parser the labeller does not improve at all on constituent parsing output. The reason being that our current architecture for the Bonsai parser is sequential (which is unsatisfactory). Following experiments on n-best parsing outputs, we observe that the labeller can drastically improve on better parses where its input is indeed correct. This suggest investigating formulating constituent parsing and functional labelling as a joint task requiring to address serious efficiency issues. We intend to tackle the two drawbacks of our current architecture (sequential process, parse forest unfolding) by formulating constituent parsing as a joint task with functional labelling in the next few months.

6.4.3. Parsing spontaneous oral text

Alpage also got involved in parsing spontaneous oral text taken from ESTER 3 data (with overlaps) generated in the ANR ETAPE project in collaboration with A. Abeillé (LLF) with the aim of preannotating a seed for a future treebank of oral French which would considerably support work in experimental linguistics led in the Labex. He has also a collaboration set up with A. Abeillé, C. Gardent and C. Cerisara for ensuring interoperability across ongoing efforts for producing oral treebanks for French. The way to carry out the task was by using a form of preprocessing of oral text to simulate a written entry to the Bonsai parser trained on written text. In the next few months we intend to test semi-supervised learning techniques to speed up the annotation process made by the LLF lab.

6.5. Named Entity Recognition and Entity Linking

Participants: Rosa Stern, Benoît Sagot.
Identifying named entities is a widely studied issue in Natural Language Processing, because named entities are crucial targets in information extraction or retrieval tasks, but also for preparing further NLP tasks (e.g., parsing). Therefore a vast amount of work has been published that is dedicated to named entity recognition, i.e., the task of identification of named entity mentions (spans of text denoting a named entity), and sometimes types. However, real-life applications need not only identify named entity mentions, but also know which real entity they refer to; this issue is addressed in tasks such as knowledge base population with entity resolution and linking, which require an inventory of entities is required prior to those tasks in order to constitute a reference.

6.5.1. Improvements of the Aleda entity database

Within the Alexina framework, we develop since 2012 the entity database Aleda [124], aimed at constituting such a reference. Aleda was first developed for French but is under development for English. Aleda is extracted automatically from Wikipedia and Geonames. It is used among others in the SxPipe processing chain and its NP named entity recognition, as well as in the Nomos named entity linking system.

In 2011, major efforts have been made for improving the coverage, precision and richness of the French Aleda: improvements in the tool for creating an XML almost-raw-text version of the wikipedia, new method for identifying and typing entities among wikipedia articles, based on infoboxes and wikipedia categories, richer database structure for storing more detailed information about each entity, and many other improvements. A paper about these advances has been submitted to LREC 2012.

6.5.2. Cooperation of symbolic and statistical methods for named entity recognition and typing

Named entity recognition and typing is achieved both by symbolic and probabilistic systems. We have performed an experiment [24] for making the rule-based system NP, SxPipe’s high-precision named entity recognition system developed at Alpage on AFP news corpora and which relies on the Aleda named entity database, interact with LIANE, a high-recall probabilistic system developed by Frédéric Béchet and trained on oral transcriptions from the ESTER corpus. We have shown that a probabilistic system such as LIANE can be adapted to a new type of corpus in a non-supervised way thanks to large-scale corpora automatically annotated by NP. This adaptation does not require any additional manual annotation and illustrates the complementarity between numeric and symbolic techniques for tackling linguistic tasks.

6.5.3. Nomos, a statistical named entity linking system

For information extraction from news wires, entities such as persons, locations or organizations are especially relevant in a knowledge acquisition context. Through a process of named entity recognition and entity linking applied jointly, we aim at the extraction and complete identification of these relevant entities, which are meant to enrich textual content in the form of metadata. In order to store and access extracted knowledge in a structured and coherent way, we aim at populating an ontological reference base with these metadata. We have pursued our efforts in this direction, using an approach where NLP tools have early access to Linked Data resources and thus have the ability to produce metadata integrated in the Linked Data framework. In particular, we have studied how the entity linking process in this task must deal with noisy data, as opposed to the general case where only correct entity identification is provided.

We use the symbolic named entity recognition system NP, a component of SxPipe, and use it as a mention detection module. Its output is then processed through our entity linking system, which is based on a supervised model learnt from examples of linkings. Since our named entity recognition is not deterministic, as opposed to other entity linking tasks where the gold named entity recognition results are provided, it is configured to remain ambiguous and non-deterministic, i.e., its output preserves a number of ambiguities which are usually resolved at this level. In particular, no disambiguation is made in the cases of multiple possible mentions boundaries (e.g., {Paris}+{Hilton} vs. {Paris Hilton}). In order to cope with possible false mention matches, which should be discarded as linking queries, the named entity recognition output is made more ambiguous by adding a not-an-entity alternative to each mention detection. The entity linking module’s input therefore consists in multiple possible readings of sentences. For each reading, this module must perform entity linking on every possible entity mention by selecting their most probable matching entity. Competing readings are
then ranked according to the score of entities (or sequence of entities) ranked first in each of them. The reading with no entity should also receive a score in order to be included in the ranking. The motivation for this joint task lies in the frequent necessity of accessing contextual and referential information in order to complete an accurate named entity recognition; thus the part where named entity recognition usually resolves a number of ambiguities is left for the entity linking module, which uses contextual and referential information about entities.

We have realized a first implementation of our system, as well as experiments and evaluation results. In particular, when using knowledge about entities to perform entity linking, we discuss the usefulness of domain specific knowledge and the problem of domain adaptation.

### 6.6. Extending wordnets

**Participants:** Benoît Sagot, Marianna Apidianaki, Valérie Hanoka.

The WOLF (see section 5.9) is a freely available, automatically created wordnet for French, the biggest drawback of which has until now been the lack of general concepts that are typically expressed with highly polysemous vocabulary that is on the one hand the most valuable for applications in human language technologies but also the most difficult to add to wordnet accurately with automatic methods on the other. In collaboration with Darja Fišer (University of Ljubljana), we have developed a self-training-like technique for acquiring a classifier that is able to assign appropriate synset ids (i.e., senses) to new words, extracted from non-disambiguated multilingual sources of lexical knowledge, such as Wiktionaries and Wikipedia [39], [40]. Automatic and manual evaluation shows high coverage as well as high quality of the resulting lexico-semantic repository. Another important advantage of the approach is that it is fully automatic and language-independent and can therefore be applied to any other language still lacking a wordnet. Indeed, it was applied to Slovene as well.

Other techniques were used as well and are the basis of various submitted conference papers. They rely, among others, on morphological derivation, on graph-based representation of highly multilingual lexicons extracted from numerous wiktionaries, and on automatically induced sense clusters.

### 6.7. Unsupervised lexical semantics

**Participant:** Marianna Apidianaki.

#### 6.7.1. Unsupervised word sense induction and disambiguation

Word sense induction (WSI) is the task aimed at automatically identifying the senses of words in texts, without the need for handcrafted resources or annotated data. Up till now, most WSI algorithms extract the different senses of a word ‘locally’ on a per-word basis, i.e. the different senses for each word are determined separately. In collaboration with Tim van de Cruys, at Alpage in 2010, now at University of Cambridge [19], [50], we have compared the performance of such algorithms to a new algorithm that uses a ‘global’ approach, i.e. the different senses of a particular word are determined by comparing them to, and demarcating them from, the senses of other words in a full-blown word space model. The induction step and the disambiguation step are based on the same principle: words and contexts are mapped to a limited number of topical dimensions in a latent semantic word space. The intuition is that a particular sense is associated with a particular topic, so that different senses can be discriminated through their association with particular topical dimensions; in a similar vein, a particular instance of a word can be disambiguated by determining its most important topical dimensions. We evaluated our model on the SemEval-2010 word sense induction and disambiguation task. All systems that participated in this task use a local scheme for determining the different senses of a word. We obtain state-of-the-art results.
6.7.2. Unsupervised cross-lingual lexical substitution

Cross-Lingual Lexical Substitution (CLLS) is the task that aims at providing for a target word in context several alternative substitute words in another language. The proposed sets of translations may come from external resources or be extracted from textual data. In 2011, we have introduced a new approach for this task [18], namely the use of an unsupervised cross-lingual word-sense induction method. This method identifies the senses of words by clustering their translations according to their semantic similarity. We evaluated the impact of using clustering information for CLLS on the SemEval-2010 CLLS data set. Our system performs better on the ’out-of-ten’ measure than the systems that participated in the SemEval task.


Participants: Pierre Magistry, Benoît Sagot.

For most languages using the Latin alphabet, tokenizing a text on spaces and punctuation marks is a good approximation of a segmentation into lexical units. Although this approximation hides many difficulties, they do not compare with those arising when dealing with languages that do not use spaces, such as Mandarin Chinese. Many segmentation systems have been proposed, some of them use linguistically motivated unsupervised algorithms. However, standard evaluation practices fail to account for some properties of such systems. New results [33] have shown that a simple model, based on an entropy-based reformulation of a language-independent hypothesis put forward by Harris in 1955, allows for segmenting a corpus and extracting a lexicon from the results. Tested on the Academia Sinica Corpus, our system allows for inducing a segmentation and a lexicon with good intrinsic properties and whose characteristics are similar to those of the lexicon underlying the manually-segmented corpus. Recent unpublished work using a slightly different model have improved these results. In parallel, preliminary experiments on other languages (Hindi, Singalese, Tamil, French) and original visualization techniques have already led to promising results.

6.9. Computational morphology

Participants: Benoît Sagot, Géraldine Walther.

Although computational morphology has been a topic of interest for Alpage for several years now, several new research topics have received attention in 2011, often in collaboration with morphologists from the Laboratoire de Linguistique Formelle (University Paris 7).

6.9.1. Inflectional morphology

Non-canonical inflection (suppletion, deponency, heteroclisis...) is extensively studied in theoretical approaches to morphology. However, these studies often lack practical implementations associated with large-scale lexica. Yet these are precisely the requirements for objective comparative studies on the complexity of morphological descriptions. We have shown [16], [43] how the Parsli model of inflectional morphology [132], which can represent many non-canonical phenomena, as well as a formalisation and an implementation thereof can be used to evaluate the complexity of competing morphological descriptions. After illustrating the properties of the model with data about French, Latin, Italian, Persian and Sorani Kurdish verbs and about noun classes from Croatian and Slovak we have conducted experiments on the complexity of four competing descriptions of French verbal inflection. The complexity is evaluated using the information-theoretic concept of description length. We show that the new concepts introduced in the model by the Parsli model enable reducing the complexity of morphological descriptions w.r.t. both traditional or more recent models.

6.9.2. Derivational morphology

This year, in relation with the ANR project EDyLex (see section 8.2.2), work has started targeted towards the acquisition of lexical information at the level of derivational morphology, both using semi- and non-supervised techniques.
Semi-supervised techniques have been used in a work dedicated to French denominal adjectives, for which we have implemented an automatic technique based on large-scale lexicons and corpora for extracting derivation links between base nouns and derived adjectives based on the same stem [46]. The resulting derivational lexicon, which is freely available, has already been partially manually validated. Future work include a full validation and adding denominal adjectives with a suppletive base.

Unsupervised techniques have been used for extraction of derivational links that appear more systematically, although their definition is less linguistically motivated as such [51].

6.9.3. Morphological issues concerning loan words

Also in the context of the ANR project EDyLex (see section 8.2.2), we have carried out a preliminary study on the morphological issues raised by borrowing phenomena, concerning in particular French nouns and verbs borrowed from English [52]. Using techniques that are similar to those used on derivational morphology, we have extracted a significant amount of loan words from a large raw corpus. We have proposed a model of the borrowing phenomenon, that takes into account graphemic (spelling), phonetic and morphological variability.

6.10. Allophony and word segmentation in language acquisition models

Participants: Luc Boruta, Benoît Crabbé.

Allophonic rules are responsible for the great variety in phoneme realizations. Infants can not reliably infer abstract word representations without knowledge of their native allophonic grammar. We have explored the hypothesis that some properties of infants’ input, referred to as indicators, are correlated with allophony. First, we provide an extensive evaluation of individual indicators that rely on distributional or lexical information. This evaluation relies on a phonetically transcribed corpus, generated automatically from a phonemically transcribed English, French and Japanese child-directed corpus. As such corpora do not exist as such, we used automatically extracted allophonic grammars of various sizes leading to various granularity levels, using our own allophonic rule extraction algorithm [57]. Then, we present a first evaluation of the combination of indicators of different types, considering both logical and numerical combinations schemes [23]. Though distributional and lexical indicators are not redundant, straightforward combinations do not outperform individual indicators.

Models of the acquisition of word segmentation are typically evaluated using phonemically transcribed corpora. Accordingly, they implicitly assume that children know how to undo phonetic variation when they learn to extract words from speech. Moreover, whereas models of language acquisition should perform similarly across languages, evaluation is often limited to English samples. Using the phonetically annotated corpora described above, that cover three typologically different languages, we evaluated the performance of state-of-the-art statistical models given inputs where phonetic variation has not been reduced. We have measured segmentation robustness across different levels of segmental variation, simulating systematic allophonic variation or errors in phoneme recognition. We have shown that these models do not resist an increase in such variations and do not generalize to typologically different languages. From the perspective of early language acquisition, the results strengthen the hypothesis according to which phonological knowledge is acquired in large part before the construction of a lexicon.

6.11. Modelling the acquisition of syntactic categories by children

Participants: Benoît Crabbé.

B. Crabbé co-supervised A. Gutman for an M2 thesis (MPRI) in collaboration with A. Christophe (LSCP/ENS) in the domain of psycholinguistic modelling. The topic was concerned with modelling and implementing psychologically motivated models of language treatment and acquisition. Contrary to classical Natural Language Processing applications, the main aim was not to create engineering solutions to language related tasks, but rather to test and develop psycholinguistic theories. In this context, the study was concerned with the question of learning word categories, such as the categories of Noun and Verb. It is established experimentally that 2-year-old children can identify novel nouns and verbs. It has been suggested that this can be done
using distributional cues as well as prosodic cues. While the plain distributional hypothesis had been tested quite extensively, the importance of prosodic cues had not been addressed in a computational simulation. We provided a formulation for modelling this hypothesis using unsupervised and semi-supervised forms of bayesian learning (EM) both offline and online.

6.12. Modelling and extracting discourse structures

Participants: Laurence Danlos, Charlotte Roze.

6.12.1. Cross-lingual lexical semantics of discourse connectives

Discourse connectives are words or phrases that indicate senses holding between two spans of text. The theoretical approaches accounting for these senses, such as text coherence, cohesion, or rhetorical structure theory, share at least one common feature: they acknowledge that many connectives can indicate different senses depending on their context. Depending on its sense, the translation of a connective into another language can vary greatly, either using an equivalent connective, or using a different construction or even no explicit connective at all.

On the basis of data provided by the bilingual concordancer TransSearch which propose statistical word alignment [64], [53] made a semi-manual annotation of the English translation of two French connectives (“en effet” and “alors que”). The results of this annotation show that the translations of these connectives do not correspond to the “transpots” identified by TransSearch and even less to the translations proposed in bilingual dictionaries.

The conclusions of this work were presented at an European workshop organized by the project COMTIS, and some members decide to use our technic for other connectives and other aligned corpora (e.g. Europarl).

6.12.2. Discourse relations inference rules

In 2011 we have developed a new methodology for building discourse relations inference rules, to be integrated into an algebra of these relations [54], [38]. The construction of such an algebra has as main objective the improvement of the comparison of discourse structures within the evaluation of discourse annotations and the creation of a gold-standard corpus. The inference rules can also help detecting inconsistencies in discourse structures, in order to improve human or machine annotation. The premises of rules already studied lead to the formulation of inference rules, established by the theoretical definition of discourse relations, manually constructed data and extracted data. By manually annotating discourses, we also compute inference probabilities. We have illustrated the adopted methodology taking as theoretical background the Segmented Discourse Representation Theory [60].

6.12.3. Discourse structure and factivity

Discursive annotations proposed in theories of discourse such as RST (Rhetorical Structure Theory) or SDRT (Segmented Representation Theory Discourse) have the advantage of building a global discourse structure linking all the information in a text. Discursive annotations proposed in PDTB (Penn Discourse Tree Bank) have the advantage of identifying the "source" of each information – thereby answering to questions such as who says or thinks what?

In collaboration with Owen Rambow (Columbia University), we have proposed [26], [28] a unified approach for discursive annotations combining the strengths of these two streams of research. This unified approach relies crucially on factivity information, as encoded in the English corpus FactBank. We intend to pursue this avenue of research by initiating in 2012 the development of a French FactBank.

6.13. Modelling and extracting temporal structures

Participants: Pascal Denis, Philippe Muller.

5 http://www.idiap.ch/project/comtis
Temporal information has been the focus of recent attention in information extraction. An elegant approach to learning temporal orderings from texts is to formulate this problem as a constraint optimization problem, which can be then given an exact solution using Integer Linear Programming. This works well for cases where the number of possible relations between temporal entities is restricted to the mere precedence relation, but becomes impractical when considering all possible interval relations.

We have proposed this year two innovations [29], inspired from work on temporal reasoning, that control this combinatorial blow-up, therefore rendering an exact ILP inference viable in the general case. First, we propose to translate the network of constraints from temporal intervals to their end-points, to handle a drastically smaller set of constraints, while preserving the same temporal information. Second, we have show that additional efficiency is gained by enforcing coherence on particular subsets of the entire temporal graphs. We evaluate these innovations through various experiments on TimeBank 1.2 using standard evaluation metrics, and compare our ILP formulations with various baselines and oracle systems.

The evaluation of temporal information extraction, i.e., the comparison of two annotations of a given text, is also a scientific challenge. This is because relations between events in a story are intrinsically interdependent and cannot be evaluated separately. A proper evaluation measure is also crucial in the context of a machine learning approach to the problem. Finding a common comparison referent at the text level is not obvious, and we have argued, in collaboration with Xavier Tannier (LIMSI), in favor of a shift from event based measures to measures on a unique textual object, a minimal underlying temporal graph, or more formally the transitive reduction of the graph of relations between event boundaries [15].


**Participants:** Emmanuel Lassalle, Pascal Denis.

Bridging descriptions are a special kind of anaphora whose interpretation requires not only identifying an antecedent, but also inferring a specific relation linking it to the anaphor. The resolution of bridging anaphora represents a very challenging task in discourse processing. It is considerably much harder than standard coreferential anaphora resolution for which shallow predictors (like distance, string matching, or morphosyntactic agreement) have been shown to be rather effective. Part of the challenge is due to an important information bottleneck. Lexical resources like WordNet are still too poor and uneven in coverage to provide a realistic solution. In turn, more recent approaches to bridging resolution have turned to web-based extraction methods. To date, the most complete and best-performing approach combines focus and lexical distance predictors using machine learning techniques [105].

We have focused on mereological bridging anaphora (that is, cases wherein the inferred relation is a part-whole relation). Moreover, we have worked on French, a language for which current lexical resources have a very low coverage. The system, presented in [32] is similar to a system developed for English [105], but it was enriched to integrate meronymic information extracted automatically from both web queries and raw text using syntactic patterns. Through various experiments on the DEDE corpus [78], we show that although still mediocre the performance of our system compare favorably to those obtained for English by the above-mentioned system. In addition, our evaluation indicates that the different meronym extraction methods have a cumulative effect, but that the text pattern-based extraction method is more robust and leads to higher accuracy than the Web-based approach.

6.15. Statistical models of word order in French

**Participants:** Juliette Thuilier, Benoît Crabbé.

We study the problem of choice in the ordering of French words using statistical models along the lines of [66] and [67]. This work aims at describing and model preferences in syntax, bringing additional elements to Bresnan’s thesis, according to which the syntactic competence of human beings can be largely simulated by probabilistic models. We previously investigated the relative position of attributive adjectives with respect to the noun.

---

6An illustrative English example is is the following discourse: *The car will not move. The engine is broken.*
This year, we mainly studied the problem of the relative ordering of postverbal complements. The focus of this investigation is the relative order of direct object and indirect object of French ditransitive verbs. The first part of this work is based on corpora data that we extracted from two journalistic corpora (French Tree Bank and Est-Républicain) and a radio corpus (ESTER). These data were manually annotated and validated for semantic categories (animacy and semantic class of the ditransitive verb). Based on these data, we built statistical models showing that the relative length of complements and verbal lemmas are the most important factors, and that, differently from English or German, categories as animacy or definiteness seem to play no role in the relative ordering.

In collaboration with Anne Abeillé (Laboratoire de Linguistique Formelle, Université Paris 7), we extended our corpora study with psycholinguistic questionnaires, in order to show that statistical models are reflecting some linguistic knowledge of French speakers. The preliminary results confirm that animacy is not a relevant factor in ordering French complements.

As regards to corpus work, we are extending the database with spontaneous speech corpora (CORAL-ROM and CORPAIX) and a wider variety of verbal lemmas, in order to enhance sample representativeness and statistical modelling. In a crosslinguistic perspective, we plan to strengthen the comparison with the constraints observed in other languages such as English or German.

As can be seen from the outline above, this line of research brings us closer to cognitive sciences. We hope in the very long run that these investigations will bring new insights on the design of probabilistic parsers or generators. In NLP the framework that is closest to implementing construction grammar is Data Oriented Parsing.

6.16. Assessing the Amazon Mechanical Turk plateform

Participant: Benoît Sagot.

In collaboration with Gilles Adda and Joseph Mariani from LIMSI and with Karën Fort from INIST, we have assessed some crowdsourced microworking systems and especially Amazon Mechanical Turk, the use of which has been steadily growing in language processing in the past few years [41], [17]. According to the mainstream opinion expressed in the articles of the domain, this type of on-line working platforms allows to develop very quickly all sorts of quality language resources, for a very low price, by people doing that as a hobby or wanting some extra cash. We have demonstrated that the situation is far from being that ideal, be it from the point of view of quality, price, workers’ status or ethics and bring back to mind already existing or proposed alternatives. Our goal was threefold:

- to inform researchers, so that they can make their own choices with all the elements of the reflection in mind,
- to ask for help from funding agencies and scientific associations, and develop alternatives,
- to propose practical and organizational solutions in order to improve new language resources development, while limiting the risks of ethical and legal issues without letting go price or quality.

6.17. Finite state formalisms for Egyptian Hieroglyphic transliteration

Participant: François Barthélémy.

The task of transliterating an Egyptian Hieroglyphic text into the latin alphabet was studied [20], as a model problem to compare two finite-state formalisms: the first one is a cascade of binary transducers; the second one is a class of multitape transducers expressing simultaneous constraints, implemented using the Karamel language [62]. The two systems were compared regarding their expressivity and readability. The first system tends to produce smaller machines, but is more tricky. On the other hand, the Karamel language provides a more abstract description of the forms, using an explicit tree structure and separating the different pieces of information on different tapes, according to semantic criteria. But the Karamel machine is much larger. Karamel is a high-level declarative formalism whereas non contextual rewrite rules are an efficient low-level language.
ARIANA Project-Team

6. New Results

6.1. Optical imagery for remote sensing

6.1.1. Phase Field-Higher Order Active Contours for Object Modelling and Image Segmentation

Participants: Ikhlef Bechar, Josiane Zerubia [contact].

This work is done in collaboration with Dr Ian Jermyn of Durham University (United Kingdom) and was funded by a contract with the EADS foundation [http://www-sop.inria.fr/ariana/Ikhlef.Bechar/hoacs/index.html].

The problem of object segmentation from imagery is an essential preliminary task for many applications (target recognition, automated navigation, organ segmentation in medical imaging, etc). The problem of adding prior knowledge about objects to the image segmentation process has received a lot of interest since recently, and active contours [38] provide us with such a tool.

We consider a new class of active contours called Higher Order Active Contours (HOACs) introduced initially in [41], and which consider an optimal contour as the one which minimizes an energy involving three additive terms; namely the length of a contour, its area and a term of interaction between all possible pairs of its points via an interaction function $\Psi(t)$. The three terms being weighed by means of three scalar parameters. The main advantage of HOACs over traditional segmentation methods is that they offer an unprecedented means for including shape prior about an object via the interaction function $\Psi(t)$.

The HOAC set up has been applied successfully to various object extraction problems such as the extraction of networks [40], circular shapes [37], etc, using a specific family of the $\Psi(t)$'s. Our main task in the framework of this project is to extend their work to more general shapes.

Our contributions so far have been in the numerical computation of the optimal HOAC parameters for a given shape. We have shown indeed that the HOAC energy can be fully made linear with respect to $\Psi(t)$, which then makes it easy to solve for $\Psi(t)$ numerically (cf. Fig. 6). This is achieved by first choosing a linear basis to represent $\Psi(t)$ and using K-K-T (Karush-Kuhn-Tucker) optimality criteria to express the fact that a target contour is a local minimum of the energy. Consequently, looking for the optimal values of the coefficients for a given shape amounts first to solve an eigen value problem, and second, to find the linear combinations of the found eigen vectors that satisfy both K-K-T minimality criteria. The computation being carried out in the Fourier domain for sake of computational efficiency. We are currently testing the proposed model on simple shapes such as the butterfly one shown in figure 6.

6.1.2. Optimization of the compression-restoration chain for satellite images

Participants: Mikael Carlavan, Laure Blanc-Féraud [contact].

This project involves the French Space Agency (CNES) and the CESBIO, on collaboration with TAS and I3S (Marc Antonini).

This work concerns the study of the optimal imaging chain in the context of satellite imaging. The main goal of this study is to propose a new method to address the problem of decoding-deconvolution-denoising and consists in a characterization and optimization of the compression/restoration processes considering the instrumental characteristics (FTM, noise, sampling). A theoretical study first showed that current processes of compression and restoration are better fitted if the restoration is performed on-board before the compression. Indeed, current restoration algorithm is designed to remove the blur and the instrumental noise but does not take into account the coding noise, and it is well-known that compression algorithms do not perform
Figure 6. An example of estimation of the $\Psi$ function for a given shape. (a) An example of a butterfly-like shape; (b) The numerical estimation of its optimal $\Psi$ function; (c) The result of a gradient descent algorithm on the HOAC energy with $\Psi$ of figure (b) until convergence (in black the original shape, and in red the shape found by the gradient descent algorithm). One can see that the estimated $\Psi$ makes the HOAC energy achieve a local minimum at a shape which is very close to the target shape of figure (a).
properly on noisy data. More generally, we concluded that the image should be the closest possible to the real image before the step of coding, encouraging, thus, to move the restoration step on-board before the compression. Figure 7 shows the global distortion w.r.t. the coding rate if the restoration step is done either before (on-board) or after (on-ground) the compression. We see that using an on-board restoration leads to a quality improvement of the final image regardless the coding rate. For example, at the usual coding rate of 2.5 bits/pixel, using an on-board restoration improves the quality of the final image about 0.5 dB.

![Figure 7. Global distortion w.r.t coding rate. The solid line is the global distortion with the restoration step done on-ground after the compression and the dashed line is the global distortion if the restoration is performed on-board before the compression.](image)

6.1.3. Aerial Image Restoration

**Participants:** Daniele Graziani, Laure Blanc-Féraud [contact].

*This project involves as partners: ATE, Coreti, and Gilles Aubert from the J.A.D. Laboratory at the University of Nice Sophia Antipolis.*

The goal of the project is to build an airborne camera system, and our part is to process aerial images provided by ATE: restoration, microscanning, video, color images. We investigate a convex variational framework to compute high resolution images from a low resolution video. We analyze the image formation process to provide a well designed model for warping, blurring, downsampling and restoration. The microscanning is modeled as a convex minimization problem, which is solved with a domain decomposition technique based on the recent work of M. Fornasier, A. Langer and C. Schonlieb. ("A convergent overlapping decomposition method for total variation Minimization", Numeriske Math. to appear), which allows parallel computing and a realization of a real time algorithm.

6.1.4. Contribution of object recognition on forest canopy images to the building of an allometric theory for trees and natural, heterogeneous forests

**Participants:** Jia Zhou, Xavier Descombes, Josiane Zerubia [contact].

*This work is done in collaboration with Dr. Pierre Couteron and Christophe Proisy at IRD, UMR AMAP, Montpellier.*
Individual tree detection methods are more and more present, and improve, in forestry and silviculture domains with the increasing availability of satellite metric imagery. Automatic detection on these very high spatial resolution images aims to determine the tree positions and crown sizes. The mathematical model based on marked point processes has showed advantages w.r.t. several individual tree detection algorithms for plantations. We used this detection method to analyze natural mangrove forests in French Guiana, eucalyptus plantations in Brazil, and other types of tropical forests. The simulated optical images were also used to improve the method and calibrate the detection parameters. To analyze a eucalyptus plantation in Brazil [23], we used 2 optical images acquired by the WorldView-2 satellite. A tentative detection simultaneously with 2 images of different dates (multi-date) was tested for the first time, which estimates individual tree crown variation during these dates. In this work, we tried to find the trees localizations and crown sizes in order to provide a plantation map, and estimate the tree crown growth during the period between 2 images, and compared these results with the field measurements and expected dynamics of corresponding populations. An example of multi-date detection result is showed in figure 9.

The detection method was also applied on simulated optical DART (Discrete Anisotropic Radiative Transfer) images, where exact field inventory could be provided on large surfaces. We assessed the detection results with these “ground-truth” maps.

6.2. SAR imagery for remote sensing

6.2.1. Stochastic modeling for very high resolution SAR image processing

Participants: Aurélie Voisin, Vladimir Krylov, Josiane Zerubia [contact].

This work is done in collaboration with DIBE, University of Genoa, with Dr Gabriele Moser and Prof. Sebastiano B. Serpico [http://spt.dibe.unige.it/] with partial financial support of the French Defense Agency, DGA [http://www.defense.gouv.fr/dga/] and the Italian Space Agency, ISA [http://www.asi.it/en].
Figure 9. Results of a multi-date detection on 2 images at different dates: in May (left) and in August (right)

We deal with the environmental risk assessment by addressing the problem of classifying SAR images of urban areas. Several difficulties need to be considered to address the SAR classification problem. The first one is related to the inherent multiplicative noise known as speckle, which degrades appreciably the registered imagery. Another difficulty is the heterogeneity of urban areas on very high resolution (VHR) images that leads to heterogeneous statistical modeling, reflecting the different ground materials such as asphalt, concrete, metal, etc. We propose a hierarchical statistical Bayesian supervised classification approach that consists of two steps. The first step deals with the SAR amplitude statistical modeling for each target class (e.g. vegetation, urban, etc.) by using a finite mixture model, estimated by resorting to a dictionary-based stochastic expectation maximization (DSEM) algorithm. More specifically, the SAR amplitude probability density functions (PDFs) are assumed to be mixtures of $K$ PDFs automatically chosen inside a predefined dictionary of SAR-specific distribution families. Such mixtures are intended to take into account the above mentioned VHR SAR statistics heterogeneity. We further consider an additional source of information obtained by extracting a textural feature map from the original SAR image in order to optimize the detection of urban areas. Typically, the textural feature is generated by using a Grey Level Co-occurrence Matrix (GLCM)-based method. The marginal PDFs of the original SAR image and the textural feature are combined via copulas, leading to a joint PDF for each class. On the second step the classification map is generated, using the joint copula-based statistics. To improve the robustness with respect to speckle noise, we consider a contextual model based on Markov random fields (MRFs), and, more specifically, a hierarchical MRF, which offers the possibility to take into account the multi-scale information and to deal with multi-resolution imagery [28]. A variety of algorithms were proposed to estimate the labels on hierarchical graphs. The consideration of a specific graph, here a quad-tree, allows to benefit from its good properties (e.g. causality) and to apply non iterative algorithms. Among the different algorithms employed in the literature, we chose to take into account an exact estimator of the marginal posterior mode (MPM). The cost function associated to this estimator offers the possibility to penalize the errors according to their number and the scale at which they occur: an error at the coarsest scale is stronger penalized than an error at the finest scale. Moreover, we introduce a prior estimation update that experimentally leads to improved results and is less affected by speckle noise when compared to a predefined prior [35]. The challenge of the problem considered here is that our given input is a single-polarized SAR image at a single resolution. To improve the classification, we extract an extra information in the form of a multi-scale wavelet decomposition from the initial image. Then, at each level, the textural feature map is obtained from each image in the decomposition stack. Finally, at each level, the wavelet image is combined with the textural image by using copulas, as described previously in Ariana activity reports. The hierarchical method was tested on real COSMO-SkyMed images. We illustrate the obtained results with an example of a SAR acquisition of the Port-au-Prince quay (Haiti). Spatially disjoint training and test areas were manually annotated. The classification is done following 3 classes: urban areas, natural landscape and wet areas. The
results are shown qualitatively in figure 10. The computation of numerical results gives an average accuracy of 95.65 percent for the considered test areas.

Figure 10. Left: Initial SAR image of Port-au-Prince (Haiti) (©ISA, 2009). Right: Classification map obtained with the hierarchical method for the 3 classes (Blue: water; Green: vegetation; Red: urban area).

Our previous work was based on single-scale MRF, thus the hierarchical approach is a direct extension. Part of our work was dedicated to the comparison of these two methods [35], [28], and we also compared the MRF-based model to a novel products of experts approach [18].

6.2.2. Parameter estimation procedures for HR SAR image classification

Participants: Vladimir Krylov, Josiane Zerubia [contact].

This work is conducted in collaboration with DIBE, University of Genoa with Dr. Gabriele Moser and Prof. Sebastiano Serpico [http://spt.dibe.unige.it/] with the support of the Italian Space Agency, ASI [http://www.asi.it/en].

Parameter estimation of probability density functions is one of the major steps in the mainframe of statistical image and signal processing. We have explored the properties and limitations of the recently proposed method of logarithmic cumulants (MoLC) parameter estimation approach which is an alternative to the classical maximum likelihood (ML) and method of moments (MoM) approaches. We have derived the general sufficient condition of strong consistency of MoLC estimates which represents an important asymptotic property of any statistical estimator [33]. We have demonstrated the strong consistency of MoLC estimates for a selection of widely used distribution families originating (but not restricted to) synthetic aperture radar (SAR) image processing. We have then derived the analytical conditions of applicability of MoLC to samples generated from several distribution families in our selection. We have conducted various synthetic and real data experiments to assess the comparative properties, applicability and small sample performance of MoLC notably for the generalized gamma and $K$ family of distributions. The synthetic-data experiments have demonstrated a competitive accuracy of MoLC estimates and a reliable behavior of this estimator for small samples which is a critical issue in applications. We have performed real-data image processing experiments to the problem of supervised classification applied to high resolution satellite SAR imagery. These experiments confirmed the stability of MoLC estimator with respect to sample size and at the same time illuminated the critical side of MoLC given by applicability restrictions. The experiments suggested the efficiency of use of the MoLC estimator for finite-mixture estimation problems [3], [19], and ML-based classification approaches [4], [28], [35].
6.2.3. Unsupervised amplitude and texture based classification of SAR images with multinomial latent model

**Participants:** Koray Kayabol, Aurélie Voisin, Vladimir Krylov, Josiane Zerubia [contact].

The participants would like to thank the Italian Space Agency (ASI) for providing the COSMO-SkyMed images. The TerraSAR-X images are provided from http://www.infoterra.de/.

We combine both amplitude and texture statistics of the Synthetic Aperture Radar (SAR) images using Products of Experts (PoE) approach for classification purpose. We use Nakagami density to model the class amplitudes and a non-Gaussian Markov Random Field (MRF) texture model with $t$-distributed regression error to model the textures of the classes. A non-stationary Multinomial Logistic (MnL) latent class label model is used as a mixture density to obtain spatially smooth class segments. The Classification Expectation-Maximization (CEM) algorithm is performed to estimate the class parameters and to classify the pixels [18]. Determining the necessary number of classes to represent the data and initialization are some drawbacks of the EM type algorithms. In [17] and [32], we combine hierarchical agglomeration, CEM and Integrated Classification Likelihood (ICL) criterion to get rid of the drawbacks of EM. We obtained some classification results of water, land and urban areas in both supervised and unsupervised cases on TerraSAR-X, as well as COSMO-SkyMed data [18], [17], [32]. The proposed unsupervised ATML-CEM (Amplitude and Texture density mixtures of MnL with CEM) method provides significantly better results, see Fig. 12, compared to the corresponding results obtained with K-MnL and its performance is close to supervised ATML-CEM.

6.3. 3D-modelling of urban scenes

6.3.1. Building reconstruction from aerial LiDAR data

**Participants:** Yannick Verdie, Florent Lafarge [contact], Josiane Zerubia.

The generation of 3D representations of urban environments from aerial and satellite data is a topic of growing interest in image processing and computer vision. Such environments are helpful in many fields including urban planning, wireless communications, disaster recovery, navigation aids, and computer games. Laser scans have become more popular than multiview aerial/satellite images thanks to the accuracy of their measurements and the decrease in the cost of their acquisition. In particular, full-waveform topographic LIDAR constitutes a new kind of laser technology providing interesting information for urban scene analysis. We study new
Figure 12. (a) The original SAR image (COSMO-SkyMed, ©ASI), (b), (c) and (d) classification maps obtained by K-MnL, supervised and unsupervised ATML-CEM methods. Blue, red and green colors represent water, urban and land areas, respectively.
stochastic models for analysing urban areas from LIDAR data. We aim to construct concrete solutions to both urban object classification (i.e. detecting buildings, vegetation, etc.) and the 3D reconstruction of these objects. Probabilistic tools are well adapted to handling such urban objects, which may differ significantly in terms of complexity, diversity, and density within the same scene. In particular, jump-diffusion based samplers offer interesting perspectives for modelling complex interactions between the various urban objects. We investigated a first approach aiming at producing accurate, watertight and compact meshes from planar patches under planar constraint especially designed for urban scenes. The LiDAR point cloud is classified through a non-convex energy minimization problem. The planar structures are extracted and connected to generate a compact, and watertight mesh of the building. Experiments highlight the potential of our method in term of performance, compactness, and accuracy. This work has been published in [22]. We illustrated our results in figure 13.

![Figure 13. Results of 3D urban reconstruction by the framework described in [22].](image)

### 6.3.2. Modeling large urban environments from unstructured point clouds

**Participant:** Florent Lafarge [contact].

We present a robust method for modeling cities from unstructured point data. Our algorithm provides a more complete description than existing approaches by reconstructing simultaneously buildings, trees and topologically complex grounds. Buildings are modeled by an original approach which guarantees a high generalization level while having semantized and compact representations. Geometric 3D-primitives such as planes, cylinders, spheres or cones describe regular roof sections, and are combined with mesh-patches that represent irregular roof components. The various urban components interact through a non-convex energy minimization problem in which they are propagated under arrangement constraints over a planimetric map. We experimentally validate the approach on complex urban structures and large urban scenes of millions of points as illustrated on Figure 14.

### 6.3.3. Parallel Monte Carlo sampler for point processes

**Participants:** Yannick Verdie, Florent Lafarge [contact], Ioan Dragan.

We designed a new parallel scheme for Markov point processes. These probabilistic models exploit random variables whose realizations are configurations of parametric objects, each object being assigned to a point positioned in the scene. The number of objects is itself a random variable. Another strength of Markov point processes is their ability to take into account complex spatial interactions between the objects and to impose global regularization constraints. Moreover, we proposed to use space-partitioning tree such as quadtree (for 2D data) or octree (for 3D data) for non-homogeneous measure adapted to the problem. We illustrate the results in figure 15.
Figure 14. Reconstruction of the cities of Marseille and Amiens, France, from Lidar point cloud.

Figure 15. Results of the new Marked Point process on (left) 2D data, and (right) LiDAR data.
6.4. Biological imagery

6.4.1. Regularizing parameter estimation with Poisson noise

Participants: Mikael Carlavan, Laure Blanc-Féraud [contact].

The problem is to automatically estimate the regularizing parameter in Poisson noisy image deconvolution using the $L_1$-norm regularization as a total variation or frame coefficients. This problem is addressed using the discrepancy principle. The standard weighted criterion composed of a data term and a regularization term is rewritten as a constrained minimization problem. The constraint is designed on the data term using the discrepancy principle and a new estimation of the bound is proposed as well as an efficient algorithm to solve this constrained minimization problem. This work is published in [10], [24], [9].

6.4.2. Brain vascular network segmentation

Participant: Xavier Descombes [contact].

This work was conducted in collaboration with Franck Plouraboué and Abdelhakim El Boustani from IMFT Toulouse and Caroline Fonta from CerCo Toulouse. It has been partially supported by a PEP II project from CNRS.

Micro-tomography produces high resolution images of biological structures such as vascular networks. We have proposed a new approach for segmenting vascular network into pathological and normal regions from considering their micro-vessel 3D structure only. We consider a partition of the volume obtained by a watershed algorithm based on the distance from the nearest vessel. Each region is characterized by its volume and the local vascular density. The volume and density maps are first regularized by minimizing the total variation. Then, a new approach has been proposed to segment the volume from the two previous restored images based on hypothesis testing. Results are presented on 3D micro-tomographic images of the brain micro-vascular network (see Fig. 16).

![Figure 16. Brain micro-vascular network segmentation](image)

6.4.3. Blind restoration of 3D biological image

Participants: Saima Ben Hadj, Laure Blanc-Féraud [contact].
Fluorescence microscopy is a powerful imaging technique providing three-dimensional images of biological living specimen. However, these images are degraded by a depth variant blur due to light diffraction phenomenon as well as refractive index mismatch between the different mediums composing the system and the biological sample. They are also distorted with noise from non-ideal imaging conditions. In order to provide biologist with more suitable images for quantitative studies, many restoration methods were developed. In most of them, the blur function, called Point Spread Function (PSF) is assumed to be piecewise constant in order to avoid the intensive computing time when using a pointwise varying PSF. However, this usually leads to blocking effect in the restored image. In our work, we extend the Space varying (SV) blur model previously proposed in [39] for 2D astronomical images to 3D microscopy images. In that model, the degraded image is a convex combination of convolutions with a space-invariant (SI) PSF. Furthermore, we fit to that model two restoration procedures which are basically developed for a SI PSF. On the one hand, we use the Richardson-Lucy method with Total Variation regularization which is carried out under Poisson noise assumption in order to restore confocal microscopy images. We employ another method with total variation regularization adapted to images with an additive Gaussian noise in order to restore Wide Field Microscopy images. For that, we rely on a fast optimization method based on a domain-decomposition technique [36]. In particular, we study its convergence properties when using the SV blur model [31], [27]. To illustrate the interest of the proposed method, we show in Fig. 17 some results obtained on a simulated bead image of Wide Field Microscopy.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure17.png}
\caption{(X,Z) slices of the (a) original image, (b) degraded image (c) restored image using a SV blur model, and (d) restored image with a space-invariant PSF.}
\end{figure}

### 6.4.4. Axon imaging

**Participant:** Florence Besse [contact].

During brain development, neurons extend cellular processes (dendrites and axons) to connect to specific targets and establish functional networks. Understanding how axonal processes migrate to reach their targets and how they form new branches to build up a complex axonal tree is thus key. To characterize the properties of axonal trees, we have generated a collection of confocal 3D pictures of normal and mutant single axons labeled using a fluorescent protein. To analyze the formation of axonal trees in real-time, we have developed a protocol to dynamically image growing axons within intact Drosophila brains. In this protocol, entire brains are cultured in conditions where they can undergo cell differentiation and maturation. Axons are labeled by the fluorescent molecule GFP and are imaged over 12h, with very low-photobleaching and no associated phototoxicity, using an ultra-sensitive 2-photon microscope. Several image sequences corresponding to the growth of axons in normal conditions have been acquired (see Fig. 18).

### 6.4.5. Detection of Axons in Neuronal Images

**Participants:** Alejandro Mottini, Florence Besse, Xavier Descombes [contact].
Imaging techniques such as confocal and two-photon microscopy provide an efficient way of analyzing supra-cellular structures. It is known that the precise shape of these objects provides information on their functioning and allows the characterization of pathological states. Therefore, the analysis of the morphological differences between normal and pathological structures is of paramount importance. In particular, the analysis of neuronal axon topologies allows biologists to study the causes of neurological diseases such as Fragile X Syndrome and Alzheimer’s disease. For this purpose biologists have acquired static 3D images of mature neuron axons using fluorescent confocal microscopy.

Due to the high volume of generated image data and the tortuous nature of the axons, manual processing is infeasible. Therefore, it is necessary to develop techniques for the automatic extraction and analysis of the neuronal structures. However, since both types of images present different characteristics, two different methodologies need to be developed. The main objective of the static case is the study of the length and number of bifurcation points of the two populations of neurons. On the other hand, the focus in the dynamic
case is put on the tracking of the axonal tips. The automatic extraction of axons from confocal microscope images is a key problem in the field of neuron axon analysis. In this work we propose a solution which combines algorithms for the denoising, binarization, skeletonization, gap filling, point detection and statistical analysis in a pipeline capable of extracting the axons. Furthermore, these algorithms were incorporated in a software, developed in Matlab, which includes an easy to use GUI along with functionalities to facilitate high data throughput analysis. The software was evaluated on several 3D confocal microscope images of normal and mutated axons. Our results support the potential use of the software in helping biologist perform automatic morphological analysis of axons in fluorescent confocal microscopy images.

6.4.6. Axons tracking

Participants: Alejandro Mottini, Huei-Fang Yang, Florence Besse, Xavier Descombes [contact].

This work was funded by project ARC-DADA (INRIA/CNRS/UNSA) [http://www-sop.inria.fr/members/Xavier.Descombes/DADA/home.html]. It is done in collaboration with Serpico team at INRIA Bretagne (C. Kervrann, P. Houllier)

To study axon growth process biologists have acquired dynamic 3D+t image sequences of developing neurons using fluorescent two-photon microscopy.

Live cell two-photon microscopy is an effective tool for the analysis of dynamical processes occurring in living samples that, when combined with fluorescence, allows the detection of objects of interest in 3D space and time. These labeled objects appear as bright spots which need to be detected. The low resolution and signal-to-noise-ratio (SNR) make this analysis difficult even for experienced biologists. As a consequence, automatic detection techniques need to be robust and flexible. To this end, the Marked Point Process (MPP) detection framework was selected. Since axonal extremities appear either as circular objects or as ending points of thin filaments, we proposed two different MPP models. These models were formulated using a Gibbs energy function and optimized with Multiple Births and Deaths, a newly proposed algorithm which guarantees a fast convergence to the global minimum. The first model is designed to detect spheres or disks and the second filaments (both in 3D). Both models were tested on several 3D static images. To evaluate the performance of the detection, results were compared against images labeled by an experienced biologist.

Figure 22 shows the results obtained on one frame of a video sequence (the shown image is a slice of the 3D frame) for the disk/sphere model. One can appreciate that the axonal tip marked by the biologist (in white, left) was detected. However, a false positive is also present.
Figure 21. Dynamic images (2D maximum intensity projection).

Figure 22. Original (left) and resulting (right) images (detected disks in pink). The true axonal tip is circled in white (original image) and red (result). In the original image, axons are labeled in white.
Figure 23 shows the result obtained on one of the slices (for the same video frame) for the filament model. Although the good filament was detected, many false positives are also present in the image. Similar results were obtained for other images.

Once the detection of the extremities in each frame of the video has been solved, the solution should be embedded into a tracking algorithm to obtain an estimation of the axon’s trajectory during the growing stage. To this end, the particle filtering technique was considered. This technique consists in estimating the posterior distribution of the current state $x_t$ of the target of interest at time $t$ based on the measurements $z_{1:t}$:

$$
p(x_t|z_{1:t}) \propto p(z_t|x_t) \int p(x_t|x_{t-1}) p(x_{t-1}|z_{1:t-1}) \, dx_{t-1},
$$

(38)

where $p(x_t|x_{t-1})$ is the transition distribution, and $p(z_t|x_t)$ is the likelihood. In the current implementation, a simple kinematic model is used for the transition distribution, and a color histogram is applied to the computation of the likelihood. Figure 24 shows the preliminary results obtained by applying the particle filtering technique on the coronal KESM (Knife-Edge Scanning Microscopy) sections of the mouse cerebellum, where the green rectangles indicate the tracked objects.

Figure 24. Tracking results on the coronal KESM sections of the mouse cerebellum. The green rectangles are the tracked objects. The user gives an initial starting region on the first image, and the algorithm tracks the region of interest in the subsequent images. Note that the images are cropped for a better view.
The next step will be to design a more sophisticated transition distribution and likelihood model that are suitable for tracking the trajectories of axons during their growth.

6.5. Dermatology

6.5.1. Statistical analysis of skin pigmentation under treatment

Participants: Sylvain Prigent, Xavier Descombes, Josiane Zerubia [contact].

This work was partially funded by a contract with Galderma R&D [http://www.galderma.com/RampD.aspx].

One of the steps to evaluate the efficiency of a therapeutic solution is to perform measurements on a series of patients who received the studied treatment. In parallel another treatment is tested on another group of people or on the same group of patients on another skin area. This second treatment is the reference one for the studied pathology or a placebo. We will call it ‘vehicle’.

For facial hyper-pigmentation, for each studied treatment, a group of $N_e$ patients receives the treatment on one cheek and the vehicle on the other. To this end, patients are selected to have the same hyper-pigmentation severity on the two cheeks. Then multi-spectral images are taken at different times $t$ along the treatment period.

We propose a methodology to estimate the efficiency of a treatment by calculating a spectral criteria that maximizes the visibility of the disease comparatively to a healthy reference area. To design such a criterion, we compare three approaches. The first one gives weights to spectral bands in order to get an equivalent of the luminance from the $CIEL^*a*b*$ decomposition. This is the standard measure in dermatology. We compare this spectral measurement to a spectral signature obtained by ICA (Independent Components Analysis) in a whole study, and a criterion that searches for the highest contrasted band.

Once a criterion is designed, we compute the hyper-pigmentation severity of a patient by normalizing the pathological area with the healthy area and the active treatment measurement by the vehicle. That gives severity measurement distributions from patients at different times. A Student paired test allows to determine if an active treatment has an effect between two measurement times. The experiments done on 3 treatments and their associated vehicles brought to the following conclusion: The severity measure based on the selection of an optimal band allows to detect a treatment effect sooner than the two other approaches. The severity measurements obtained by the optimal band method is illustrated in figure 25.

Figure 25. Evolution of skin pigmentation in the time and its quantification.
AROBAS Project-Team

6. New Results

6.1. Perception and autonomous navigation

Participants: Patrick Rives, Pascal Morin, Andrew Comport, Alexandre Chapoulie, Gabriela Gallegos, Cyril Joly, Maxime Meillard, Glauco Scandaroli.

6.1.1. Indoor SLAM: Self-calibration of the camera frame with respect to the odometry frame

Fusing visual data and odometry information is a standard technique to improve the robustness of the SLAM solution. Odometry data is considered as an input of the motion prediction equation while the visual data constitutes the filter observation. However, such method requires that the system is well calibrated: the pose of the camera frame with respect to the odometry frame has to be known. Usually, this pose is directly obtained by an hand made measurement yielding to incorrect values. We propose a new self calibration method to get these calibration parameters automatically. In practice, the state in the SLAM formulation is augmented with the unknown camera parameters (with respect to the odometry frame). This method requires to adapt a few Jacobians with respect to the original SLAM algorithm which assumes that these parameters are known. The accuracy and the stability of the estimation scheme clearly depends on the observability and the conditioning properties of the new system.

In 2010, we presented results in the case where the camera frame location has only 3 degrees of freedom (two translations and a rotation with respect to the vertical z axis). This year, these results were extended to the full calibration problem. As in the previous case, we assume that the robot is moving on a horizontal ground and observes 3D landmarks from the images delivered by the on board camera. The five parameters introduced by the calibration problem - 2 translations and 3 rotations (only 2 translations since the z component is not observable due to the planar motion of the robot) - are estimated simultaneously in addition to the "classical" SLAM parameters. The implementation of the algorithm is based on a Smoothing And Mapping (SAM) approach which computes a solution by considering the whole trajectory (instead of only the current pose as with the EKF approach).

As a theoretical result, we prove that this augmented system remains observable if and only if the curvature of the robot trajectory changes. This analysis was validated on real data with our indoor robot. Fig. 2 shows the mobile platform and the camera. It can be seen that an important rotational offset was added on the camera to test the capability to deal with large rotational values (the parameters are initialized with identity). Results are provided on Fig. 3 - 4 and table 1. They show that the trajectory and the map seem consistent; moreover, the algorithm was able to correct the odometry drift (green trajectory on Fig. 3). Then, the observability analysis was validated since the estimation of the camera frame parameters begins when there is a significant change in the radius of curvature of the trajectory (see the confidence bounds on Fig. 4). Finally, the estimation of these parameters was consistent with the ground truth (table 1). These results were presented at IROS’11 conference [ 30 ].

6.1.2. Outdoor Visual SLAM

Safe and autonomous navigation in complex outdoor urban-like environment requires a precise and real time localization of the robot. Standard methods, like odometry, typically performed by wheel encoders or inertial sensors, are prone to drift and not reliable for large displacements. Low cost GPS stations are inaccurate and satellite masking effect happens too frequently due to corridor-like configurations. We develop a real time and accurate localization method based on vision only without requiring any additional sensor.
Figure 2. Left: Robot used for the experiment – Right: The omnidirectional camera mounted with a rotational offset

Figure 3. Red: trajectory and map provided by the algorithm — Green: odometry integration — Black cross: end of the trajectory — Blue: 99% confidence region for the last robot position
Figure 4. Evolution of the curvature and the estimation of the camera parameters

Table 1. Numerical results concerning the camera parameters

<table>
<thead>
<tr>
<th>Camera param.</th>
<th>Reference</th>
<th>Estimation</th>
<th>3σ bounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_x$ (m)</td>
<td>-0.16</td>
<td>-0.149</td>
<td>[-0.165, -0.132]</td>
</tr>
<tr>
<td>$t_y$ (m)</td>
<td>-0.35</td>
<td>-0.38</td>
<td>[-0.064, -0.013]</td>
</tr>
<tr>
<td>$\gamma_x$ (deg)</td>
<td>6.05</td>
<td>6.609</td>
<td>[6.276, 6.942]</td>
</tr>
<tr>
<td>$\gamma_y$ (deg)</td>
<td>-7.34</td>
<td>-7.714</td>
<td>[-7.889, -7.539]</td>
</tr>
<tr>
<td>$\gamma_z$ (deg)</td>
<td>0</td>
<td>0.668</td>
<td>[0.330, 1.007]</td>
</tr>
<tr>
<td>Final Pose</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_{end}$ (m)</td>
<td>-0.06</td>
<td>-0.11</td>
<td>[-1.32, 1.09]</td>
</tr>
<tr>
<td>$y_{end}$ (m)</td>
<td>0.15</td>
<td>0.56</td>
<td>[-1.76, 2.89]</td>
</tr>
<tr>
<td>$\theta_{end}$ (deg)</td>
<td>0</td>
<td>-1.18</td>
<td>[-12.96, 10.58]</td>
</tr>
</tbody>
</table>
Our approach relies on a monocular camera on board the vehicle and the use of a database of spherical images of the urban scene acquired during an offline phase. This geo-referenced database allows us to obtain a robust “drift free” localization. Basically, the database is constituted of spherical images augmented by depth which are positioned in a GIS (Geographic information system). This spherical robot centered representation accurately represents all necessary information for vision based navigation and mapping ([26]). During the online navigation, the current vehicle position is computed by aligning the current vehicle camera view with the closest reference sphere extracted from the database.

A spherical augmented acquisition system has been developed and tested on our Cycab vehicle. This system is composed of six wide angle stereo cameras in overlap, which permits to extract depth information by dense correspondence. Since the depth is available, we are able to construct 360 degrees spherical images with a unique center of projection. Those 3D spheres are then used in an image-based spherical odometry algorithm to obtain the trajectory of the vehicle ([31]), fuse the spheres and construct the database.

During the online navigation, we consider a vehicle equipped with a simple camera (perspective, omnidirectional...). Here the aim is to register the current view on the closest sphere stored in the database. To achieve this we have developed a spherical image-based registration which allows efficient and accurate localization. But since the database of augmented visual spheres can be acquired under different illumination conditions than the online camera is experiencing, a robust algorithm combining model based localization and online visual odometry has been developed [32]. This method performs in real-time (45 Hz), and allows to handle large illumination changes and outliers rejection (see figure 5).

As a part of the ANR CityVIP project, the localization and mapping system has been successfully tested in Clermont Ferrand (France). A database of augmented images has been built along a learning trajectory. The aim was to automatically “replay” the learning trajectory using the database and a monocular camera. To avoid collisions and pedestrians, a laser was mounted on the front of the Cycab. The system was able to autonomously follow large scale trajectories (over 400 meters), in crowded urban environments (see figure 6).

Figure 5. Top left, robust outliers weights. Top center, augmented reference image. Top right, reference depth-map. Bottom left, intensity error after alignment. Bottom center, registered image. Bottom right, original current image.
6.1.3. Loop closure detection for spherical view based SLAM

Although more precise than the odometry computed from the wheels encoders, the visual odometry also suffers from the problem of drift when large displacements are performed. It is possible to correct this drift if the robot is capable to determine if the place it is visiting has already been visited. This is often referred as the loop closure detection problem and several methods exist in the literature using perspective cameras. We develop new methods more reliable by exploiting the peculiar properties of spherical cameras.

Standard perspective cameras have a limited field of view leading to an incapability to encompass all the surrounding environment. This limitation of the field of view drastically limits the performances of visual loop closure algorithms. We propose to use spheres of vision computed by mosaicing images from 6 wide angle cameras mounted on a ring. Such a representation offers a full 360° field of view and keeps the spherical image invariant to the changes in orientations.

Loop closure detection can be exploited in a SLAM context at two levels: firstly, in the metric representation to retro-propagate along the robot’s trajectory the cumulative errors due to the drift, secondly, in the topological representation, to fusion in the graph representation the nodes corresponding to a same place.

Existing algorithms are not point of view independent: loop closures are detected uniquely when a place is revisited by the robot coming from the same direction but if the robot comes back in a different direction, the algorithms fail. Our solution relies on the presented spherical view and an efficient way of information extraction from it. We extract local information describing the points of interest of the scene. We enhance this local information with a global descriptor characteristic of the distribution of the points of interest over the sphere thereby describing the environment structure. These informations are used to retrieve the already visited places. Our algorithm performs well and is robust to the point of view variation [27]. This has led to an accepted paper at OMNIVIS 2011.
Figure 7. Spherical view acquisition

The figure 8 below presents obtained results. The trajectory is corrected (drift reduction) using the loop closure constraint. Red and green dots represent the loop closing places, they are linked by a red line.

Figure 8. Loop closure detection and trajectory drift correction

6.1.4. Context-based segmentation of scenes

In a topological SLAM framework based on vision, the places are often represented by images gathered at regular time/distance intervals. It is nevertheless a meaningless representation in the context of topology. We
would prefer a definition like "in front of a building" or "entrance of the campus" instead of "image number \( i \). Places are thus a set of images we need to group. This is what we call context based segmentation. In order to achieve this segmentation a criterion for "changing place" is needed, we propose to evaluate the environment structure using a global spatial descriptor (computed on the spherical view) called GIST. The algorithm relies on a statistical process control monitoring for an out-of-control signal involving a changing place event. The algorithm still needs to be improved for better robustness on the localization of the changing place events when we come back on previous visited paths.

The figure 9 presents the preliminary results. On the bottom left is the similarity matrix of the images GIST while on the bottom right is the segmented trajectory followed by our robot.

Figure 9. Context-based segmentation of scenes

6.1.5. Nonlinear observers for visual-inertial fusion with IMU-bias and camera-to-IMU rotation estimation

This work concerns the fusion of visual and inertial measurements in order to obtain high-frequency and accurate pose estimates of a visual-inertial sensor. While cameras can provide fairly accurate pose (position and orientation) estimates, the data acquisition frequency and signal processing complexity limit the capacities of such sensors in the case of highly dynamic motions. An IMU (Inertial Measurement Unit) can efficiently complement the visual sensor due to its high frequency acquisition, large bandwidth, and easy-to-process signals. IMU biases and calibration errors of the displacement between the camera frame and the IMU frame, however, can severely impair the fusion of visual and inertial data. Identification of these biases and calibration of this displacement can be achieved with dedicated measurement tools, but this requires expensive equipment and it is time consuming. We propose instead to address these issues via the design of observers. Last year, we had proposed a nonlinear observer to fuse pose and IMU measurements while identifying additive IMU biases on both gyrometers and accelerometers. We have extended this work to the self-calibration of the
rotation between the pose sensor frame (camera) and the IMU frame. Simulation and experimental results have confirmed that this calibration significantly improves the final pose estimation and allows to process motions with faster dynamics. This work has been presented at the IROS conference in October [35]. It is a joint work with G. Silveira from CTI in Brazil. We are currently extending this result to include the self-calibration of the translation displacement between both sensors.

6.2. Control of mobile robots

Participants: Claude Samson, Pascal Morin, Minh-Duc Hua [Post Doc, I3S, CNRS-Univ. of Nice-Sophia Antipolis], Daniele Pucci, Glauco Scandaroli, Luca Marchetti, Tarek Hamel [Univ. of Nice-Sophia Antipolis].

6.2.1. New developments of the Transverse Function control approach

We are pursuing the development of the Transverse Function control approach for highly nonlinear systems via the application of the approach to challenging mechanical systems with various structural control properties.

6.2.1.1. Control of a redundant wheeled snake mechanism using transverse functions on SO(4)

The Transverse Function approach is applied to the control of a nonholonomic three-segments/snake-like wheeled mechanism, similar to the planar low-dimensional version of Hirose’s Active Cord Mechanism (ACM) previously studied [65], but with two additional internal degrees of freedom (d.o.f.) whose actuation yields more flexible and efficient control solutions (see figure 10). From a theoretical point of view, these complementary d.o.f. modify the Control Lie Algebra of the system so that only first-order Lie brackets of the control vector fields are needed to satisfy the Lie Algebra Rank Condition (LARC). The fact that four independent (angular velocity) control inputs are used also implies for this system the existence of Transverse Functions (TF) defined on the six-dimensional special orthogonal group \(SO(4)\). Several examples of mechanisms whose control involve TF defined on \(SO(3)\) have been pointed out in the past [54], [64], [65]. Beyond the specific control problem addressed here, a motivation for the present study is to illustrate for the first time how functions defined on the larger set \(SO(4)\) can be determined and used for the control of a physical system. This study is complemented with recalls concerning the parametrization of \(SO(4)\) by pairs of isoclinic quaternions and with the derivation of complementary differential calculus relations associated with this parametrization. The results will soon be submitted for presentation at an international conference.

6.2.1.2. Control of three hooked vehicles with off-axle hitches

An extension of the study [65] performed last year on Hirose’s Active Cord Mechanism (ACM) concerns the case when one of the wheeled-trains (the middle one, for instance) possesses actuators giving it tractive and rotating capacities (alike a unicycle-like vehicle), while the other two vehicles are passively hooked to this tractive vehicle. This type of actuation departs from the one of Hirose’s Active Cord Mechanism for which the tractive capacity of the mechanism relies exclusively on the deformation of the system of vehicles via the control of the inter-connecting angles, and it makes an important difference at the control level. This system may also be seen as a unicycle-type vehicle with two trailers and off-axle hitches. Unlike the simpler hitch-on-axle case commonly addressed in the literature, this system is not differentially flat and “complete” feedback solutions ensuring practical stabilization of any, feasible or non-feasible, trajectory remained an open issue. This actuation allows for the complete alignment of the three vehicles without going through actuation singularities, and for the asymptotic tracking of a reference frame moving along a straight line or a circle. On the other hand, in order to fully take advantage of the extra possibilities offered by it, one has to consider higher-order Lie bracket maneuvering motions that significantly complicate the feedback control design. The Transverse Function approach is applied using the fact that a dynamic extension of this two-control-inputs system is left-invariant on a 6-dimensional Lie group. Transverse functions calculated as the group product of “elementary” functions defined either on toruses or on \(SO(3)\), and yielding feedback controls ensuring asymptotic stabilization of “feasible” reference trajectories under common “persistent excitation” properties (as in the case of classical feedback control solutions based on a linear approximation of the associated tracking error system) are proposed. As usual, the superiority of the transverse function solution over more classical solutions comes from that it also applies to the case of non-feasible reference trajectories for which (practical) stabilization involves complex maneuvers. The results of this study will be submitted next year for presentation at an international conference.
Figure 10. Three segments snake robot with two steering wheels. View from above
6.2.1.3. Control of an extended trident-snake vehicle

This study is part of a thesis work on the control of non-standard nonholonomic mobile robots by W. Magiera under the dual supervision of Prof. K. Tchon (Wroclaw University of Technology) and C. Samson. This collaboration involves several long term visits of the PhD student at INRIA, starting this year (2 months), and for the next two years. This year’s objective is to address a particularly challenging control problem and evaluate the possibilities offered by the Transverse Function approach to solve it. The system under consideration is based on the "common" trident snake mechanism [54] complemented with one, two, or three additional "passive" wheeled extensions, each of them subjected to the rolling-without-slipping constraint (see figure 6.2.1.3). Transverse Functions solutions tested so far involve a mixt (product) of functions defined either on the torus, or on special orthogonal groups, and future improvements may involve the search for new transverse functions.

![Trident-snake mechanism with passive extensions. View from above](image)

6.2.2. Control of aerial vehicles

6.2.2.1. Vehicles subjected to lift forces

The development of a general theory for the control of underactuated (ground, marine, and aerial) vehicles whose main propulsion relies on a thrust force exerted in a single (vehicle’s related) direction was continued
this year. Part of this program, more specifically devoted to aerial vehicles, is the subject of D. Pucci’s thesis research project. This year’s focus was the prolongation of the work initiated last year on the modelling of lift forces and on their effects on the flight and control of aerial vehicles. Among the new results obtained on the subject, an extension and generalization of a previous feedback control strategy developed for spherically-shaped vehicles only subjected to drag forces [52], based on an “ideal” generic model of lift and drag forces associated with bi-symmetric wings, has been accepted for presentation at an international conference ([34]). The proposed solution involves a change of thrust control input in order to render the dynamics of the transformed system independent of the angle of attack associated with the vehicle’s main wing. A weakness of the aforementioned model is that it does not account for the so-called stall phenomenon, which is an abrupt loss of lift when the angle of attack increases beyond a certain value called stall angle. Taking it into account adds considerable complexity to the vehicle’s dynamics, especially in the case of a vehicle moving within a fluid endowed with a large Reynolds number for which the stall phenomenon can no longer be neglected. We showed that, although this phenomenon never forbids the existence of an attitude equilibrium given a desired reference velocity, the uniqueness of this equilibrium is not always granted. As a consequence, modifications of the desired velocity may result in the abrupt disappearance of an equilibrium so that the asymptotic stabilization of a desired velocity profile may become an ill-conditioned problem. To avoid this complication a possibility consists in characterizing “good” velocity profiles –associated, for instance, to transition maneuvers between hovering and high-velocity cruising– for which the existence of continuously changing equilibria is ensured. First results on this topic and research direction have been submitted for presentation to an international conference.

6.2.2.2. Nonlinear control of VTOL UAVs with uncertain position measurements

This work concerns the feedback control of VTOL UAVs (Vertical Take-Off and Landing Unmanned Aerial Vehicles). The objective is to asymptotically stabilize a reference equilibrium configuration with a “semi-global” convergence domain, i.e. global convergence domain in position and semi-global in orientation, knowing that a global convergence domain in orientation cannot be obtained with continuous feedback laws due to the topology of the rotation space $\mathbb{S}^3$. Several solutions to this problem have been proposed in the past years, under the assumption that the pose (i.e. position and orientation) is completely known. This work concerns the case when the relation between the “position measurements” and the true position vector is uncertain. In practice, such uncertainties are related, e.g., to ill-calibrated sensors or to incomplete knowledge of the environment in the case of proximetry sensors. It is assumed that position measurements are given by $\overline{p} = Mp$ where $p$ is the true position error with respect to the reference position, and $M$ is an unknown invertible matrix. As a first contribution, we propose a class of feedback laws that achieve semi-global stability of the equilibrium $p = 0$ for any matrix $M$ that satisfies the stability criteria $\|M - I_3\| < \delta(k)$ where $I_3$ is the $3 \times 3$ identity matrix, $\delta$ is a strictly positive function, and $k$ is the vector of control parameters. An explicit expression of the function $\delta$ is provided, thus relating the control parameters to the stability margin. The second contribution of this work is the application of this control approach to the visual servoing of VTOL UAVs with respect to a planar vertical structure (wall, etc). From the homography matrix that relates the current camera image to a reference image (taken at the reference pose), we derive a signal output of the form $\overline{p} = Mp$. The matrix $M$ typically depends on unknown parameters but we show that a very rough knowledge of them is sufficient to design a stable controller based on the above-mentioned stability criterion. These results have been submitted for publication at an international conference. This is a joint work with H. de Plinval and P. Mouyon from ONERA Toulouse.

6.2.3. Development of an autonomous shopping cart

This work, which consists in developing a shopping cart with autonomy capabilities (automatic user following, obstacle avoidance, etc), is part of the national INRIA PAL project (Personally Assisted Living) which aims at developing robotic tools for disabled persons or elderslies.

The architecture of “Autonomous Shopping Cart” has been developed in three layers. The first one is responsible for connecting the services layer to physical (or virtual) devices. During this year, all necessary components to access the devices have been implemented:
• the Phidgets library wraps the API of Phidgets devices and abstracts the access to the peripherals on the wheelchair robot;
• the Hannibal library wraps the interface to access the Hannibal robot (through Carmen library);
• the Simulator library wraps the simulator interface.

All of them expose a common interface to the software modules. Thus, the higher components do not have to be changed if the test platform changes.

The second level is the core of this year’s work. It is composed by the Control module and Modeling module. The Control module aims at stabilizing the trajectory of the robot w.r.t a given reference motion. In practice, this reference motion corresponds to the cart user that needs to be followed, but it could be any virtual reference motion. A first implementation of the Control module has been made using Matlab software. The result is a Control library that contains two different methods for controlling the trajectory:

• position control only;
• full-state control (position+orientation).

These methods have been implemented in Matlab language and then converted in C++. The resulting library has been utilized within the Control module deployed on the robotic platform.

The main objective of the Modeling module is to detect the cart user within the sight of the sensors. This task is generally non-trivial, due to noise in the sensor signals and variations of the environment. For this reason a Multiple Hypothesis Tracker has been used to allow for the presence of several persons in the environment. The method uses the laser scans to extract potential persons and then a Selector algorithm extracts the best hypothesis for the cart user. This hypothesis is then converted into a virtual reference point given to the Control module for trajectory tracking.

The third layer is represented by the Behavior module. This component manages the other modules, starts and stops services on request, enables the initialization procedure and so on. As for now, it starts all modules and initiates the starting procedure. In particular, it selects the first person to be tracked, among possible candidates.

Experiments have been successfully conducted both on the mobile robot Hannibal and on the wheeled walking aid ANG (Assistive Navigation Guide) developed by the EPI Coprin.
6. New Results

6.1. Lighting and Rendering

Participants: Mahdi Bagher, Laurent Belcour, Georges-Pierre Bonneau, Eric Bruneton, Cyril Crassin, Jean-Dominique Gascuel, Olivier Hoel, Nicolas Holzschuch, Fabrice Neyret, Cyril Soler, Fabrice Neyret, Charles de Rousiers, Cyril Soler.

6.1.1. Non-linear Pre-filtering Methods for Efficient and Accurate Surface Shading

Participants: Eric Bruneton, Fabrice Neyret.

Rendering a complex surface accurately and without aliasing requires the evaluation of an integral for each pixel, namely a weighted average of the outgoing radiance over the pixel footprint on the surface. The outgoing radiance is itself given by a local illumination equation as a function of the incident radiance and of the surface properties. Computing all this numerically during rendering can be extremely costly. For efficiency, especially for real-time rendering, it is necessary to use precomputations. When the fine scale surface geometry, reflectance and illumination properties are specified with maps on a coarse mesh (such as color maps, normal maps, horizon maps or shadow maps), a frequently used simple idea is to pre-filter each map linearly and separately. The averaged outgoing radiance, i.e., the average of the values given by the local illumination equation is then estimated by applying this equation to the averaged surface parameters. But this is really not accurate because this equation is non-linear, due to self-occlusions, self-shadowing, non-linear reflectance functions, etc. Some methods use more complex pre-filtering algorithms to cope with these non-linear effects. In [14] we presented a survey of these methods. We have started with a general presentation of the problem of pre-filtering complex surfaces. We then present and classify the existing methods according to the approximations they make to tackle this difficult problem. Finally, an analysis of these methods allows us to highlight some generic tools to pre-filter maps used in non-linear functions, and to identify open issues to address the general problem.

6.1.2. Frequency-Based Kernel Estimation for Progressive Photon Mapping

Participants: Laurent Belcour, Cyril Soler.

We have developed an extension to Hachisuka et al.’s Progressive Photon Mapping (or PPM) algorithm [32] in which we estimate the radius of the density estimation kernels using frequency analysis of light transport [29]. We predict the local radiance frequency at the surface of objects using a Gaussian approximation, and use it to drive the size of the density estimation kernels, in order to accelerate convergence (see Figure 3). The key is to add frequency information to a small proportion of photons: frequency photons. In addition to contributing to the density estimation, they will provide frequency information. This work has been published in [20].

6.1.3. Efficiently Visualizing Massive Tetrahedral Meshes with Topology Preservation

Participant: Georges-Pierre Bonneau.
Figure 3. In this figure we compare against progressive photon mapping with our algorithm for the convergence of an indirectly lit part of the scene. In the closeup, we show that our algorithm produces a lower varying estimate at an earlier stage of its execution. The images were produced using 100,000 photons per pass and 25% of frequency photons to make timing comparable.

Figure 4. Left: Variable resolution visualization of a volume mesh with multiple linear features. The topology of the substructures is guaranteed to be preserved. Right: Snapshot of the multiresolution visualization tool to explore simulation data with embedded structures on a desktop PC.
This work is the result of a collaboration with S. Hahmann from the EVASION team-project and Prof. Hans Hagen partly done during a sabbatical of G.-P. Bonneau in the University of Kaiserslautern, Germany. Interdisciplinary efforts in modeling and simulating phenomena have led to complex multi-physics models involving different physical properties and materials in the same system. Within a 3d domain, substructures of lower dimensions appear at the interface between different materials. Correspondingly, an unstructured tetrahedral mesh used for such a simulation includes 2d and 1d substructures embedded in the vertices, edges and faces of the mesh. The simplification of such tetrahedral meshes must preserve (1) the geometry and the topology of the 3d domain, (2) the simulated data and (3) the geometry and topology of the embedded substructures. This work focuses on the preservation of the topology of 1d and 2d substructures embedded in an unstructured tetrahedral mesh, during edge collapse simplification. We derive a robust algorithm, based on combinatorial topology results, in order to determine if an edge can be collapsed without changing the topology of both the mesh and all embedded substructures. Based on this algorithm we have developed a system for simplifying scientific datasets defined on irregular tetrahedral meshes with substructures, illustrated in Figure 4. We presented and demonstrated the power of our system with real world scientific datasets from electromagnetism simulations in the Springer book chapter [27].

6.1.4. Real-Time Rough Refraction

Participants: Nicolas Holzschuch, Charles de Rousiers.

Figure 5. Left: ground truth with total internal reflection. Right: the result with our Real-time Rough Refraction technique. While images have some differences, the result remains plausible.

We have developed an algorithm to render objects of transparent materials with rough surfaces in real-time, under distant illumination. Rough surfaces cause wide scattering as light enters and exits objects, which significantly complicates the rendering of such materials. We present two contributions to approximate the successive scattering events at interfaces, due to rough refraction: First, an approximation of the Bidirectional Transmittance Distribution Function (BTDF), using spherical Gaussians, suitable for real-time estimation of environment lighting using pre-convolution; second, a combination of cone tracing and macro-geometry filtering to efficiently integrate the scattered rays at the exiting interface of the object. We demonstrate in I3D paper [24] the quality of our approximation by comparison against stochastic raytracing. This work is illustrated in Figure 5.
6.1.5. Interactive Indirect Illumination Using Voxel Cone Tracing

**Participants:** Cyril Crassin, Fabrice Neyret.

Indirect illumination is an important element for realistic image synthesis, but its computation is expensive and highly dependent on the complexity of the scene and of the BRDF of the involved surfaces. While offline computation and pre-baking can be acceptable for some cases, many applications (games, simulators, etc.) require real-time or interactive approaches to evaluate indirect illumination. We present in the Pacific Graphics paper [16] a novel algorithm to compute indirect lighting in real-time that avoids costly precomputation steps and is not restricted to low-frequency illumination. An illustration is given in Figure 6. It is based on a hierarchical voxel octree representation generated and updated on the fly from a regular scene mesh coupled with an approximate voxel cone tracing that allows for a fast estimation of the visibility and incoming energy. Our approach can manage two light bounces for both Lambertian and glossy materials at interactive framerates (25-70FPS). It exhibits an almost scene-independent performance and can handle complex scenes with dynamic content thanks to an interactive octree-voxelization scheme. In addition, we demonstrate that our voxel cone tracing can be used to efficiently estimate Ambient Occlusion. A primer of this work has been published as a poster (Best Poster Award [22]). Insights of the method were given in the Siggraph Talk 2011 [23].

The publication [22] has received the Best Poster Award at I3D’2011.

6.1.6. Fast multi-resolution shading of acquired reflectance using bandwidth prediction

**Participants:** Mahdi Bagher, Laurent Belcour, Nicolas Holzschuch, Cyril Soler.

Shading complex materials such as acquired reflectances in multi-light environments is computationally expensive. Estimating the shading integral involves stochastic sampling of the incident illumination independently at several pixels. The number of samples required for this integration varies across the image, depending on an intricate combination of several factors. Ignoring visibility, adaptively distributing computational budget across the pixels for shading is already a challenging problem. In the paper [28] we present a systematic approach to accelerate shading, by rapidly predicting the approximate spatial and angular variation in the local light field arriving at each pixel. Our estimation of variation is in the form of local bandwidth, and accounts for combinations of a variety of factors: the reflectance at the pixel, the nature of the illumination, the local geometry and the camera position relative to the geometry and lighting. An illustration is given in Figure 7. The speed-up, using our method, is from a combination of two factors. First, rather than shade every pixel, we use this predicted variation to direct computational effort towards regions of the image with high local variation. Second, we use the predicted variance of the shading integrals, to cleverly distribute a fixed total budget of shading samples across the pixels. For example, reflection off specular objects is estimated using fewer samples than off diffuse objects.
Figure 7. The technique developed by Mahdi Bagher allows to predict in real time the local bandwidth of the image obtained by shading a measured material with all frequency distant illumination (See inset colored top-right image). This information allows a drastic economy of samples in the computation of the integrals that are needed to produce an accurate image. In particular this allows to perform multi-sampling anti-aliasing in a deferred shading pipeline with much less image-space samples than the brute-force solution.
6.2. Expressive Rendering and Visualization

Participants: Pierre Bénard, Georges-Pierre Bonneau, Alexandre Coninx, Joëlle Thollot.

6.2.1. Temporal Coherence for Stylized Animation

Participants: Pierre Bénard, Joëlle Thollot.

Figure 8. In our state-of-the-art report we review and carefully compare Temporal Coherence techniques for stylized animations.

Non-photorealistic rendering (NPR) algorithms allow the creation of images in a variety of styles, ranging from line drawing and pen-and-ink to oil painting and watercolor. These algorithms provide greater flexibility, control and automation over traditional drawing and painting. Despite significant progress over the past 15 years, the application of NPR to the generation of stylized animations remains an active area of research. The main challenge of computer generated stylized animations is to reproduce the look of traditional drawings and paintings while minimizing distracting flickering and sliding artifacts present in hand-drawn animations. These goals are inherently conflicting and any attempt to address the temporal coherence of stylized animations is a trade-off. We have published the state-of-the-art report [15] motivated by the growing number of methods proposed in recent years and the need for a comprehensive analysis of the trade-offs they propose. We formalize the problem of temporal coherence in terms of goals and compare existing methods accordingly. We propose an analysis for both line and region stylization methods and discuss initial steps toward their perceptual evaluation. The goal of our report is to help uninformed readers to choose the method that best suits their needs, as well as motivate further research to address the limitations of existing methods.

6.2.2. Visualization of data with uncertainty using perceptually guided procedural noise

Participants: Alexandre Coninx, Georges-Pierre Bonneau.

This work is the result of a collaboration with EdF R&D and Jacques Droulez, Director of Research at CNRS in Collège de France. In his PhD work, Alexandre Coninx has introduced a new method to visualize uncertain scalar data fields by combining color scale visualization techniques with animated, perceptually adapted Perlin noise. The parameters of the Perlin noise are controlled by the uncertainty information to produce animated patterns showing local data value and quality, as illustrated in Figure 9. In order to precisely control the perception of the noise patterns, we perform a psychophysical evaluation of contrast sensitivity thresholds for a set of Perlin noise stimuli. We validate and extend this evaluation using an existing computational model. This allows us to predict the perception of the uncertainty noise patterns for arbitrary choices of parameters. We demonstrate and discuss the efficiency and the benefits of our method with various settings, color maps and data sets. This work has been published at APGV’2011 [21].
Figure 9. Left: classical colormap visualization of scalar data without uncertainty. Right: in our technique, we perturb the input of the colormap using a perceptually guided procedural noise, scaled by the uncertainty of the data. The data and its uncertainty can be visualized in the same image.

6.3. Modeling and Animation

Participants: Georges-Pierre Bonneau, Alexandre Derouet-Jourdan, Nicolas Holzschuch, Nassim Jibai, Cyril Soler, Joëlle Thollot.

6.3.1. Multiscale Feature-Preserving Smoothing of Tomographic Data

Participants: Nassim Jibai, Nicolas Holzschuch, Cyril Soler.

Figure 10. Left: A contour surface extracted from noisy tomographic data contains surface noise and several topological artifacts such as small handles and holes. Right: The surface extracted from our smoothed volume is clean, and yet small features, such as the thread in the screw, and sharp edges have been preserved.

Computer tomography (CT) has wide application in medical imaging and reverse engineering. Due to the limited number of projections used in reconstructing the volume, the resulting 3D data is typically noisy. Contouring such data, for surface extraction, yields surfaces with localised artifacts of complex topology. To avoid such artifacts, we propose a method for feature-preserving smoothing of CT data, illustrated in Figure 10. The smoothing is based on anisotropic diffusion, with a diffusion tensor designed to smooth
noise up to a given scale, while preserving features. We compute these diffusion kernels from the directional histograms of gradients around each voxel, using a fast GPU implementation. This work has been published as a Siggraph’2011 Poster [26].

6.3.2. 3D Inverse Dynamic Modeling of Strands
Participants: Alexandre Derouet-Jourdan, Joëlle Thollot.

![Figure 11. Our model can physically animate the tail. The tail will retrieve its initial shape at the end of slight (possibly strong) motions.](image)

In this work, we propose a new method to automatically and consistently convert 3D splines into dynamic rods at rest under gravity, bridging the gap between the modeling of 3D strands (such as hair, plants) and their physics-based animation. An illustration is given in Figure 11. This work is done in collaboration with F. Bertails from the BIPOP team-project. It has been published in a Siggraph’2011 poster [25].

6.3.3. Lagrangian Texture Advection: Preserving both Spectrum and Velocity Field
Participants: Eric Bruneton, Nicolas Holzschuch, Fabrice Neyret.

Texturing an animated fluid is a useful way to augment the visual complexity of pictures without increasing the simulation time. But texturing flowing fluids is a complex issue, as it creates conflicting requirements: we want to keep the key texture properties (features, spectrum) while advecting the texture with the underlying flow — which distorts it. In this context we present a new, Lagrangian, method for advecting textures: the advected texture is computed only locally and follows the velocity field at each pixel (see Figure 12). The texture retains its local properties, including its Fourier spectrum, even though it is accurately advected. Due to its Lagrangian nature, our algorithm can perform on very large, potentially infinite scenes in real time. Our experiments show that it is well suited for a wide range of input textures, including, but not limited to, noise textures. This work has been published in the IEEE Transactions on Visualization and Computer Graphics (TVCG) [18].

6.3.4. Feature-Based Vector Simulation of Water Waves
Participant: Fabrice Neyret.

We have developed a method for simulating local water waves caused by obstacles in water streams for real-time graphics applications. Given a low-resolution water surface and velocity field, our method is able to decorate the input water surface with high resolution detail for the animated waves around obstacles. We construct and animate a vector representation of the waves. It is then converted to feature-aligned meshes for capturing the surfaces of the waves (see Figure 13). Results demonstrate that our method has the benefits of real-time performance and easy controllability. The method also fits well into a state-of-the-art river animation system. This work has been published in the Journal of Computer Animation and Virtual Worlds [19].
Figure 12. Left: Our method advects open-domain textures preserving both the spectrum and the motion field, in real-time. Right: Various applications in 2D and 3D, with procedural, image, bump, displacement textures.

Figure 13. Our method permits the real-time rendering of highly detailed animated features on large river scenes.
6.3.5. Volume-preserving FFD for Programmable Graphics Hardware

Participant: Georges-Pierre Bonneau.

Figure 14. Left: FFD deformation of the armadillo mesh, without volume preservation. Right: our technique: GPU-based volume preservation of the FFD deformation.

This work is the result of a collaboration with S. Hahmann from the EVASION team-project, Prof. Gershon Elber from Technion and Prof. Hans Hagen from the University of Kaiserslautern.

Free Form Deformation (FFD) is a well established technique for deforming arbitrary object shapes in space. Although more recent deformation techniques have been introduced, amongst them skeleton-based deformation and cage based deformation, the simple and versatile nature of FFD is a strong advantage, and justifies its presence in nowadays leading commercial geometric modeling and animation software systems. Since its introduction in the late 80’s, many improvements have been proposed to the FFD paradigm, including control lattices of arbitrary topology, direct shape manipulation and GPU implementation. Several authors have addressed the problem of volume preserving FFD. These previous approaches either make use of expensive non-linear optimization techniques, or resort to first order approximation suitable only for small-scale deformations. In this work we take advantage from the multi-linear nature of the volume constraint in order to derive a simple, exact and explicit solution to the problem of volume preserving FFD (see Figure 14). Two variants of the algorithm are given, without and with direct shape manipulation. Moreover, the linearity of our solution enables to implement it efficiently on GPU. The results have been published in a Visual Computer journal paper [17].
6. New Results

6.1. Visual Analytics of EA Data

Participants: Jean-Daniel Fekete, Évelyne Lutton [correspondant].

An experimental analysis of Evolutionary Algorithms (EAs) usually generates a huge amount of multidimensional data, including numeric and symbolic data. It is difficult to efficiently navigate in such a set of data, for instance to be able to tune the parameters or evaluate the efficiency of some operators. Usual features of existing EA visualisation systems consist in visualising time- or generation-dependent curves (fitness, diversity, or other statistics). When dealing with genomic information, the task becomes even more difficult, as a convenient visualisation strongly depends on the considered fitness landscape. In this latter case the raw data are usually sets of successive populations of points of a complex multidimensional space.

The purpose of this study was to evaluate GraphDice on complex sets of EA data (for artificial and real test-cases), and to sketch future developments of this tool, in order to better adapt it to the needs of EA experimental analysis (Fig. 6). An output of this study is the acceptation of the EASEA-Cloud ANR-Emergence project, in which developments will aim at adding tools in GraphDice specific for:

- visualisation of Evolutionary Algorithms written in the EASEA language,
- monitoring the execution of these algorithms on a cloud of computers (CPU + GPU).
6.2. Interactive Evolutionary Algorithms for Visual decision making

Participants: Nadia Boukhelifa, Waldo Cancino, Jean-Daniel Fekete, Evelyne Lutton [correspondant].

When dealing with very large datasets with many dimensions, it is often difficult to efficiently navigate and find interesting viewpoints, significative compound variables, unexpected behaviour, and other remarkable characteristics.

Our aim within the System@tic CSDL project (Complex Systems Design Lab, 2009–2012) is to use interactive evolutionary algorithm to assist the user in its exploration task. Finding an interesting, non obvious, viewpoint on a complex dataset can be formulated as an interactive optimisation problem. Population-based evolutionary search mechanisms can then efficiently be exploited for suggesting new viewpoints on data, that progressively adapt to the needs of an user.

In September 2011 (arrival of Nadia Boukhelifa and Waldo Cancino) we started to build a prototype based on GraphDice, that proposes new dimensions in the scatterplot matrix. These secondary set of dimensions are compositions of the dimensions of the initial dataset. Starting from an initial set of suggested dimensions (PCA analysis of the dataset), an evolutionary algorithm progressively refines the compound dimensions according to a measurement of the activity of the user on the corresponding views.

6.3. Optimisation of Food Models

Participant: Évelyne Lutton [correspondant].

In collaboration with Alberto Tonda and Romain Reuillon, ISC-PIF

The European project DREAM ( http://dream.csregistry.org/ ) managed by INRA-CEPIA, aims at building decision support tools for better managing product quality and, by the way, manufacturing processes in the domain of agrifood industry.

Our contribution to this project is focused on the evolutionary optimisation of Bayesian Networks models, on the development of efficient cooperative-co-evolution schemes to solve some food modeling problems (milk gel, cheese ripening), and on the efficient visualisation of output data of these algorithms.

6.4. A Study on Dual-Scale Data Charts

We presented the results of a user study that compares different ways of representing dual-scale data charts (see Fig. 7). Dual-scale charts incorporate two different data resolutions into one chart in order to emphasize data in regions of interest or to enable the comparison of data from distant regions. While some design guidelines exist for these types of charts, there is currently little empirical evidence on which to base their design. We filled this gap by discussing the design space of dual-scale cartesian-coordinate charts and by experimentally comparing the performance of different chart types with respect to elementary graphical perception tasks such as comparing lengths and distances. Our study suggests that cut-out charts which include collocated full context and focus are the best alternative, and that superimposed charts in which focus and context overlap on top of each other should be avoided.

6.5. Information Visualization Evaluation

Petra Isenberg has contributed to three articles on evaluation methodologies: The first article “Collaborative Visualization: Definition, Challenges, and Research Agenda” [12] deals with challenges of collaborative visual analytics and includes a discussion on the challenges of evaluating tools during multi-person use. The second article, “Information Visualization Evaluation in Large Companies: Challenges, Experiences and Recommendations” [14], discusses challenges of evaluating and deploying visual analytics tools in a large company setting. It lists several challenges and provides concrete guidance to others who seek to evaluate tool within domain experts in their work environment. Finally, “Seven Guiding Scenarios for Information Visualization Evaluation” [33] is a pre-print of a journal article (in press) which provides a new viewpoint on evaluation in information visualization. Instead of giving an overview of methods, it cites evaluation goals and questions and can, thus, provide clear considerations for practitioners and researchers in the area.
5. New Results

5.1. Introduction

This year we obtained new results in our three sub-objectives and also related to Focuslab platform and software valorization:

1. **Sub-objective 1 - Mining for Knowledge Discovery in Information Systems**: this we get five results (with one achieved PhD thesis).

   Let us note that six past works on this sub-objective described in previous AxIS annual reports have been published this year as articles in international journals ([22],[11]) or conferences, one in a national journal [50], two in a french-speaking conference [35],[46], one book [20] and one book chapter [52] at international level. Indeed

   - The work in 2009 on mining data streams by Marascu in her thesis [96] has been published in [11] with more details in the algorithms and in the experiments.
   - The work published in 2008 on discovering frequent behaviors [107] has been published in [22] with more details in the algorithms and in the experiments.
   - Our previous work on satellite image mining in 2010 [99] has been published in French at EGC [35].
   - Our past work on Functional data analysis involving data described by regular functions rather than by a finite number of real valued variables has been published as a scientific book chapter [52]. In this paper we propose to use a clustering approach that targets variables rather than individual to design a piecewise constant representation of a set of functions. The contiguity constraint induced by the functional nature of the variables allows a polynomial complexity algorithm to give the optimal solution.
   - In the context of the WRUM project (Morocco) and Zemmouri’s PhD thesis, we have a long paper accepted at JFO 2011 related to past works (2010) on how to integrate domain knowledge in a multi-view KDD process [46].

2. **Sub-objective 2 - Information and Social Networks Mining for Supporting Information Retrieval**: Three results (with one achieved Ph-D thesis). Let us remind the best paper [15] obtained By E. Smirnova at ECIR 2011 for her research on expert finding.

3. **Sub-objective 3 - Multidisciplinary Research For Supporting User Oriented Innovation**: this interdisciplinary research is dedicated to the design, tailoring and refinement of methodologies and tools for a better users’ involvement in innovation processes. We have seven results this year..

Concerning our activity in terms of **FocusLab Experimental Platform and Software**, a) we first applied ATWUEDA on another context of evolutive data (on system monitoring data at EDF) which is different of Web usage data) to show the genericity of the approach [9], and b)we develop a Web-based version of the FocusLab experimental platform for analysis usage data ( hardware and software parts).

5.2. Mining for Knowledge Discovery in Information Systems

5.2.1. Mining Data Streams: Clustering and Pattern extraction

**Participant**: Chongsheng Zhang.
In Zhang’s thesis [19] (supervised by F. Masseglia), which was partially founded by ANR MIDAS (cf. 6.2.1), we present our study of the management and mining issues on data streams with evolving tuples, caused by model updates or tuple revisions. For instance, in an online auction system where bids on auction items are streaming, it is possible that some users may bid for more than one item within the user-specified time interval. As a result, the profiles of the users can be updated or revised in such applications. Data streams having evolving tuples bring new challenges as well as research opportunity. In this work, he develops novel and efficient models and methods for managing and mining data streams with evolving tuples. (I) To model data streams with evolving tuples, we propose the Anti-Bouncing Streaming model (ABS) for usage streams. ABS fits data streams with evolving tuples and it enables methods for processing of data streams to handle tuple updates or revisions. (II) To find frequent itemsets from data streams with evolving tuples over pane-based sliding windows, we conduct theoretical analysis and propose theorems which can avoid scanning the past slides to check for possible itemsets that may become frequent. We also design novel data structures which can manage the data streams with evolving tuples efficiently and facilitate the mining of frequent itemsets. Moreover, we devise an efficient counting algorithm to verify the frequentness of the candidate frequent itemsets. We also propose two running frameworks for this problem. (III) To extract important feature set from data streams (including the ones with evolving tuples), based upon ABS, we devise the streaming feature set selection algorithm for data streams which is the first in the literature. This method is based on information theory to extract the informative feature sets. To further accelerate the extraction of the most informative feature set from high-dimensional data, we propose a framework that reduces the huge search space to a rather small subset while still guarantee the quality of the discovered feature sets.

In 2011, Chongsheng Zhang has mainly worked on a data stream mining method, intending to extract frequent itemsets. This method has not been published yet and is described in Chapter 5 (page 79) of his thesis document [19].

5.2.2. Clustering on Multiple Dissimilarity Matrices
Participants: Yves Lechevallier, Francisco de A.T. de Carvalho, Thierry Despeyroux, Alessandra Silva Anyzewski.

In [23] we introduce hard clustering algorithms that are able to partitioning objects taking into account simultaneously their relational descriptions given by multiple dissimilarity matrices [49]. The aim is to obtain a collaborative role of the different dissimilarity matrices in order to obtain a final consensus partition. These matrices could have been generated using different sets of variables and a fixed dissimilarity function or using a fixed set of variables and different dissimilarity functions, or using different sets of variables and dissimilarity functions.

These methods, which are based on the dynamic hard clustering algorithm for relational data as well as on the dynamic clustering algorithm based on adaptive distances, are designed to furnish a partition and a prototype for each cluster as well as to learn a relevance weight for each dissimilarity matrix by optimizing an adequacy criterion that measures the fitting between clusters and their representatives.

These relevance weights change at each algorithm iteration and can either be the same for all clusters or different from one cluster to another. The usefulness of these partitioning hard clustering algorithms are shown on two time trajectory real world datasets.

5.2.3. Clustering of Constrained Symbolic Data
Participants: Marc Csernel, Francisco de A.T. de Carvalho.

In the context of our FACEPE collaboration with Brazil (cf. section 6.4.3.1), we have presented a method which allows clustering of symbolic descriptions constrained by presence rules in a polynomial time instead of a combinatorial one. This method allows to deal with “false missing values”. Such a method can be applied on various classification problems [26].

5.2.4. Web Page Clustering based on a Community Detection Algorithm
Participants: Yves Lechevallier, Yacine Slimani.
Extracting knowledge from Web user’s access data in Web Usage Mining (WUM) process is a challenging task that is continuing to gain importance as the size of the web and its user-base increase. That is why meaningful methods have been proposed in the literature in order to understand the behaviour of the user in the web and improve the access modes to information. In this work [42], we are interested in the analysis of the user browsing behavior. The objective is to understand the navigational practices of users (teachers, students and administrative staff). First we clean the data by removing irrelevant information and noise. During the second step, remaining data are arranged in a coherent way in order to identify user sessions. After we defined a new approach [42] of knowledge extraction. This approach treats the data resulting from the preprocessing phase (first and second steps) as being a set of communities. Our approach extends the Modularity measure, proposed by Newman and Girvan [97], in the Web Mining context in order to benefit from their classifying capacity in the communities discovery.

This work is done in collaboration with the LRIA laboratory – Université Ferhat Abbas, Sétif, Algérie

5.2.5. Critical Edition of Sanskrit Texts

Participants: Marc Csernel, Nicolas Béchet, Ehab Hassan, Yves Lechevallier.

New progresses concerning the computer assisted elaboration of Sanskrit texts have been made. First Nicolas Béchet and Marc Csernel have worked on the problem of moved texts. After an alignment between two versions of the texts, we discover that some parts of the text appears to have been moved according to the technics developed in [48]. Until now, we were not able to discover when a text has been moved in a manuscript.

Now using a words-grams technique proposed in [48], we were able to obtain quite good results on the moved texts problem and we were able to optimize the different possible parameters. A paper on the subject has been submitted to the Cicling 2012 conference (http://www.cicling.org/2012/).

After the new treatment related to the moved text problem, we need to provide an interactive display of the critical edition. During his internship, Ehab Hassan has been working on the subject and obtained good results. These results need to be deeply examined by Sanskritists to see if they always fulfill their needs.

5.3. Information and Social Networks Mining for Supporting Information Retrieval

5.3.1. Clustering of Relational Data and Social Network Data

Participants: Yves Lechevallier, Amine Louati.

The automatic detection of communities in a social network can provide this kind of graph aggregation. The objective of graph aggregations is to produce small and understandable summaries and can highlight communities in the network, which greatly facilitates the interpretation.

Social networks allow having a global view of the different actors and different interactions between them, thus facilitating the analysis and information retrieval.

In the enterprise context, a considerable amount of information is stored in relational databases. Therefore, relational database can be a rich source to extract social network. The extracted network has in general a huge size which makes its analyses and visualization difficult tasks. In [45], we propose a social network extraction approach from relational database.

Often, the network has a large size which makes its analysis and visualization difficult.

The aggregation step is a necessary task, so we offer [33] and [32] an aggregation step based on the k-SNAP algorithm [109] that produces a summary graph by grouping nodes based on attributes and relationships selected by the user.

This work is done in collaboration with Marie-Aude Aufaure, head of the Business Intelligence Team, Ecole Centrale Paris, MAS Laboratory.
5.3.2. Networks Solutions for Expert Finding and People Name Disambiguation

**Participants:** Elena Smirnova, Yi-Ling Kuo, Brigitte Trousse.

The task of finding people who are experts on a given topic has cently attracted close attention. State-of-the-art expert finding algorithms uncover knowledge areas of candidate experts based on textual content of associated documents. While powerful, these models ignore social structure that might be available. Therefore, we develop a Bayesian hierarchical model for expert finding that accounts for both content and social relationships. The model assumes that social links are determined by expertise similarity between candidates. The results of EGC experiments on UvT expert collection have demonstrated the effectiveness of our algorithm [43].

E. Smirnova visited Intellius, people search technology company (Aug 8 - Oct 5, 2011): the goal of this visit was to validate the research on expert finding in social networks on real dataset and further advance it. As a real dataset, we have taken a sample of United States LinkedIn public profiles. We built an organizational network by connecting a LinkedIn user and his colleagues at different workplaces. We also constructed a geographical network from user’s current location in the United States. We used Amazon’s Mechanical Turk framework (http://aws.amazon.com/code/923) to collect user-oriented judgements for model evaluation. We found that the user-oriented model is statistically significantly preferred to the baseline model on 72.5% of queries.

Her work on name disambiguation done in 2010 has been integrated in an article related to the problem of quick detection of top-k Personnalized PageRank (PPR) in [24]. The effectiveness of the chosen approach based on Monte Carlo methods for quick detection of top-k PPR lists has been demonstrated on the Web and Wikipedia graphs.

Yi-Ling Kuo during her internship has worked on Person Name Disambiguation and started by managing the analysis of the very huge Yahoo! Web graph.

This topic has been done in the context of Smirnova’s thesis [18] which has been defended on december 15 (thesis supervised by B. Trousse (AxIS) and K. Avrachenkov (Maestro)).

5.3.3. Towards an On-Line Analysis of Tweets Processing

**Participant:** Nicolas Béchet.

Tweets exchanged over the Internet represent an important source of information, even if their characteristics make them difficult to analyze (a maximum of 140 characters, etc.). In [25], we define a data warehouse model to analyze large volumes of tweets by proposing measures relevant in the context of knowledge discovery. The use of data warehouses as a tool for the storage and analysis of textual documents is not new but current measures are not well-suited to the specificities of the manipulated data. We also propose a new way for extracting the context of a concept in a hierarchy. Experiments carried out on real data underline the relevance of our proposal.

This work is done inside a collaboration with LIRMM and CEMAGREF.

5.4. Multidisciplinary Research For Supporting User Oriented Innovation

5.4.1. Usability Design and Evaluation Methods

**Participants:** Dominique Scapin, Yves Lechevallier, Pascal Marie-Dessoude, Claudia Detraux.

We pursued our work on articulation of usage mining approach and human factors expertise for the design and evaluation of information systems. Namely, collaborative clustering techniques were used to analyze data issued from users via a card sorting technique, with respect to an a priori (“expert”) clustering. Considering the difficulties that people have in managing large information sets in their everyday life, for either professional or non-professional purposes (administration, social relationships, leisure, etc.), our recent research focuses on personal information space for which information bits are currently scattered many places.
In this PIMs field, there is little research with a user-centric approach, with the view that users-based knowledge might help specifying computer-based tools and a state-of-the-art [37] showed little work specifically on usability. Studies address a variety of questions from user needs to accessibility (including studies on older people) or user acceptance, among others. In the context of user-centered and long-term studies to understand the evolution of user information practice, we looked in a study at the intuitive way people organize their personal information, with or without computer systems, in order to help the design of future systems. Also, we recently surveyed 15 tools that claim to support personal information management.

5.4.2. Living Lab Landscape

Participants: Marc Pallot, Brigitte Trousse, Bernard Senach, Dominique Scapin.

In order to provide to the research community a comprehensive landscape of research streams in the Living Lab domain, we launched a study on the state-of-the-art about the ubiquitous notion of User Experience. During this continuous study, a landscape [34], [69] has progressively emerged that we organized through 4 main axes: focus granularity (individual/group), user’s role in the design process (observed subject/value creator), collaboration style (structured/unstructured), and evaluation purpose (reliability/adoptability). Our landscape of research streams has been used by Finnish colleagues who conducted an empirical study on the use of the Living Lab research domain landscape as a tool for assessing the maturity level of 16 Finnish Living Labs [36]. The Living Lab research domain Landscape has allowed the study team to identify four categories of Living Labs.

5.4.3. Future Internet Domain Landscape

Participants: Marc Pallot, Brigitte Trousse, Bernard Senach.

There are many different Internet research areas and corresponding technologies that were already investigated, experimented and progressively deployed such as peer-to-peer, autonomous, cognitive and ad hoc networking, that have already demonstrated how to improve network performance and user experience. Peer-to-peer networking for large-scale distributed systems and widely used applications has proved both the feasibility and economic potential for delivering services to millions of users. Others emerged more recently in the context of the future Internet (FI), such as Cloud Computing for transparently sharing among users scalable elastic resources over a limitless network. As it remains difficult to visualise the conceptual evolution and articulate the various Internet research areas, we conducted a study for identifying the appropriate concepts that could populate the FI domain landscape [69], [71] over three different periods of time (1990-1999, 2000-2005 and 2006-2011). Several INRIA research teams are involved in FIRE (Future Internet Research Experimentation) Testbed projects, namely: PlanetLab, OneLab, TEFIS, SensLAB, and BonFIRE whose scientific leaders were interviewed during the development of the FI domain landscape. Four dimensions were used for landscaping the Future Internet research domain: evolution approaches (from incremental evolution design to Clean Slate re-design or radical evolution), Internet routing (from the basic data packet delivery towards more sophisticated content distribution and retrieval capacities such as content Centric Networking), network type (from wired communication to wireless communication networks), evolution trend (from computer network towards network computing). The resulting tentative landscape of FI research areas shown in Figure 1 is intended to provide a faster and broader understanding of the different Internet research streams and related topics.

5.4.4. Future of Internet and User-open innovation for Smart Cities

Participants: Caroline Tiffon, Marc Pallot, Brigitte Trousse, Bernard Senach.

The goal of the Fireball project is to bring together three different constituencies: user driven open innovation, Future Internet, and Smart Cities [34], [39]. It aims at defining a roadmap [29], based on analysis of needs, opportunities and gaps, to benefit a wide scale implementation of the methodologies and concepts elaborated. A first objective in the project was to get a clear picture of the state-of-the-art in each domain. During the review, progressively emerges a landscape [69] that we organized along 4 main axes: wiring (wired/wireless), user’s role (subject/actor), Internet evolution approach (structured/unstructured), evaluation purpose (reliability/adoptability). A large variety of FI research have been engaged. If initial efforts in Future
Internet research have been directed towards the goal of providing the technical infrastructure supporting the next network generation, a rising trend in this research field is to consider now a higher level layer, the layer of services.

5.4.5. Method and Tool for Selection of Open Innovation Software Tools

Participants: Mylène Leitzelman, Brigitte Trouse.

In spite of an important number of tools supporting open innovation, there is few comparative evaluation and no grid or evaluation criteria helping to choose a product. A 2011 review in the Computer Aided Innovation field provides a large overview of available tools in relation with a wide range of innovation cycle features. This useful top-down categorization approach is of little help to choose a specific tool. To find the best OI tool supporting idea exchanges among a community of participants, we built an exploratory method on the Web and we elaborate an assessment grid of OI tools based on the QSOS method which is a method designed to qualify, select and compare free and open source software in an objective, traceable and argued way. It publicly available under the terms of the GNU Free Documentation License. In our QSOS method, evaluation criteria are organized in a tree-hierarchy grid with, a scoring method procedure of each tree-leaf criteria (from 0 - not covered to 2 - completely covered). To achieve the construction of the OI tools criteria assessment grid (first step), we used different mining tools for Web crawling, network analysis, criteria classification and from the 29 top rated OI tools, we finally selected 6 of them from which we extracted the tree-map categorization used to build the reference software criteria sheet. In the following steps, after appropriate weighting, we used the provided OS3 Web application and we were able to compare 4 top selected OI tools.

To support our method, we developed a QSOS-based OI Grid for supporting the OI tool selection. The QSOS Grid to compare OI tools, which is in an XML format, has been translated into a MindMap. As a first mock-up, the QSOS Web Interface O3S will be installed at the beginning of 2012 on our server. For the future, we

---

6 Hüsig, University Regensburg & Khon, Otto Beisheim School of Management
will extend our own QSOS comparative method to other softwares and to other categories interesting for the Living Lab community.

5.4.6. Extension of Methods for Idea Generation Process

Participants: Anne-Laure Negri, Caroline Tiffon, Bernard Senach, Brigitte Trousse.

Internet of Things (IoT) is supposed to be a physical world where everyday objects, rooms and machines are connected to one another and to the larger digital world. In this web of people and objects, individuals as well as things will have their own unique URL and, according to interactions between all these entities, social networks will be articulated with a ring of connected objects. This mash up of “hyper groups” and “hyper objects” will be the next step towards a deeper level of automation in which the user interface has vanished, the explicit control over the world is no longer the rule and where the users will have to understand the dynamic changes of their environment in relation to their social interactions and to their physical behaviour. The design of IoT based services therefore raises many challenges related to the user experience and requires a deep understanding of users’ needs in their real life environment as well as many field experiments; this is why the Living Lab approach appears as the qualification device which has been lacking so far. As described in a paper accepted for ServDes2012 conference (February 2012) among 80 papers submitted, we developed and tested two methods designed for Internet of Things service idea generation: GenIOT (Generative Technique for the Internet of Things) and Aloha! (Animation Lens: Object/Human Actoring!) methods.

- The GenIOT method aims at providing citizen a tool for identifying and sharing examples of situations where they may benefit from an Internet of Things based service. Indeed the objective is to help the introspective effort of analysing one’s own daily behavior and unveil situated data in order to develop grounded internet of things services ideation. Fake sensors are placed by the participants in their daily environment and pictures of the sensors in situ are shared and tagged on a collaborative platform. Other participants are invited to discuss online the ideas behind the pictures. Gamification rules are implemented in the platform in order to enhance collaboration and participation.

- The Aloha! method tackles another creative dimension as it asks participants to role or play characters or objects and bodystorm a collaborative scenario meeting the characters need. Participants report it to be a fun and effective method, alternative to traditional brainstorming and integrating serendipity as a creative asset.

For the exploration step in the context of ELLIOT, we developed NiceAir, an android mock-up for visualizing air quality data on Google map in Nice as well as some interest points (free bikes, free park places, bus stop, etc.). Such a mock-up will be available in the co-creation step of green services in the ELLIOT project and during the european Mobility Week.
5.4.7. Mock-ups for two innovation processes (exploration step)  
**Participants:** Lucile Gramusset, Guillaume Pilot, Mohamed Gaieb, Bernard Senach, Brigitte Trousse.

In the context of two contracts (TICTAC and ELLIOT) related to user oriented innovation process, we have developed two mock-ups:

- Based on the user feedback from the first experiment in TICTAC (cf. section 6.1.1), we decided to develop a mock-up *MobilTIC* of a real-time information service related to public transportation (Envibus & SNCF) usable for any smartphone with internet access for Sophia-Antipolis workers. We designed a Web interface the most simple based on PHP/Mysql technologies and accessible from a computer or 3G mobile phone. Usage analysis of MobilTIC has been anticipated by providing rich logs. An exploration task with citizen is planned in January 2012 with a new experiment with an improved version.

- A first mock-up called *Nice Air* has been developed in the context of ELLIOT (cf. section 6.3.1.1) by L. Gramusset and M. Gaieb on android smartphone for providing information related to air quality and noise on a map of the area of Nice Cote d’Azur with some interest points such as bus stops and Vélib parkings.

5.5. FocusLab Experimental Platform and Software

5.5.1. ATWUEDA based Clustering Approach for System Monitoring  
**Participant:** Yves Lechevallier.

Progressive advances in hardware and software technologies have enabled the production and storage of system monitoring data streams in a wide range of fields (e.g. telecommunications, sensor networks, etc.). Traditional clustering methods are unable to deal with data of such a voluminous and dynamic nature. In this work [51], we propose an efficient clustering approach (ATWUEDA) for monitoring massive time-changing data streams. This work considers a real case study on condition monitoring data streams of an electric power plant provided by EDF.

This work is done in collaboration with Alzennyr Da Silva of BILab laboratory (Telecom ParisTech and EDF R&D Common Laboratory)

5.5.2. FocusLab Experimental Platform (CPER Telius 2008-2012)  
**Participants:** Xavier Augros, Mohamed Gaieb, Brigitte Trousse, Yves Lechevallier.

The FocusLab platform aims to be a major delivery mechanism of previous and current work in AxIS. It is a way to make methodological contributions (including software) available for the scientific community, but also a way for stimulating further research. This work has slowly started on the software part due to the absence of human resource funding and due to the absence of engineers in the team until 2010. Mid 2011, we started the specification and the development of a first version of a platform with three parts (hardware, software and methods) with the arrival of two engineers on the ELLIOT and TIC TAC contracts. A first version is available since the end of September as a Web portal and a second one is planned for the end of 2012 with advanced features for the software part (cf. service oriented platform, SOA architecture and interoperability).

Related to the software part we are in the process of developing several AxIS methods as Web services: we started with SCDS (cf. 4.2.3) which was demonstrated in the context of the MIDAS project on two applications (Orange labs mobile portal and vehicle trajectories) and in the ELLIOT platform (linked to the San rafaelle Hospital media use case) at the first review meeting (cf. 6.3.1.1) applied on data issued from San Rafealle Hospital use case. ATWUEDA (cf. 4.3.2), GEAR (Marascu’s thesis) are under development as web services.

Our work on mining evolutive data (ATWUEDA) and data streams (such as SCDS) have been used in real applications in the context of Internet of things and sensors: ATWUEDA (system monitoring for EDF (cf. section 5.5.1) and SDMS (cf. section 4.2.3).
6. New Results

6.1. Robotics

6.1.1. Calibration and identification

6.1.1.1. Calibration of a cable-driven robot

Participants: David Daney, Julien Alexandre dit Sandretto, Jean-Pierre Merlet, Gilles Trombettoni.

To improve the accuracy of a cable manipulator, it is necessary to identify the uncertainties of its model. The robots, studied in Cogiro, an ANR National initiative, are redundantly actuated: the number of powered wires is larger than the number of degrees of freedom of the manipulator. Under some cable properties hypothesis, this over-constraint mechanism allows to perform a self-calibration - i.e. the identification of the parameters does not need additional external measurement. A first experimentation, done in Montpellier, validated a novel approach which consists in a simultaneous identification of parameters and robot position (unknown in self-calibration process).

6.1.1.2. Cable properties

Participants: Julien Alexandre dit Sandretto, Gilles Trombettoni, David Daney.

The majority of researches done on cable-driven robot modeling need to take into account that the mass and the elasticity of wires are neglectable. However, they can not prove that these hypotheses are acceptable regarding these objectives. We have proposed an algorithm based on interval analysis to judge the validity of these assumptions for a cable-robot in a specific workspace. This method have been tested on the Tecnalia/LIRMM’s prototype and used for the construction of the Cogiro robot.

6.1.1.3. Optimal calibration poses of a 3-RPR planar parallel robot

Participant: David Daney.

The choice of the measurement configuration is crucial to improve the robustness of the calibration for measurement uncertainties. This year, a geometrical approach has been used to determine formally the set of the optimal poses for the identification of the kinematic parameter of a 3-RPR planar parallel robot. This result is important because it explains the influence of the location of some particular robot poses in their workspaces during the model identification process. A generalization is explored to construct automatically an optimal set for robot calibration and moreover, to improve experimental design algorithms. The aim is now to obtain similar results for cable driven robots calibration.

6.1.1.4. Geometric calibration of a space telescope

Participants: Thibault Gayral, David Daney, Jean-Pierre Merlet.

In October 2010 begun a collaborative work with Thales Alenia Space on the calibration of the mechanical structure of a space telescope. Its architecture is based on a parallel manipulator (type active wrist 6-PUS) used to correct the relative position of two mirrors. The aim is to identify the parameters of this robot, to improve its accuracy and then increase the quality of the images provided by the telescope. Thus, a geometric calibration procedure was considered and a campaign of photogrammetry was performed on the telescope. Using a kinematic description, a final accuracy of at worst 10 \( \mu m \) was reached on the position of the platform of the telescope. The aim of these measures were also to valid or not the model of comportment of the flexible parts of the device. This campaign brought to light the necessity to consider forces and torques acting on the structure in the deformation of the flexible parts in order to reach a submicrometric accuracy.

6.1.1.5. Modelization of flexible articulations of the telescope

Participants: Thibault Gayral, David Daney, Jean-Pierre Merlet.
In order to improve the final accuracy of the above-mentioned space telescope, a novel model including the statics equations in order to calculate deformations of the flexible articulations is currently under study. The main difficulty is to identify parameters (stiffness matrix and geometric parameters) that have different units and are not of the same order of magnitude. To solve this issue, we are focusing our effort to write the problem in a better robust form.

6.1.1.6. Interval Identification

**Participants:** David Daney, Julien Alexandre dit Sandretto, Gilles Trombettoni.

There are many approaches to identify the parameters of a model. In most cases, it consists in providing a particular solution of an over-constraint set of equations which must be robust to measurement to errors: in least square sense, with some statistical properties.... However, the interpretation and the validity of the result can be difficult and prone error. We propose to investigate some interval approaches in order to associate to the result some information and a certification of solutions.

6.1.2. Rehabilitation and biomechanics

**Participants:** Sami Bennour, David Daney, Mandar Harshe, Jean-Pierre Merlet [correspondant].

The focus of the work is on analyzing knee joint motion during a walking activity. The measurement system is based on the wire actuated parallel robot architecture. To increase the reliability of our analysis, and decrease the influence of Skin Tissue Artifacts (STA), we also incorporate a passive wire measurement system, IR camera based motion capture system, accelerometers, and force sensors to measure human motions.

The main principle of the system is to observe relative motions of the collars attached to tibia and femur. These are connected to the base by wires and also hold the other sensors. Measurements in the global frame and collar specific local frames give precise data to reconstruct collar (and thus, knee joint) motion.

Over the past year we have finalized the experimental setup, by calibrating the collars and the sensor systems, and adapting the existing wire robot system (MARIONET-REHAB) to work along with the other sensors. The software developed uses a single unified input file to specify all sensor configurations, streamlining experiments. We performed our preliminary experimental trials for walking motion on three subjects using the wire sensors, accelerometers and optical motion capture system.

We began work on processing the data obtained from these trials. Post-processing functions have also been developed to calculate additional collar properties, perform sensor data processing (filtering, noise removal and estimation) and access files in the C3D file format, which is used a binary file format used by the motion capture system.

The main challenge we are working on is to perform sensor data fusion and increase reliability of results. For this we must identify parameters that correlate the different sensor measurements and perform error analysis. Possible solutions include using interval analysis methods to address the uncertainties.

6.1.3. Kinematics of wire parallel robots

**Participant:** Jean-Pierre Merlet.

The kinematics of wire robot is a complex problem because a solution is possible only if the tension in the wire is positive. Hence the static equilibrium has to be taken into account. This problem is not well addressed in the literature. Curiously the forward kinematics of robot (i.e. finding the possible poses of the platform for given wire lengths) with at least 6 wires is straightforward: the distance equations allows to determine all poses and then we use the static equations to calculate the wire tensions and discard the one having at least one negative tension. For robot having less than 6 wires we have to consider simultaneously the distance equations and the static equations in order to get a square system (of \( n + 6 \) equations for a \( n \)-wires robot). We have investigated the case of a 3-wires robot with all wires attached at different points on the platform and have shown that all solutions can be computed provided the solving of an univariate polynomial of degree 158 [17]. Although we are not able to guarantee that the degree of this polynomial cannot be decreased, we believe that nevertheless the order will be too high for robust determination of the solutions and can only be used to determine an upper bound for the maximal number of solutions. We have also investigated theoretically and experimentally the
kinematics of a $n$-wires ($n \geq 4$) robot with all wires attached at the same point (i.e. only the position of the center of the platform can be controlled). Although this robot is apparently redundant, we have shown that in any pose at most 3 wires will be simultaneously under tension and therefore that the redundant wires cannot be used to control the wire tensions.

As the wire length measurements are not sufficient to determine the current pose of the platform (which is necessary for control purposes) we are investigating the use of additional sensors. Our prototypes MARIONET-ASSIST and MARIONET-VR are instrumented to measure wire directions, but with a large uncertainties. We have started a theoretical investigation to determine under which conditions these uncertainties may lead to a non-unique solution and we will validate the results on the two prototypes.

6.1.4. Rehabilitation robots for the immersive space

**Participants:** Michael Burman, Jean-Pierre Merlet.

The on-site immersive room provides 3D visualization but is lacking of haptic feedback and motion capabilities. We plan to implement in this room a movable system, constituted of:

- a 6 degrees-of-freedom motion base: the motion system 710-6-500-220 by Servos Simulation Inc. has been selected and is now operational and fully calibrated. If necessary the user may stand on this motion base
- the MARIONET-VR wire-driven parallel robot: this robot uses the same actuation principle than the MARIONET-REHAB robot (linear actuator with a pulley system for coiling and uncoiling of the wires), but is able to lift a person. The prototype is basically functionnal but its installation in the immersive room has been delayed because of lack of appropriate fixing elements

The full system will be installed in the immersive room at the beginning of 2012.

6.1.5. Assistance robotics

**Participants:** Michael Burman, David Daney, Jean-Pierre Merlet.

As mentioned earlier in the report we have started in 2008 a long term strategic move toward assistance robotics, with the objectives of providing low-cost, simple to control, robotized communicant devices that may help disabled, elderly and handicapped people in their personal life, with the credo that they have to be adapted to the end-user and to its everyday environment (by contrast with the existing trend of focusing on a "universal" robot, to which the end-user and its environment have to adapt) [18], [14], [21]. We have started last year the development of a simulated flat in order to explore various full scale scenarii that cover a part of the daily life of an elderly, to develop specific assistance devices and to test them. We describe in the following sections several devices that have been developed/improved during this year¹. Note that our demonstration in assistance are highlighted during the visit of Sophia (275 visitors have attended our demonstration during 14 visits) and have received serious press coverage (5 papers, 2 TV interviews).

6.1.5.1. Walking aids

Wheeled walking aids are usually the first tools that are used when motricity problems occur. We are developing the family of robotized Assistive Navigation Guide (ANG), which are based on commercially available Rollators, with several objectives (we mention only a few of them):

- fall prevention/detection: fall is a major problem for elderly (it is estimated that fall is the main cause of 10 000 elderly deaths per year in France).
- mobility help: provide an on-demand mobility help
- gait pattern monitoring: we believe that being able to monitor the trajectory of the walking aid will provide useful information on the gait pattern of the user

¹ pictures of this assistive flat are available at [http://www-sop.inria.fr/coprin/developpements/main.html](http://www-sop.inria.fr/coprin/developpements/main.html)
For reaching these objectives we have developed two walking aids:

- **ANG-light**: a walking aid with encoders in the wheels, 3D accelerometer, gyrometer. These sensors allow to measure the trajectory of the walking aid and several features of the user’s gait (step pattern, gait asymmetry,...). ANG-light has been tested by the CHU of Nice-Cimiez that was willing to perform an in-depth investigation of its use. For that purpose we have asked in September 2009 for the necessary formal authorization to the local CPP, which has been granted only in December 2011. To prepare this study we have organized a large scale experiment at INRIA, where 24 users were asked to perform the trajectories of the protocol twice, with and without the aid. When not using the aid the users were equipped with 3D accelerometer on the wrists and knees and were using specific shoes with force sensors in the sole. Initial analysis of the records shows that indeed we are able to obtain significant information on the gait pattern, that are not available using the existing tools, and detect differences in the gait pattern for user having even a light pathology in the lower limb. The experiment with elderly patients at CHU will take place in January 2012.

- **ANG-II**: this aid is an evolution of the motorized walker ANG, with a lower weight and better integration.

### 6.1.5.2. MARIONET-ASSIST

This wire-driven parallel robot is installed in the ceiling of the flat. It has been used this year in the 4-1 configuration (4 wires attached at the same point), which allows for controlling the position of the platform, but not its orientation. Several platforms have been developed, all of them incorporating a webcam and allowing for a free rotation around the vertical axis, while an accelerometer measure the tilt angle of the platform (which is used to determine in which direction the end-user is willing to move). One of the platform incorporate a 4 d.o.f. robot that may grasp light object (one of our objective is to use also the robot as a manipulator for bringing object back to the user in a more or less autonomous way, which is the subject of the PhD thesis of R. Ramadour).

We have shown that the the robot can be used for sit-to-stand transfer and for lifting handicapped people. A specific attention has been devoted to propose very simple control interface: joystick, remote TV set, control box whose tilt determine the motion axis.

### 6.1.5.3. Other flat equipments

Our scenario includes the management of emergency situations such as the fall. Fall detection can be performed by the ANG walkers but we have also started investigating the inclusion of fall detection system in the clothes of the end-user either through a GEO-300 devices or by incorporating an Arduino Lilypad processor. When a fall is detected indoor an alert is transmitted to a coordinator (a Nabaztag) which will order the walker and the MARIONET-ASSIST robot to move close to the user to provide a support. At the same time two mobile robots will converge to the same location: a remote-controlled, webcam equipped ROVIO (which can provide images of the end-user to a rescue center) and a Pekee II, that we will equip to provide first aid.

An important point in assistance is to be able to have at all time a rough idea of the localization of the patient. Although we plan to use a Kinect for that purpose, we will also investigate the use of non-vision sensors (which are much less intrusive and therefore can be more easily accepted) such as RFID tag (ANG-II has a RFID tag reader), directive distance sensors and light barriers.

Another axis for assistance is to reduce the risks of fall by using the principle that the objects has to come to the hand of the user (or of the robot), not the opposite. This implies instrumenting the environment with drawer openers and doors manipulation and we have started implementing them on drawers and on the fridge of the flat.

### 6.2. Interval analysis

#### 6.2.1. Inner Regions and Interval Linearizations for Global Optimization

**Participants:** Gilles Trombettoni [correspondant], Bertrand Neveu.
Researchers from interval analysis and constraint (logic) programming communities have studied intervals for their ability to manage infinite solution sets of numerical constraint systems. In particular, inner regions represent subsets of the search space in which all points are solutions. Our main contribution is the use of recent and new inner region extraction algorithms in the upper bounding phase of constrained global optimization.

Convexification is a major key for efficiently lower bounding the objective function. We have adapted the convex interval taylorization proposed by Lin & Stadtherr for producing a reliable outer and inner polyhedral approximation of the solution set and a linearization of the objective function. Other original ingredients are part of our optimizer, including an efficient interval constraint propagation algorithm exploiting monotonicity of functions.

We end up with a new framework for reliable continuous constrained global optimization. This interval Branch & Bound significantly outperforms the best reliable global optimizers [22], [25], [28].

6.2.2. An Interval Extension Based on Occurrence Grouping

Participants: Bertrand Neveu [correspondant], Gilles Trombettoni.

We proposed last year a new “occurrence grouping” interval extension \([f]_og\) of a function \(f\). When \(f\) is not monotonic w.r.t. a variable \(x\) in a given domain, we try to transform \(f\) into a new function \(f^{og}\) which is monotonic w.r.t. two subsets \(x_a\) and \(x_b\) of the occurrences of \(x\): \(f^{og}\) is increasing w.r.t. \(x_a\) and decreasing w.r.t. \(x_b\). \([f]_og\) is the interval extension by monotonicity of \(f^{og}\) and produces a sharper interval image than the natural extension does.

This year we have improved the linear program and algorithm that minimize a Taylor-based over-estimate of the image diameter of \([f]_og\). We have detailed the proofs of correctness and reliability of this occurrence grouping algorithm [8], [29].

6.3. Miscellaneous results

6.3.1. Equilibrium strategies for linked Electricity and CO2 markets

Participant: Odile Pourtallier.

In collaboration with M. Bossy (INRIA -TOSCA Team) and N. Maïzi (CMA - Mines ParisTech) O. Pourtallier the study of equilibrium model for coupled electricity and CO2 allowance exchange markets has been pursued. (see also Section 7.1 ). A static equilibrium model has been studied under various assumptions on the CO2 market design. All the CO2 market designs do not lead equilibrium, which interferes on the (short term day ahead) electricity market, which in turn interferes on the electricity mixe and consequently on the total emission. Together with El-Hadj Dia (INRIA -TOSCA Team) we have also pursued an indifference pricing methodology which is presented in more details in INRIA -TOSCA Team section.

6.3.2. Symbolic tools for modeling and simulation

Participant: Yves Papegay.

This activity is the main part of a long-term ongoing collaboration with Airbus whose goal is to directly translate the conceptual work of aeronautics engineers into digital simulators to accelerate aircraft design.

An extensive modeling and simulation platform has been designed which includes a dedicated modeling language for the description of aircraft dynamics models in term of formulae and algorithms, and a symbolic compiler producing as target an efficient numerical simulation code ready to be plugged into a flight simulator, as well as a formatted documentation compliant with industrial requirements of corporate memory.
Implementation of this platform is a modeling and simulation environment based on symbolic computation tools. It contains several components:

- a model editor, that makes it possible and easy to enter the whole set of equations describing large and complex industrial models,
- an highly interactive and modular evaluation workbench allowing to simulate the models and to visualize the results inside the modeling environment with the benefits for the designer of being able to directly use all its computational functionalities,
- a C code generator which, using these models, automatically generates the numerical real-time simulation engines
- a technical documentation generator

During the year 2011 the technology demonstrated by our prototype has been transferred to our industrial partner. A lot of work has been done on our modeling and simulation environment to improve its robustness and its development level of quality toward industrial standards. Final version of our prototype is to be delivered to Airbus at the end of the year.

### 6.3.3. Multi-agent aircraft design

**Participant:** Yves Papegay.

The modeling environment described in the previous section is used, in collaboration with other teams at Airbus, in the framework of the ID4CS project founded by ANR and dedicated to multi-agent optimization of large scale system. Several models of aircraft engines and of aircrafts have been developed as user cases for the project. Automatic generation of extended models namely computing first order derivatives of the original models has been implemented.
5. New Results

5.1. Specification and verification of database driven systems

Participants: Serge Abiteboul, Pierre Bourhis, Luc Segoufin, Szymon Toruńczyk, Victor Vianu.

Modelization and verification of data centric systems. We have intensively studied the Active XML model. It is a high-level specification language tailored to data-intensive, distributed, dynamic Web services. Active XML is based on XML documents with embedded function calls. The state of a document evolves depending on the result of internal function calls (local computations) or external ones (interactions with users or other services). Function calls return documents that may be active, so may activate new subtasks. Our first line of result is a comparison of the specification power of various workflow control mechanisms within the Active XML framework and beyond [23].

AXML is very powerful and many static analysis problems are undecidable. We have also introduced a model of automata designed for modeling infinite runs of systems equipped with static relational databases. The automata model is equipped with finitely many variables, each of which can store values from a linearly ordered domain, such as the natural numbers. The transitions of the automata depend on a conjunctive query involving the database and the current values of the variables. For verifying infinite runs of such automata, an extension of temporal logic is considered, capable of comparing values stored in the variables and the database, at different times of the run. The main contribution of the work is the proof that automated verification of such temporal properties of the system can be carried out in PSpace. For more details, see [35].

Static analysis of query languages. XPath is arguably the most widely used XML query language as it is implemented in XSLT and XQuery and it is used as a constituent part of several specification and update languages. Hence in order to perform static analysis on a system manipulating XML data it is important to master the static analysis for XPath. Most of the important static analysis problems reduce to satisfiability checking: does a given query return a non-empty answer on some data. In general, in the presence of data values, the satisfiability of XPath is undecidable. We have shown that when restricted to its vertical navigational power, XPath becomes decidable [30].

5.2. Distributed data management

Participants: Serge Abiteboul, Emilien Antoine, Daniel Deutch, Alban Galland, Wojciech Kazana, Yannis Katis, Luc Segoufin, Cristina Sirangelo.

Distributed knowledge base. As a foundation for managing distribution, we have proposed a model of a distributed knowledge base, that handles data and meta-data, as well as access control and localization, in a unique integrated setting. To support automatic reasoning on this knowledge base, we also introduced a novel rule-based language supporting the exchange of rules, namely Webdamlog. This work has been presented [21] and demonstrated [26] at major database conferences.

Probabilistic XML. Data from the Web are imprecise and uncertain. To manage this imprecision in a well-principled way, we have made significant advances in the field of probabilistic databases, and specifically, probabilistic XML. We have introduced new tractable probabilistic models for representing uncertain hierarchical information, and carried out in-depth studies of query evaluation, aggregation, and updates in various probabilistic XML models. These results have matured and some of the results are available in journal articles, e.g., [14].
Enumeration of query answers. In many applications the output of a query may have a huge size and enumerating all the answers may already consume too many of the allowed resources. In this case it may be appropriate to first output a small subset of the answers and then, on demand, output a subsequent small numbers of answers and so on until all possible answers have been exhausted. To make this even more attractive it is preferable to be able to minimize the time necessary to output the first answers and, from a given set of answers, also minimize the time necessary to output the next set of answers - this second time interval is known as the delay. We have shown that this was doable with a linear preprocessing time and constant enumeration delay for first-order queries over structures of bounded degree [19].

Data exchange and Web incomplete information. We have addressed the problem of restructuring data exchanged between communicating applications on the Web. We have proposed and analyzed a new language to specify data restructuring rules (schema mappings). This language generalizes existing mapping dependencies, by allowing a more flexible specification mechanism [20].

Jorge. We also invested a lot of effort in a textbook (undergraduate and graduate level) on Web data management (nicknamed Jorge) to be published at Cambridge University Press [38]. The book is already available on the Webdam Web site http://webdam.inria.fr/Jorge

5.3. Tree automata theory

Participants: Stéphane Demri, Florent Jacquemard, Luc Segoufin.

Most of our results for this section concerns data words and data trees. Those are words and trees where each position contains a data value together with the classical label. Data trees can be seen as a model for XML data. We have studied automata model using registers or memory or allowing constraints that can involve data comparisons in its transitions.

Register Automata. These extend the classical model of finite automata with auxiliary registers storing data values for later comparison.

We have introduced a new model of automata over data trees and shown the decidability of its emptiness problem [30]. These automata were used for obtaining decidability results for the static analysis for some fragments of XPath presented in the previous section.

Automata with counters. In [39], a survey chapter on the verification of infinite-state systems is presented that is focused on the verification of counter systems. Verification problems for vector addition systems and recursive Petri nets are considered. Moreover, we introduce subclasses of counter systems for which reachability questions can be solved in Presburger arithmetic viewed as a means to symbolically represent sets of tuples of natural numbers.

Automata with isomorphism tests among subtrees. We have also considered some models described by tree automata enriched with a feature testing for isomorphisms between subtrees. Such constraints could be used for testing monadic key constraints over XML documents. For these models, the main challenge is to establish the decidability of the non-emptiness of the language specified by a given automaton [18].

Rewriting Controlled by Selection Automata. Motivated by the problem of static analysis of XML update programs, we have studied [33] the combination, called controlled term rewriting systems (CTRS), of term rewriting rules with constraints selecting the possible rewrite positions. These constraints are specified, for each rewrite rule, by a selection automaton which defines a set of positions in a term based on tree automata computations. We have established several decidability and complexity results for several cases of the reachability and regular model checking problems for this tree transformation formalism.
DREAM Project-Team

6. New Results

6.1. Diagnosis of large scale discrete event systems

**Participants:** Marie-Odile Cordier, Christine Largouët, Sophie Robin, Laurence Rozé, Yulong Zhao.

The problem we deal with is monitoring complex and large discrete-event systems (DES) such as an orchestration of web services or a fleet of mobile phones. Two approaches have been studied. The first one consists in representing the system model as a discrete-event system by an automaton. In this case, the diagnostic task consists in determining the trajectories (a sequence of states and events) compatible with the sequence of observations. From these trajectories, it is then easy to determine (identify and localize) the possible faults. In the second approach, the model consists in a set of predefined characteristic patterns. We use temporal patterns, called chronicles, represented by a set of temporally constrained events. The diagnostic task consists in recognizing these patterns by analyzing the flow of observed events.

More recently, we started research on interacting with large-scale systems in a decision-oriented way. Scenario patterns were defined for exploring complex systems, based on the use of model-checking techniques.

6.1.1. Distributed monitoring with chronicles - Interleaving diagnosis and repair - Making web services more adaptive

Our work addresses the problem of maintaining the quality of service (QoS) of an orchestration of Web services (WS), which can be affected by exogenous events (i.e., faults). The main challenge in dealing with this problem is that typically the service where a failure is detected is not the one where a fault has occurred: faults have cascade effects on the whole orchestration of services. We have proposed a novel methodology to treat the problem that is not based on Web service (re)composition, but on an adaptive re-execution of the original orchestration. The re-execution process is driven by an orchestrator Manager that takes advantage of an abstract representation of the whole orchestration and may call a diagnostic module to localize the source of the detected failure. It is in charge of deciding the service activities whose results can be reused and may be skipped, and those that must be re-executed. A paper has been submitted to the CAISE conference.

6.1.2. Scenario patterns for exploring qualitative ecosystems

Our work aims at giving means of exploring complex systems, in our case ecosystems. We proposed to transform environmental questions about future evolution of ecosystems into formalized queries that can be submitted to a simulation model. The system behavior is represented as a discrete event system described by a set of interacting timed automata, the global model corresponding to their composition on shared events. To query the model, we have defined high-level generic patterns associated to the most usual types of scenarios. These patterns are then translated into temporal logic formula. The answer is computed thanks to model-checking techniques that are efficient for analyzing large-scale systems. Five generic patterns have been defined using TCTL (Timed Computation Tree Logic): WhichStates, WhichDate, WhichStates, Stability, Safety. Three of them have been implemented using the model-checker UPPAAL.

The approach has been experimented on a marine ecosystem under fishing pressure. The model describes the trophodynamic interactions between fish trophic groups as well as interactions with the fishery activities and with an environmental context. A paper has been accepted for publication by the Environmental Modelling Software Journal [4].

We extended the approach to deal with “How to” queries. As before, we rely on a qualitative model in the form of timed automata and use model-checking tools to answer queries. We have recently proposed two approaches to answer questions such as “How to avoid a given situation?” (safety query). The first one exploits controller synthesis and the second one is a “generate and test” approach. We compared these two approaches in the context of an application that motivates this work, i.e the management of a marine ecosystem and the evaluation of fishery management policies. The results have been accepted for publication in [14].
6.2. Machine learning for model acquisition

Participants: Thomas Guyet, René Quiniou.

Model acquisition is an important issue for model-based diagnosis, especially as modeling dynamic systems. We investigate machine learning methods for temporal data recorded by sensors or spatial data resulting from simulation processes. We also investigate efficient methods for storing and accessing large volume of simulations data. Our main interest is extracting knowledge, especially sequential and temporal patterns or prediction rules, from static or dynamic data (data streams). We are particularly interested in mining temporal patterns with numerical information and in incremental mining from sequences recorded by sensors.

6.2.1. Mining temporal patterns with numerical information

We are interested in mining interval-based temporal patterns from event sequences where each event is associated with a type and time interval. Temporal patterns are sets of constrained interval-based events. This year, we have been working on improving the formal setting of the approach as well as its efficiency [8]. We have introduced the notion of $\epsilon$-covering of temporal patterns over sequences to cope with the dual nature, symbolic and numerical, of temporal patterns. The parameter $\epsilon$ specifies the tightness of the similarity used for matching patterns and sequences. It complements the parameter $\sigma$ representing the minimal support which is used to prune candidate patterns. The $\epsilon$-similar occurrences of some pattern, precisely their associated temporal intervals, are classified to characterize the different classes of numerical temporal intervals that correspond to different patterns sharing the same symbolic part. This process have been embedded in two sequential pattern mining algorithms, GSP and PrefixSpan, and we have compared their performance.

6.2.2. Incremental sequential mining

We investigate the problem of mining and maintaining frequent sequences in a window sliding on a stream of itemsets. We propose in [11] a complete and correct incremental algorithm based on a tree representation of frequent sequences inspired by PSP [52] and a method for counting the minimal occurrences of a sequence. Instead of the frequence, to a node representing a pattern is associated the set of occurrences of this pattern. The algorithm updates efficiently the tree representation of frequent sequences and their occurrences by means of two operations on the tree: deletion of the itemset at the beginning of the window (obsolete data) and addition of an itemset at the end of the window (new data). Experiments were conducted on simulated data and on real data of instantaneous power consumption.

6.2.3. Multiscale segmentation of satellite image time series

Satellite images allow the acquisition of large-scale ground vegetation. Images are available along several years with a high acquisition frequency (1 image every two weeks). Such data are called satellite image time series (SITS). In [9], we present a method to segment an image through the characterization of the evolution of a vegetation index (NDVI) on two scales: annual and multi-year. We test this method to segment Senegal SITS and compare our method to a direct classification of time series. The results show that our method using two time scales better differentiates regions in the median zone of Senegal and locates fine interesting areas (cities, forests, agricultural areas).

6.3. Decision aiding with models and simulation data

Participants: Tassadit Bouadi, Marie-Odile Cordier, Véronique Masson, Florimond Ployette, René Quiniou, Karima Sedki.

Models can be very useful for decision aiding as they can be used to play different plausible scenarios for generating the data representing future states of the modeled process. However, the volume of simulation data may be very huge. Thus, efficient tools must be investigated in order to store the simulation data, to focus on relevant parts of the data and to extract interesting knowledge from these data.
6.3.1. Exploring models thanks to scenarios: a generic framework

In the framework of the APPEAU project (see 8.2.1), that ended in December 2010, a paper, describing a generic framework for scenario exercises using models applied to water-resource management, has been written during 2011 in cooperation with all the partners and submitted to Environmental Modelling and Software. It is currently under revision.

6.3.2. A datawarehouse for simulation data

The ACASSYA project aims at providing experts or stakeholders or farmers with a tool to evaluate the impact of agricultural practices on water quality. As the simulations of the deep model TNT2 are time-consuming and generate huge data, we have proposed to store these simulation results in a datawarehouse and to extract relevant information, such as prediction rules, from the stored data. We have devised a general architecture for agro-environmental data on top of the framework Pentaho. An article presenting the principles of this architecture as well as a set of realistic scenarios and their transformation into OLAP queries has been submitted to Compag (Computers and Electronics in Agriculture).

6.3.3. Efficient computation of skyline queries in an interactive context

Skyline queries retrieve from a database the objects that optimizes multiple criteria, related to user preferences for example, or objects that are the best compromises satisfying these criteria. When data are huge such objects may shed light on interesting parts of the dataset. However, computing the skylines (i.e. retrieving the skyline points) may be time consuming because of many dominance tests. This is, especially the case in an interactive setting such as querying a data cube in the context of a datawarehouse. We have worked on how to answer efficiently to skyline queries by the materialization of precomputed skyline queries related to dynamic user preferences. An article has been submitted to the conference SIGMOD 2012.

6.3.4. Influence Diagrams for Multi-Criteria Decision-Making

For multi-criteria decision-making problems, we propose in [6] a model based on influence diagrams able to handle uncertainty, represent interdependencies among the different decision variables and facilitate communication between the decision-maker and the analyst. The model makes it possible to take into account the alternatives described by an attribute set, the decision-maker’s characteristics and preferences, and other information (e.g., internal or external factors) that influence the decision. Modeling the decision problem in terms of influence diagrams requires a lot of work to gather expert knowledge. However, once the model is built, it can be easily and efficiently used for different instances of the decision problem. In fact, using our model simply requires entering some basic information, such as the values of internal or external factors and the decision-maker’s characteristics.

6.3.5. Recommending actions from classification rules

In the framework of the SACADEAU project (see 8.2.1), a paper dedicated to building recommendation actions for a given situation, from the set of classification rules, learnt from simulation results, has been published in the KAIS journal [7].

6.4. Causal reasoning and influence diagrams

Participants: Philippe Besnard, Louis Bonneau de Beaufort, Marie-Odile Cordier, Yves Moinard, Karima Sedki.

This work stems on [23], [24], [25], [26], [27] and, for the logic programming translation, on [53], [54]. It is related to diagnosis (observed symptoms explained by faults).

The previously existing proposals were ad-hoc or, as in [29], [41], they were too close to standard logic in order to make a satisfactory diagnosis. Our proposal starts from a restricted first order logic (of the Datalog kind: no function symbols) and introduces causal formulas, built on causal atoms such as (causes) intended to mean: “α causes β”. The system is described thanks to these causal formulas, classical formulas, and taxonomy atoms such as (IS_A) (α is of kind β).
The system produces *explanation atoms* of the kind \( (\alpha \text{ explains } \beta \text{ if possible } \{\gamma_1, \ldots, \gamma_n\}) \), meaning that \( \beta \) can be explained by \( \alpha \) if all the \( \gamma_i \)'s are possible together in the context of the given data.

This year, we have improved our logic programming translation in ASP. The aim is to improve efficiency and also reduce the work of the programmer, taking advantage of the declarative aspect of this type of programming. We have applied some of these improvements to two classic riddles, in order to illustrate the power and limitations of current answer set programming systems, and we proposed a few improvements which could make the present systems yet easier to use [12], [13].

We are starting a work with some similarities to automatize the treatment of cognitive maps. The aim is to extract relevant information from these maps, which means: building a graph formalism for representing mixed causal and influence relations, and defining a framework (argumentation theory is a good candidate) to aggregate the graphs and provide inference rules in order to infer new information and relations. This work is done in the framework of the RADE2BREST project, involving Agrocampus Ouest and CNRS (GEOMER/LETG), funded by "Ministère de l’Ecologie". The goal of this project is to model shellfish fishing in order to assess the impact of management pollution scenarios on the Rade de Brest. The cognitive maps result from interviews with fishermen.

---

1. This project is not mentioned in section 8.1 because DREAM is not a partner of this project.
5. New Results

5.1. Dynamic World Perception and Evolution Prediction

5.1.1. Environment modeling and sensor data acquisition


An overall architecture of our environment-modeling module with the inputs from heterogeneous sensors is shown in Fig. 6. The combined use of two lidars and stereo-vision helps mitigate uncertainty and allows for detection of partially occluded objects. The data processing includes the computation of probabilistic occupancy grids for each sensor and their subsequent fusion with the Bayesian Occupancy Filter (BOF). The output of the module is an estimation of the position, velocity and associated uncertainty of each observed object, which are used as input to the risk assessment module.

![Figure 6. Architecture of the environment modeling module.](image)

This architecture is implemented on our experimental platform, a Lexus LS600h car shown in Fig. 7. The vehicle is equipped with a variety of sensors including two IBEO Lux lidars placed toward the edges of the front bumper, a TYZX stereo camera situated behind the windshield, and an Xsens MTi-G inertial sensor with GPS.

The stereo camera baseline is 22 cm, with a field of view of 62°. Camera resolution is 512x320 pixels with a focal length of 410 pixels. Each lidar provides four layers of up to 200 impacts with a sampling period of 20 ms. The angular range is 100°, and the angular resolution is 0.5°. The on-board computer is equipped with 8GB of RAM, an Intel Xeon 3.4 GHz processor and an NVIDIA GeForce GTX 480 for GPU. The observed region is 40 m long by 40 m wide, with a maximum height of 2 m. Cell size of the occupancy grids is 0.2x0.2 m.
The Lexus experimental platform provides to acquire sensor data in real traffic environments: eight layers of laser scans, stereo images, IMU data (accelerations), velocity, GPS position, steering angle. The experiments are conducted in various road environments (country roads, downtown and highway), at different time of the day, with various driving situations (light traffic, dense traffic, traffic jams). The datasets are acquired online and are used for testing of our sensor fusion and risk assessment algorithms.

5.1.2. Bayesian fusion of visual and telemetric information

Participants: Igor Paromtchik, Christian Laugier, Mathias Perrollaz, Amaury Nègre.

5.1.2.1. Concept of BOF and obstacle detection in occupancy grids

Obstacle detection is a widely explored domain of mobile robotics. It presents a particular interest for the intelligent vehicle community, as it is an essential building block for Advanced Driver Assistance Systems (ADAS). In the ANR project LOVe (Logiciel d’Observation de Vulnérables) and ArosDyn project, the e-Motion team proposed to perform obstacle detection within the occupancy grid framework. In order to work efficiently with occupancy grids, we have previously developed a probabilistic framework with the Bayesian Occupancy Filter (BOF) [40] [88] (patent 0552736 (2005)), which provides filtering, data fusion, and velocity estimation capabilities while allowing for parallel computation. The Fast Clustering and Tracking Algorithm (FCTA) [73] is then used to identify and track individual objects. The BOF is designed with the intent of its implementation in hardware as a system-on-chip. Like other grid based approaches, the BOF framework performs sensor fusion at the cell level [40]. The BOF evaluates probabilities of both cell occupancy and cell velocity for each cell in a four-dimensional spatio-temporal grid. The monitoring of traffic scenes includes detection and tracking of objects by the FCTA [73].

Fig. 8 shows examples of occupancy grid mapping with the proposed approach. The arrows indicate the pedestrian, the car, and the bicycle, which appear in the camera images and the occupancy grids. Because the accuracy of stereo-vision tends to become poor at large distance, the corresponding grid has been attenuated beyond 20 m and the system is tuned to give more confidence to the lidars than to the stereo-vision. One of advantages of sensor fusion is a larger viewfield so that the vehicles overtaking the ego-vehicle (they are not seen in the camera images) are correctly mapped on the resulting BOF grid. Moreover, the sensor fusion as well as the Bayesian estimation provide to filter out the laser impacts with the road surface, e.g. right lidar in Fig. 8.

Note that a large number of dynamic objects in the traffic scenes may lead to a failure of object-based fusion because of a large number of association hypotheses. The grid-based approach allows us to avoid the object association problem for sensor fusion.

5.1.2.2. Disparity space approach for a vision based occupancy grid

Figure 8. Examples of occupancy grid mapping in typical urban traffic scenes, from left to right: left image from the stereo pair, an occupancy grid from the left lidar, an occupancy grid from the right lidar, an occupancy grid from stereo-vision, an occupancy grid estimated by data fusion with the BOF, and a probability scale.

To use sensors in the BOF framework, it is essential to develop an associated probabilistic sensor model that takes into consideration the uncertainty over measurements. In 2009, we proposed such a sensor model for stereo-vision [79]. The originality of the approach relied on the decision to work in the disparity space, instead of the classical Cartesian space. In 2010, we improved our sensor model, in order to mimic some features of the sensor models used for range finders. Particularly, we worked on managing visible/occluded areas of the scene [81], and on including the information from the road/obstacle segmentation of the disparity image [80]. Our approach was also designed to allow highly parallel computation of the occupancy grid. A. Nègre implemented the approach on GPU using NVIDIA CUDA to enhance the performance. The complete processing of the stereo data can now be done in 6 ms, while more than 150 ms were necessary with the CPU implementation. The complete approach for occupancy grid computation using stereovision is described in [30].

Figure 9 shows an example of the occupancy grid computed by our new approach. We can observe that most objects are detected (light color), even if partially occluded (e.g. the sign on the right). Information from the road surface is also taken into consideration (dark areas). Moreover, similar to a laser scanner, it appears that regions in front of objects are seen as partially unoccupied, while less information is available behind obstacles (occupancy probability is closer to 0.5).

In 2011, we focused on including the approach into the risk estimation framework on our Lexus experimental platform. We implemented a demonstration to estimate a distance measurement to the closer object situated in the future trajectory of the vehicle. The future trajectory is estimated either by using a lane detection algorithm (in the highway) or by combining velocity and steering information of the vehicle. Figure 10 shows the HMI displayed in the car while driving.

5.1.2.3. Processing of multi-layer telemetric data in probabilistic framework

Participants: Mathias Perrollaz, Juan-David Adarve, Alexandros Makris.
The occupancy grid computation based on a laser scanner uses the classical independent beam sensor model [90]. Since our vehicle is equipped with two four-layers laser scanners, it is necessary to merge the data from the multiple layers. In the original BOF framework, the fusion was performed through the classical Bayesian Fusion methodology. As shown in figure 11, this method causes problems of misdetection when some beams go over an object. In 2011, we proposed and implemented another approach. The fusion is now obtained through a weighted sum of the occupancy grids provided by each layer. The weight of each layer is obtained by computing a confidence grid. This confidence depends both on the inclination of the layer and on the possible occlusions. The new approach provides a more precise description of the environment.

5.1.3. Sensor Fusion and parameters estimation

Participants: Agostino Martinelli, Chiara Troiani.

This is the follow up of the research activity started in 2009, when a self-calibration problem for a wheeled robot has been investigated. The main results achieved during that year were published in [69], [71] and [70].

This calibration problem allows us to introduce a general framework able to deal with any estimation problem. This framework is based on a new theoretical concept, the concept of continuous symmetry. Detecting the continuous symmetries of a given system has a very practical importance. It allows us to detect an observable state whose components are non linear functions of the original non observable state. The general theory has been developed during the last two years. Preliminary results have been published in 2010 [72] and a more complete version of these results, which include several extensions, has been published on Transaction on Robotics, in 2011 [9].

In 2011, this general framework has been extensively applied to investigate the problem of the fusion of visual and inertial data in the framework of the European project sFly. Special emphasis has been devoted to the structure from motion problem (SfM) when fusing these data. This problem has particular interest and has been investigated by many disciplines, both in the framework of computer science ([35], [54], [56], [87] and references therein) and in the framework of neuroscience and vision perception ([67], [95] and references therein). Even though prior work has answered the question of which are the observable...
Figure 10. a) segmentation of the environment with the stereo-vision algorithm. Blue areas belong to the road surface, while red areas belong to the obstacles. b) HMI shown in the car during the demonstration of risk estimation. The trajectory is estimated by considering the velocity and steering angle of the ego vehicle. Here the car in front is not considered as dangerous because it is more than 2 seconds ahead. c-d) Another example, on the highway. For this example, the trajectory is estimated by considering the road markings.
Figure 11. Occupancy grid computed after fusion of eight layers of laser data. Above: with the previous approach, some objects are not correctly represented (e.g. the barrier on the left). Below: with the new approach, the description is more precise.
modes, i.e. the states that can be determined by fusing visual and inertial measurements [35], [54], [56], the questions of how to compute these states in the absence of a prior, and of how many solutions are possible, were still unanswered. During 2011, we have derived, for the first time, a closed form solution to the SfM problem in this case, allowing the determination of the observable modes without the need for any prior knowledge. The proposed solution analytically expresses all the observable modes in terms of the visual and inertial measurements acquired during a given (short) time-interval allowing the determination of all the observable modes without the need for any prior knowledge. Additionally, we have shown that this problem can have a unique solution or two distinct solutions or infinite solutions depending on the trajectory, on the number of point-features and on the number of monocular images where the same point-features are seen. Our results are relevant in all the applications which need to solve the structure from motion problem with low-cost sensors and which do not demand any infrastructure. Typical examples are the emergent fields of space robotics [77], humanoid robotics and unmanned aerial navigation in urban-like environments [93], where the use of the GPS is often forbidden. Furthermore, our results could play an important role in neuroscience by providing a new insight on the process of vestibular and visual integration. To this regard, we remind the reader that the influence of extra retinal cues in depth perception has extensively been investigated in the last decades. In the case when this extra retinal cue is the motion parallax induced by self-motion relative to a stationary environment, the scale factor is provided by the head velocity [65], [66]. The vast majority of these studies, consider the case when the head motion is active [38], [94]. This prevents the possibility to understand the contribution of the vestibular signals because of efference copy generated by active self movement. However, a very recent study investigates this problem by performing trials with passive head movements [43]. The conclusion of this study is that the combination of retinal image with vestibular signals can provide rudimentary ability to depth perception. Our findings could provide a new insight to this problem of depth perception since by combining retinal image with vestibular signals it is possible to determine the scale factor even without any knowledge about the initial speed. New trials would be necessary in order to verify whether a mechanism reproducing our closed form solution is present in humans and/or in other animals (especially the ones without binocular vision). Our findings also show that it is possible to easily distinguish linear acceleration from gravity. Specifically, our closed form solution perform this determination by a very simple matrix inversion. This problem has also been considered in neuroscience [75], [31]. Our results could provide a new insight to this problem since they clearly characterize the conditions (type of motion, features layout) under which this determination can be performed.

Our results have been published in three conference papers [14], [11], [15] and have been accepted for publication in transactions on robotics (a version is currently available as a technical report, [29]).

In parallel to this theoretical activity an experimental activity has started in order to experimentally validate our findings in the near future and to deploy our technologies to industrial partners. To this regard, a contact with the company Delta Drone in Grenoble has been established and a valorization contract with a SME in the field of civil drone applications is currently in preparation.

5.1.4. Analysis of dynamic scenes for collision risk assessment


The grid-based environment representation is used for dynamic scene analysis in the Arosdyn project [78]. The original idea behind the risk estimation approach developed in the e-Motion team consists in considering the possible behaviors of the vehicles in the scene. Indeed, with the classical TTC(time to collision)-based approach, the risk is estimated based on the prediction of the trajectory, considering the current state of the objects. This is only valid for very short term predictions, and in some cases it can result in a over-estimation of the collision risk. Understanding the intention of the other participants of the road scene allows a longer term, more precise prediction of trajectories.

Our approach is divided into two steps: behavior recognition and behavior realization. The behavior recognition aims at estimating the probability for a vehicle to perform one of its feasible behaviors. The behaviors are semantic representations of driving maneuvers (e.g. turn left, turn right, go straight, ...). The probability
distribution over possible behaviors is obtained by inference using layered HMMs. Driving behavior realization is modeled as Gaussian Process (GP). This model allows us to obtain the probability distribution over the physical realization of the vehicle motion (i.e. trajectories) by assuming a usual driving, for a given behavior.

Finally, a complete probabilistic model of the possible future motion of the vehicle is given by the probability distribution over driving behaviors, and by the realization of these behaviors. The risk calculation is performed by sampling of the paths from the corresponding GP. The fraction of the samples in collision gives the risk of collision.

In 2011, we conducted some early experiments on sensor fusion, using real data acquired with our Lexus experimental vehicle [16]. Moreover, the global framework of the Arosdyn project has been presented in [8].

5.1.5. Recognition for intelligent vehicles


We developed a generic object class recognition method. The method uses local image features and follows the part based detection approach. The state-of-the-art visual object class recognition systems operate with local descriptors and codebook representation of the objects. Various local features (e.g. gradient maps, edges) are used to create the descriptors. Then kernel based classifiers are commonly employed to classify the detected features in one of several object classes [32] [45]. The recognition of vehicles or pedestrians from sensors mounted on a moving platform is achieved by different approaches using various types of sensors, e.g. stereo camera, laser [51] [44]. The approaches that perform data fusion from various sensors have proven to be the more robust in a variety of road conditions [86].

This work focuses on the development of an object class recognition system which follows the part based detection approach [64]. The system fuses intensity and depth information in a probabilistic framework. To train the system for a specific object class, a database of annotated with bounding boxes images of the class objects is required. Therefore, extending the system to recognize different object classes is straightforward. We apply our method to the problem of detecting vehicles by means of on-board sensors. Initially, depth information is used to find regions of interest. Additionally, the depth of each local feature is used to weight its contribution to the posterior of the object position in the corresponding scale. The votes are then accumulated in a 3d space-scale space and the possible detections are the local maxima in that space. Figure 12 presents the steps of our approach.

![Figure 12. Detection procedure steps. The stereo information is used to define the regions of interest for the subsequent steps. Intensity and depth features are extracted from a dense grid within these regions. In the following the features are matched with the codebook clusters which are in turn used to estimate the posterior for the object in each position. The detections are the local maxima of the posterior.](image-url)
The novelty of our approach is the fusion of depth and intensity information to form a probabilistic part-based detector. Using depth information is beneficial for the robustness of the approach, because we avoid including many noisy detections resulting from false matches between features of different scales. The method is tested with stereo video sequences captured in an urban environment. Fig. 13 shows some example detections. The proposed method detects side-views of cars in various scales, in cases with partial occlusions, and under significant background clutter.

![Car-side detection examples.](image)

**Figure 13. Car-side detection examples.** True and false positive detections are represented by red and yellow bounding boxes respectively. (a) Cars in different scales with significant background clutter and significant occlusions are detected. (b) Precise detection of the un-occluded vehicle, whereas a vehicle that is heavily occluded in the left is not detected. (c) Difficult detection of a vehicle which is far and partially occluded and a false detection in the region between the road surface and the trees. (d) Detection with partial occlusion. (e) Partial detection of a taller than normal vehicle (on the left). The training dataset does not contain vehicles of this type. (f) Successful detection of a partially occluded car and a false positive arising from a bus and a van. Training separate detectors for these type of vehicles as well will help to avoid these false alarms.

5.1.6. **Context-aware Bayesian estimation of risk at road intersections for cooperative vehicles**

**Participants:** Stéphanie Lefèvre, Christian Laugier.

The work developed in this PhD is done in collaboration with Renault (CIFRE thesis) and concerns safety applications for cooperative vehicles.

In a few years, car manufacturers will start equipping vehicles with V2X communication devices, which will allow vehicles to share information with other vehicles and with roadside units using a dedicated communication channel. This new sensor on the car opens a whole new world of possibilities for Advanced Driver Assistance Systems (ADAS). In particular, the fact that the vehicle is able to “see” a car before it even enters the field-of-view of the driver allows for a better assistance in the tasks of perceiving, analyzing, predicting, and estimating the risk of a situation.

Early in the PhD we identified safety applications at road intersections as a relevant application domain for V2X technologies. The variety and complexity of scenes at road intersections makes reasoning and interpretation particularly difficult. On the other hand, intersections are a location of many accidents (they represent up to 50% of accidents in some countries), therefore reducing the accident rates in these areas would
have a considerable impact of global traffic safety. We also identified the key issues (and challenges) to be 1) situation understanding and 2) risk assessment, to be carried out from incomplete models and uncertain data. The focus of the year 2010 was on the first of these two problems. We developed a Bayesian Network that could estimate a driver’s intended exit lane at an intersection based on the current state of the vehicle (position, orientation, turn signal state) and on contextual information extracted from the digital map. The idea was to use the information on the geometry of the road network and on the connectivity between lanes to build a statistical model of the relationship between the position and turn signal of a vehicle and the driver’s intended exit lane. Initial results of this work were published in IEEE CIVTS’11 [12], then in IEEE IV’11 [13] with a more thorough evaluation.

The objective of the work conducted in 2011 was twofold:

1. Extend the initial system: add some filtering and take into account the priority rules.
2. Estimate the risk of a situation, based on the estimated behavior/intention of the drivers in the scene.

We proposed a probabilistic motion model for vehicles approaching and traversing an intersection that incorporates some knowledge about how the context (i.e. the traffic rules, the presence of other vehicles, the geometry and topology of the intersection) influences vehicle behavior. The distinctive features of our algorithm are:

- The explicit use of priority rules
  Priority rules are explicitly taken into account in the motion model: the necessity for a driver to stop and/or yield to another vehicle at an intersection is estimated, jointly with the driver’s intention to comply. This allows for a flexible and computationally inexpensive computation of risk. Flexible because depending on the final application one can decide to compute different types of risk, e.g. the probability that a specific vehicle is a violator, or the probability that a crash will occur between two vehicles, or the risk of a specific maneuver for a vehicle. Inexpensive because these can be computed without performing trajectory prediction for the vehicles in the scene.

- The assumption that drivers generally respect traffic rules
  Instead of making the classical assumption that vehicles’ trajectories are independent, we model their mutual influences by introducing a prior knowledge that drivers generally respect priority rules. The motion model therefore takes into account the priority rules and the presence of other vehicles to better interpret correctly a vehicle’s behavior. The advantages are twofold. Firstly, we are able to better estimate the maneuver intention of the drivers, which means our situation assessment capabilities are improved. Secondly, risk is estimated with a higher sensitivity. We avoid risk overestimation while still being able to detect dangerous situations as well as the conventional, more conservative, methods.

This reasoning is implemented using a Bayesian filter which estimates the hidden variables M (maneuver intention), D (distance to intersection), H (intention to stop) and H’ (necessity to halt) jointly for all the vehicles in the scene, using the position, speed and heading information shared between the vehicles via V2X communication. Inference on the hidden variables is carried out by a particle filter. The algorithm was described in an INRIA research report [27]. In this report we showed by reasoning on theoretical scenarios that our assumption that drivers tend to respect priority rules should lead to improved situation assessment and risk assessment (see Fig. 14).

Recently, data has been collected at an intersection using the Renault demonstrator vehicles, so that our algorithm can be tested on real data. Preliminary results seem to confirm that the intuitions described in the research report were correct. A Graphical User Interface is in the process of being developed so that demonstrations of the system can be carried out live in the Renault demonstrator vehicles (see Fig. 15).

5.2. Human Centered Navigation in the physical world
Figure 14. Illustration of a scenario where the advantage of taking into account the interactions between vehicles for maneuver prediction is obvious for ADAS applications. The behavior of the red vehicle is interpreted differently depending on whether or not the interactions with the green vehicle are considered.

Figure 15. Graphical User Interface for warning a driver of a violation of priority rules at an intersection (the violator vehicle is displayed in red).
5.2.1. **Goal oriented risk based navigation in dynamic uncertain environment**

**Participants:** Anne Spalanzani, Jorge Rios-Martinez, Amaury Nègre, Arturo Escobedo-Cabello, Christian Laugier.

Navigation in large dynamic spaces has been addressed often using deterministic representations, fast updating and reactive avoidance strategies. However, probabilistic representations are much more informative and their use in mapping and prediction methods improves the quality of obtained results.

Since 2008 we have proposed a new concept to integrate a probabilistic collision risk function linking planning and navigation methods with the perception and the prediction of the dynamic environments [47]. Moving obstacles are supposed to move along typical motion patterns represented by Gaussian Processes or Growing HMM. The likelihood of the obstacles’ future trajectory and the probability of occupation are used to compute the risk of collision. The proposed planning algorithm, call Risk-RRT, is a sampling-based partial planner guided by the risk of collision. Results concerning this work were published in [48] [49] [50].

In 2011, our algorithms were integrated into an embedded software architecture for social aware navigation (see fig. 16). For this purpose we started to migrate our algorithms to a new experimental platform. Moreover, we adapted the code to the open source software called ROS (Robot Operating systems 2) which offers tools to develop robot applications based in state of the art algorithms. Particularly, localization and visualization tools have been used. We have linked the control of our robotic wheelchair, the Risk-RRT planning and the social filter modules described in 5.2.2 into the framework ROS as shown in figure 16. The main objective was to increase the visibility of our approach and develop common libraries with research groups in robotics. In 2011, in the scope of the AEN PAL project, we started a collaboration with the EPI Arobas and complementary developments have been put on the INRIA forge.

![Architecture for the social navigation system in ROS](image)

Next two sections are conducted under the french project PAL “Personally Assisted Living” with a goal to enhance the quality of living by providing more autonomy in the daily activities of the disabled.

5.2.2. **Social conventions based navigation**

**Participants:** Jorge Rios-Martinez, Anne Spalanzani, Christian Laugier.

---

2 Willow Garage Inc., [http://www.ros.org](http://www.ros.org)
The objectives of this work are to integrate the notion of comfort in the classical safe navigation methods. If one consider that the navigation system transports a person, the integration of social conventions in the navigation strategy starts to be crucial. In this work, we propose to integrate the notions of personal space and interaction between people. We propose to enrich the knowlge the robot has, with a representation of the social conventions. The robot must take into consideration interactions to avoid groups of people (even if passing through the group is the “best” path for a conventional planning algorithm), or to join a group with a behavior close to the one of a human. To understand the behaviors of interaction between humans and the management of space, the works developed in the area of sociology to define some concepts as personal space, \( o \)-space and F-formations are used.

- **Personal Space**
  In [53], Hall describes the use of space between humans, he observed the existence of some rules that conducted people to keep distances from others. He proposed a classification of the space around a person (its **Personal Space**) in social interaction in four zones:
  - the public zone > 3.6m,
  - the social zone > 1.2m
  - the personal zone > 0.45m
  - the intimate zone < 0.45m

  This is a useful tool for a robot to understand the intentions of the humans. It is well known that these measures are not stricts and that they change depending on age, culture and type of relationship but the categories proposed explain very well reactions like the uncomfortable sense of a stranger invading your intimate zone or the perception of somebody looking for social interaction because he is entering to your social zone.

- **F-formation**

  ![Examples of F-formations](image)

  \( \text{Figure 17. Examples of F-formations: (a) Vis-a-vis, (b) L-Shape, (c) C-Shape.} \)

In [57], Kendon observed that people interacting in groups follow some spatial patterns of arrangement. When people are executing some activity they claim an amount of space related to that activity, this space is respected by other people and Kendon referred it as individual’s **transactional segment**. This transactional segment can vary depending on body size, posture, position and orientation during the activity. Moreover the groups can establish a joint or shared transactional segment and only the intervenants have permitted access to it, they protect it and others tend to respect it. The \( o \)-space is that shared transactional segment. A F-formation system is the spatial-orientation arrangement that people create, share and maintain around their \( o \)-space. We can see in fig. (17) three examples of F-formations.

The first stage in order to achieve an integration of social concepts with robot navigation was to include estimations of the risk of disturbing personal space and interaction space in the general risk estimation. A strategy to detect interactions in the environment based in the velocity, position and orientation of humans was implemented.
Figure 18. Detecting conversations in the environment lets the robot to take navigation decisions that avoid humans activity interruption

In fig. 18 we observe the results of the proposed integration, the robot (green rectangle) can use the detections of conversations (light ellipses) between humans (blue circles) for add more risk to paths that invade the space of conversations. When a conversation is detected, a bi-dimensional Gaussian $G$ is created to represent the interaction space, also called o-space, the center of this space is approximated by taking into account the the participants’ poses. Then, $G$ is used to obtain an estimation of risk of disturbing by passing around the conversation. The navigation strategy is based on the Risk-RRT algorithm. Details of this approach were published in [18].

5.2.3. Autonomous Wheelchair for the Elderly People’s Assistance

Participants: Arturo Escobedo-Cabello, Anne Spalanzani, Christian Laugier.

The elderly and the disabled are expected to benefit from the new technologies in the field of autonomous navigation robotics. Normal users of electric wheelchairs will also benefit from the development of more automatic functionalities bringing an extra driving comfort, especially during delicate maneuvers such as narrow door passages. This contribution is similar to the installation of driving assistance on a car. A simple improvement of the classical powered wheelchair can often diminish several difficulties of control.

Comfort defined as a state of ease and satisfaction of bodily wants, with freedom from pain and anxiety, has recently emerged as a design goal in autonomous navigation systems. Designers are becoming more aware of the importance of the user when scheming solution algorithms. The idea of comfort is especially important in the case of wheelchairs where the occupants are weak as result of their age or disease.

For any robot that is designed to transport people, the trajectory should be smooth and correspond to the user’s understanding as much as possible. Since human interpretation of the environment often differs from a robot’s interpretation, the decisions taken by the system might seem incomprehensible to a human observer. For example an autonomous vehicle could refuse to move forward due to some obstacle, while a human user would easily be able to move its way through. This undesirable behaviors may prove irritating and with time may lead to users stopping from using the system.

In 2011 we setup a robotic wheelchair as a trial platform. The wheelchair is a differential drive robot equiped with a SICK LMS-200 lidar to get 2D range information from the environment, odometry sensors, and a velocity controller we have also added a kinect sensor in order to perform some in the field of social
interactions. Some basic functions can be executed including the mapping of the environment using a Rao-Blackwellized Particle Filter [52], localization using an Adaptive Monte Carlo Localization approach (AMCL) [91], global planning using an A* algorithm [60] and local reactive planning using the Dynamic Window Algorithm [46].

Alongside we started working with the Kinect sensor to detect and track people. Using the given tracking information, the wheelchair is able to follow a human located in front of it. This behavior is aimed to bring assistance not only to the user but also to the caregiver by allowing him to move without pushing the wheelchair. The technical implementation of the related approaches has been done on the basis of the ROS middleware due to easy integration with other opensource robotics software which benefit sharing and testing developed software.

In 2012 we shall focus on the estimation of the user intentions by learning models of behavior. We’ll then use these models to propose an adaptive autonomous navigation method that best answer the user needs.

5.2.4. Multi-Robot Distributed Control under Environmental Constraints

Participants: Agostino Martinelli, Alessandro Renzaglia.

This research has been carried out in the framework of the European project sFly. In recent years it is revealed more and more the importance of using multi-robot systems for security application, otherwise impossible to be performed by a single robot.

The main problem approached is the optimal surveillance coverage of an unknown and complex environment, i.e. finding the optimal deployment for the robots and the way to safely reach such configuration. The solution for the 2D case without obstacles is already known in literature [39]. On the other hand, for the non-convex case, it is still a difficult problem. In [84] we firstly proposed a possible strategy based on a combination of the repulsive potential field method and the Voronoi partition. Then, in the last two years we have mainly approached the coverage problem by using a new stochastic optimization method. This work is in collaboration with professor Elias Kosmatopoulos, from CERTH (Thessaloniki), and professor Lefteris Doitsidis, from TUC (Crete), partners in the sFly project.

The Kosmatopoulos’s group has proposed a new adaptive stochastic optimization algorithm for a general class of multi-robot passive and active sensing applications [59], [58]. This method possesses the capability of being able to efficiently handle optimization problems for which an analytical form of the function to be optimized is unknown, but the function is available for measurement at each iteration of the algorithm employed to optimize it. As a result, it perfectly suits for multi-robot optimal coverage in non-convex environments, where the analytical form of the function to be optimized is unknown but the function is available for measurement (through the robots’ sensors) for each multi-robot configuration.

The main results obtained for the 2D case by using this method has been published in [85], [83]. We assume the robots are equipped with global positioning capabilities and visual sensors able to monitor the surrounding environment. The goal is to maximize the area monitored by the team, by identifying the best configuration of the team members. Moreover, in 2011, a distributed version of the algorithm was presented in [17]. In multi-robot systems, a distributed approach is desirable for several fundamental reasons. The most important are failure of the central station and limited communication capabilities. The proposed approach has the following key advantages with respect to previous works:

- it can solve the problem in a distributed way;
- it does not require any a priori knowledge on the environment;
- it works in any given environment, without the necessity to make any kind of assumption about its topology;
- it can incorporate any kind of constraints, for instance regarding a possible existing threshold on the maximum distance on the monitored region, or a limited visibility angle;
- it does not require a knowledge about these constraints since they are learnt during the task execution;
- its complexity is low allowing real time implementations.
Figure 19. (a) Wheelchair used in the emotion team, (b) Two people being tracked using the Kinect and the map of the environment done by the wheelchair.
The previous approach has been also extended for the more important and realistic 3D case. Working in collaboration with the ETHZ (Zurich), some simulations using real data, which were collected with the use of a miniature quadrotor helicopter specially designed for the needs of the European project sFly, have been performed (see fig. 20). This work has lead to two joint publications with CERTH and TUC: one conference paper to present (CDC2011) and one journal papers under review, and two joint publications with CERTH, TUC and ETHZ: one conference paper ([10]) and one journal papers under review.

In 2011, this approach has been combined with human aware navigation technics presented in section 5.2.5.

In the next months, the algorithm will be implemented on real MAVs for the final demo of the project. This demo will include experimentation both in indoor and outdoor complex environments.

Finally, a new collaboration with professor Kosmatopoulos has recently begun. The objective of this work is to develop a new efficient and scalable algorithm for multi-robot active control to perform cooperative simultaneous localization and mapping (CSLAM) and target tracking. The main idea is to use a convex optimization algorithm based on Semi-Definite Programming and Sum-of-Squares polynomials. Preliminary simulation results are very promising and a journal paper is under preparation.

5.2.5. Exploring stochastic optimization method to navigate between humans


Suppose that we have a robot navigating in an unknown and complex environment where people are moving and interacting. In such scenario the respect of the humans’ comfort becomes an important goal to achieve. The discomfort concept could be very general but we focus on the one mentioned before, i.e., the discomfort caused by disturbing one interaction or a personal space of humans. The approach here is to minimize the discomfort while the robot is navigating. As we cannot measure directly the value of discomfort, we can...
infer it by modeling the concepts presented before using simple equations and after by applying a method of optimization. We propose to exploit a new stochastic and adaptive optimization algorithm (CAO) [59]. This method is very useful in particular when the analytical expression of the optimization function is unknown but numerical values are available for any state configuration. Furthermore, the proposed method can easily incorporate any dynamical and environmental constraints. To validate the performance of the proposed solution, several simulation results are provided.

Figure 21. Simulation of the robot navigating in an environment populated by people at three different times, three humans walking and two in conversation. Above discomfort function, below image of scenario, people represented by circles, robot’s positions represented by small triangle.

In fig. 21 the model for discomfort function is shown together with robot navigation. At each step the robot randomly generate configurations in the environment and selects the one that takes it closer to the goal while minimizing values for the discomfort function of humans in the environment, this is repeated until goal is reached. Several executions of proposed approach in different scenarios can be observed in fig. 22.

The details of this approach have been submitted to ICRA2012.

5.3. Bayesian Modelling of Sensorimotor Systems and Behaviors

Results proposed in this section were done in collaboration with the LPPA collège de France.

5.3.1. Bayesian programming applied to a multi-player video games

Participants: Gabriel Synnaeve, Pierre Bessière.

The problem addressed in this work is the autonomous replacement of a human player. It is the continuation of last year’s work on the same topic as well as a follow-up of previous E-Motion Ph.D Ronan Le Hy [61]. This year, we focused on real-time strategy (RTS) games, in which the players have to build an economy, advance technology, produce and control an army to kill the opponents. From a research point of view, multi-player games are interesting because they stand for a good in-between of the real world and simulations. The world is finite and simulated (no sensors problems) but we didn’t wrote the simulation and the other players are humans (or advanced robots in the case of AI competitions).
This year’s research work focused on plan recognition from noisy and incomplete observations. Previous plan recognition works in multiplayer games were mainly based on planning and case-based reasoning (CBR) \cite{92,68,55,76} or HMMs \cite{41}. CBR allows for taking domain knowledge into account easily while not dealing efficiently with uncertainty/incompleteness of information. HMMs deal with uncertainty quite well but domain knowledge is harder to structure. We found different ways to decompose the joint \( P(\text{Observation}_1:N, \text{Plan}_1:M) \) which allows for tractable and robust inference. For instance with the help of intermediate variables which can be derived from domain knowledge (as we did) or found automatically (e.g. cross-validation on a HMM). Particularly, we were able to structure dependencies between domain knowledge extracted variables using coherence variables. We then learn the parameters of such joint distributions from data. Supervised (labeled), and semi-supervised learning (when we label automatically from clustering) have led to a publication at CIG (IEEE) 2011 \cite{19} and unsupervised learning (using only raw game data) led to a publication at AIIDE (AAAI) 2011 \cite{20}.

On top of the research/evaluation implementation, we also implemented it in our StarCraft: Broodwar’s bot implementation BroodwarBotQ. With this bot, we took part in AIIDE and CIG conferences AI tournaments placing respectively 9th (out of 18) and 4th (out of 10). We also published last year’s result on multiple units control in real-time engagements (see \cite{23}) at CIG (IEEE) 2011 \cite{21}. As optimal micro-management is almost always intractable (P-space) in real situations, we considered each unit as a Bayesian sensory motor robot which makes a fusion of its sensory inputs about the world, the enemy units, but also its allies (without explicit communication for less complexity) and higher level directions. So the units only take short term decision on where to go and who to attack, higher level planning is done at a squad (and then army) level and given as a sensory input. Results in micro-management tournaments are state of the art. In the more general case, they could be improved by reinforcement learning of the models parameters.
We are now working on concurrent goals resources attribution, still in the context of incomplete knowledge about the opponent. We are also working on correlating low-level observations (effects) and high-level inferences (causes) about the enemy strategy to be able to predict its future behavior.

Figure 23. A real-time engagement: where should we go (we consider only the wide arrows)? Who should we fire on (we can fire only on orange arrows pointer units, while violet units are also potentially interesting targets)?

5.3.2. Bayesian modelling to implement and compare different theories of speech communication

Participants: Raphael Laurent, Pierre Bessière, Julien Diard, Jean-Luc Schwartz.

A central issue in speech science concerns the nature of representations and processes involved in communication. The search for phoneme or syllable specific invariants led to three major sets of approaches: motor, auditory and perceptuo-motor theories, which have been widely argued for and against. The debate appears to be stagnating. This work is based on the belief that mathematical modeling of these theories could provide breakthroughs. More precisely, it is proposed that casting these theories into a single, unified mathematical
framework would be the most efficient way of comparing the theories and their properties in a systematic manner.

Bayesian modeling provides a mathematical framework that precisely allows such comparisons. The same tool, namely probabilities, can be used both for defining the models and for comparing them. Moreover, the use of a unified framework implies that common hypotheses would have common mathematical translations. This helps toward more principled studies of the competing theories.

Following this integrative approach, the motor, auditory and perceptuo-motor theories are thus cast into one unifying Bayesian framework in which they all appear as instances of various questions asked to one probabilistic communication model. This allows to compare these theories through quantitative testing in various paradigms. The work is aimed at understanding the differences in the predictions given by the different theories, and from these predictions to suggest experiments involving human subjects.

The model was used first to work on purely theoretical simulations aimed at studying with diverse paradigms the decrease in the performances predicted by the different theories due to communication noise. It was then used to work on plosive syllables production and perception, thanks to VLAM, a vocal tract simulation tool, which allows to map articulatory parameters to acoustic signals.
6. New Results

6.1. Graph Based Knowledge Representation

6.1.1. Knowledge Graph Abstract Machine

Participants: Olivier Corby, Catherine Faron-Zucker, Fabien Gandon.

KGRAM (Knowledge Graph Abstract Machine) is a generic interpreter for W3C SPARQL Query Language that operates not only on RDF graphs but on labelled graphs. The interpreter interacts with the target graph through proxies that implement an interface: \textit{Producer} enumerates edges from the target graph, \textit{Evaluator} evaluates filters and \textit{Matcher} takes entailments into account.

This year, work has been done to leverage KGRAM up to SPARQL 1.1 Query Language & Update. It implements most of current version of the recommendation, except the \textit{service} statement. It passes almost all W3C SPARQL 1.1 test cases.

In addition, the Corese Semantic Web Factory has been redesigned and modularized into release 3.0 entirely based on KGRAM interfaces and proxies. Corese 3.0 is a new lightweight RDF/S implementation with SPARQL 1.1. We ported the former Inference Rule engines (forward and backward engines) onto Corese 3.0. We also ported former SPARQL extensions: approximate search based on ontological distance, SQL and XPath in SPARQL 1.1, edge enumeration and length of Property Path, pragmas.

This new version is already used in several applications among which: cartography at IGN [28], design constraint modeling at CSTB [35], technological watch in ISICIL ANR project. It is also used in several PhD Theses in the team. A list of applications can be found on Corese Web site\(^6\).

6.1.2. Semantic Web Graph Visualization

Participants: Olivier Corby, Nicolas Delaforge, Erwan Demairy, Fabien Gandon [contact].

Thanks to an INRIA grant (ADT), we design and develop a Semantic Web Gephi Plugin. This plugin is coupling Corese and the Gephi Open Graph Visualization Platform to provide a framework to query and visualize RDF data taking into account their schemas. See the web pages\(^7\).

6.1.3. Semantic Social Network Analysis

Participants: Guillaume Erétéo, Fabien Gandon.

The PhD thesis of Guillaume Erétéo [14] in the context of the ANR project ISICIL allowed us to analyze the characteristics of the heterogeneous social networks that emerge from the use of web-based social applications, with an original contribution that leverages Social Network Analysis with Semantic Web frameworks. Social Network Analysis (SNA) proposes graph algorithms to characterize the structure of a social network and its strategic positions.

Semantic Web frameworks allow representing and exchanging knowledge across web applications with a rich typed graph model (RDF), a query language (SPARQL) and schema definition frameworks (RDFS and OWL). In this thesis, we merged both models in order to go beyond the mining of the link structure of social graphs by integrating two approaches: (1) semantic processing of the network typing and (2) emerging knowledge of online activities.

\(^6\) http://www-sop.inria.fr/teams/edelweiss/software/corese/
\(^7\) http://wiki.gephi.org/index.php/SemanticWebImport
\(^8\) https://gforge.inria.fr/projects/segviz-public/
In particular we investigated how (1) to bring online social data to ontology-based representations, (2) to conduct a social network analysis that takes advantage of the rich semantics of such representations, and (3) to semantically detect and label communities of online social networks and social tagging activities.

This work was published at [15], [14].

6.1.4. Index Summarizing the Content of RDF Triple Stores

Participants: Adrien Basse, Fabien Gandon, Isabelle Mirbel.

We are interested in designing an architecture to support the distribution of a SPARQL query on a small and fixed number of RDF repositories. To do so, the key stage is to characterize the content of the base of each server in order to be able to predict if a server could contribute or not to the answer of a query. In the context of the PhD Thesis of Adrien Basse we propose an algorithm to extract a compact representation of the content of an RDF store. We improved the canonical representation of RDF graphs based on DFS code proposed in the literature by providing a join operator to reduce the number of generated redundant patterns.

6.1.5. Rules for the Web of Data

Participants: Oumy Seye, Olivier Corby.

In the context of this PhD thesis, the focus is on Rules for the Web of data. We are interested in integrating Rule Interchange Format (RIF) - W3C recommendation for exchanging rules on Web - to others W3C technologies. The aim of this year is to study the integration possibilities of RIF-BLD into semantic Web technologies. RIF-BLD is the dialect of RIF for logic-based systems. Firstly, we have studied the state of the art. Secondly we improved the RIF-BLD parser for presentation syntax and XML syntax. As RIF-BLD can be used with RDF data and OWL ontologies, it is interesting to consider RIF inferences in queries on RDF graph structure. That is why we finally study the integration of RIF-BLD into the Corese Semantic Web engine. In this last step, we have implemented the mapping of abstract syntax tree of RIF-BLD to abstract syntax tree of SPARQL. Thus, we can now execute logic inferences of RIF-BLD in the backward engine of Corese.

We have a paper accepted at EGC 2012 presenting RIF2SPARQL [44], a translation of RIF-BLD statements in SPARQL to perform the logical inferences of RIF-BLD on the Corese Semantic Web Factory. These inferences are implemented in backward chaining approach. We have designed and implemented the mapping of RIF-BLD to SPARQL.

6.1.6. Collaborative Management of Interlingual Knowledge

Participants: Maxime Lefrançois, Fabien Gandon.

We are interested in bridging the gap between the world of natural language and the world of the Semantic Web, in particular to support multilingual access to the Web of Data and management of interlingual knowledge bases. We introduce the ULiS approach, that aims at designing a pivot-based NLP technique called Universal Linguistic System, using Semantic Web formalisms, and being compliant with the Meaning-Text theory. Through ULiS, a user could interact with an Interlingual Knowledge base (IKB) in controlled natural language. Linguistic resources themselves (e.g. dictionary, grammar) are part of a specific IKB, thus, actors may enhance them (i.e. the model of the controlled natural language), through requests in controlled natural language (e.g., add a new lexical units, add grammar rules).

In [30] we proposed a novel approach to define Interlingual Lexical Units classes in the Interlingual Lexical Ontology so that they support the projection of their lexicographic definition on themselves using the OWL formalism. This approach is compliant with the Meaning-Text Theory.

In [31], [40] we introduced three basic interaction scenario for ULiS and we proposed and overviewed the layered architecture of ULiS: meta-ontology, ontology, facts; and ontology, interlingual knowledge, situational knowledge.

We have started a collaboration with the RELIEF project that deals with the construction of a French Lexical Network (Alain Polguère, CNRS-ATILF).
6.1.7. Reuse of Data Analytics Contents and Processes

Participants: Corentin Follenfant, Fabien Gandon, Olivier Corby.

Industrial Business Intelligence (BI) proposes tools and methods to perform data analysis over heterogeneous enterprise sources. They allow one to harvest, federate, clean, annotate, query, organize and visualize data in order to support decision making with human-readable documents such as reports, dashboards, mobile visualizations. Such processes currently require expertise in technical domains like relational modeling in order to produce relevant content.

Users willing to do so without following the learning curve have to reuse existing content to create new one, and need to be guided throughout the workflow. Recommender systems can contribute to easing their progression, but most of them will operate inside walled garden for specific tasks instead of assisting the user throughout his workflow.

Semantic Web tools allow us to provide a common ground for modeling the different operations that compose BI workflows with RDFS vocabularies, capturing usage of the underlying transformations operators within document repositories with RDF graphs, and enabling further composition and reuse of BI operations to achieve new analysis. We introduced with [38] an extension of the RDF Data Cube vocabulary to describe these operations as flexible services that are composed by matching multidimensional data structures interfaces, and validated this model on a production repository containing 900 BI documents decomposed into 8000 documents snippets.

The underlying sequence of operations specific to each snippet was then extracted into a unique RDF graph. Aggregate SPARQL queries allow us to compute basic usage statistics for BI operations that can feed recommender systems such as BI workflows wizards. Besides refining the proposed model, next steps include evaluating the technical usability of SPARQL property paths patterns for data lineage and to identify frequent patterns in sequences of BI operations.

This PhD Thesis is done with a CIFRE industrial grant from SAP Research.

6.2. Interaction Design

6.2.1. Question Answering over Linked Data

Participant: Elena Cabrio.

While an increasing amount of semantic data is being published on the Web, the crucial issue of how typical Web users can access this body of knowledge comes to light. This PostDoc project focuses on the development of methods for a flexible mapping between questions in natural language, and data objects. The main purpose is to allow an end user to submit a query to an RDF triple store in English and get results in the same language, masking the complexity of SPARQL expressions and RDFS/OWL inferences involved in the resolution, but at the same time profiting from the expressive power of these standards. In particular, we address the problem of automatic identification of the relevant relations in Question Answering (QA), to capture the context in which the requests should be interpreted, to be able to determine the constraints on the database query.

We aim at investigating the applicability of the Textual Entailment (TE) approach, recently proposed as a general framework for applied semantics, where linguistic objects are mapped by means of semantic inferences at a textual level [55]. According to such framework, entailment relations can be detected between an input question and a set of relational patterns that represent possible lexicalizations of the relations of interest. Such relations, collected in a pattern repository, can be associated to a SPARQL query to the database. A TE system should therefore first try to establish an entailment relation between an input question and each of the relations in the pattern repository. Then, the SPARQL queries associated to the relations for which the entailed patterns have been found are composed in a single query to the database.


Since this PostDoc research work has just started, our early efforts were directed toward the study of the state of the art on QA over Linked Data. We are currently carrying out a feasibility study on the extraction of the relational patterns from Wikipedia (as the source of free text) and the use of DBpedia as a linked data resource. For the experimental part, we are considering energy and environment as the reference scenario.

6.2.2. Mobile Access to the Web of Data

Participants: Luca Costabello, Fabien Gandon.

This thesis, directed by F. Gandon and I. Herman (CWI and Semantic Web Activity Lead at W3C) deals with accessing the Web of Data from mobile environments. The first year addressed the multi-faceted relationship between ubiquitous consumption of Linked Data and mobile context. More specifically, focus has been put on RDF adaptive representation and on context-aware SPARQL endpoints access control.

When accessed from devices immersed into ubiquitous environments, RDF resources must be properly adapted to the mobile context in which the consumption is performed. A domain-independent, lightweight vocabulary for displaying Web of Data resources in mobile environments has been designed (PRISSMA, Presentation of Resources for Interoperable Semantic and Shareable Mobile Adaptability [36]). The vocabulary is the first step towards an adaptive rendering engine for RDF data coupled with a declarative framework to share and re-use presentation information for context-adaptable user interfaces for Linked Data.

Another line of research regards the role of mobile context in restricting access to the Web of Data. Ubiquitous connectivity enables new scenarios in consuming Linked Data and access control in such pervasive environments must not ignore the mobile context in which RDF consumption takes place, as uncontrolled access in given situations may be undesired by data providers. The work led to enhance the access control framework for SPARQL endpoints proposed by teammate Serena Villata (see 6.2.3) with the notion of mobile context provided by PRISSMA.

6.2.3. Access Control for the Web of Data

Participant: Serena Villata.

This research activity is mainly focussed on the field of Knowledge Representation. First, we have proposed a new access control model for the Web of Data and the Social Web. In particular, we have introduced the S4AC ontology where the meaning of the access policies and their components is defined. This access control model proposes, also, a contextual evaluation of the accessors’ information. This model has been applied both to the world of Linked Data and to the world of social networks. This research activity has been addressed in the context of the DataLift ANR project [21], [20].

Second, we have continued a research activity in the area of argumentation theory. In this context, we are exploring the use of argumentation theory for modeling trust in those systems which deal with incomplete knowledge, and for providing explanations about the agents’ choices [22], [19], [23], [25].

6.2.4. ISICIL

Participants: Nicolas Delaforge, Fabien Gandon.

As the leading team of the ISICIL project, we have developed many software components (client-side and server-side) to enrich the ISICIL platform. First, the whole server mechanism was redesigned, in order to improve the server performance, to strengthen and modularize the framework as well. Many semantic REST services were added (activity stream, syndication, subscription/notification, graphs and charts visualizations).

In collaboration with Erwan Demairy, in charge of the SegViz ADT, a Gephi-ISICIL connector was implemented, allowing ISICIL users to visualize the results of their SPARQL queries directly into a dynamic graph. A demo of this work was presented during the ISICIL public seminarium in September. Furthermore, projects such as Datalift and ISICIL had brought out the need of an access control model for the Web of Data. For this purpose, we designed the S4AC model and ontology and we realized a prototype to evaluate it based on the ISICIL dataset Figure 4.

\[ \text{http://dbpedia.org} \]

\[ \text{http://ns.inria.fr/s4ac/} \]
Since the Philoweb conference in 2010, a workshop dedicated to philosophical engineering was attached to the French IC conference in Chambery. We presented there the advancement of a brand new bookmark model called Webmarks which semantically models the user interest on a web resource (Figure 5 & 6). A long paper on this work was accepted in the EGC 2012 conference and will be published in the RNTI journal (Hermann editions) [27].

We also collaborate with the I3S team in a task of semantization of a commercial wiki, called Mindtouch. This wiki is enhanced with semantic description of its content (Figure 7), its users are part of the ISICIL social network and their activities on the wiki are reported on the metadata server. This tool represents the editorial layer of the software bundle developed to improve the business intelligence tasks. This work was also accepted as a long paper in the EGC 2012 conference [26].

Figure 4. S4AC Access Policy Editor

6.2.5. Models and Methods for Representing and Identifying Groups of Individuals and Their Activities

6.2.5.1. Models and Methods for Representing and Identifying "Collective Personas"

Participant: Alain Giboin.

Context of the work: ISICIL project.

As opposed to Individual Personas (which are user models represented as specific, realistic humans), Collective personas are models representing specific, realistic groups of people as such (e.g., teams, communities). Collective personas are aimed to design groupware more closely adapted to groups. In 2010, we updated our review of the existing methods for elaborating collective personas. This year, we published the updated review [17].
Figure 5. Webmark GUI
Figure 6. Webmark Model
Figure 7. SweetDeki Resource Model
6.2.5.2. Models and methods for Representing and Identifying Relationships between Individuals

Participants: Alain Giboin, Neji Bouchiba.

Context of the work: AVISICIL project, in collaboration with researchers from the Kewi team (I3S, UNS) and from the Laboratoire de Psychologie Cognitive et Sociale (UNS) who are involved in affective computing design projects (designing systems intended to help elderly people maintain their relationships, or autistic children to build relationships with others).

Digital technologies have been claimed to contribute to prevent elderly people from social isolation or loss of social ties. For example, ubiquitous computing, online social networking and affective computing have been reported to facilitate social interaction [64] or to enhance social connectedness [61] among the elderly.

Participating to a project aimed to design a system for recognizing, through various sensors, the affective states (emotions) that indicate a loss or maintenance of social ties, we conducted a social ergonomic study to provide elements of design and evaluation of such a system. Noting that depressive states are among the most significant signs of an actual or potential loss of social ties (see, e.g., [65]), we focused the study in particular on: (a) the models describing the depressive states and the process of their recognition, and the links between these states and the state of social ties; (b) the sensors that can contribute to this recognition. In order to evaluate our solution (so-called GeREmo) with the elderly, we also identified, from an analysis of existing studies on the acceptability of digital technologies, criteria for assessing the acceptability of the GeREmo solution [50].

6.2.6. Comparing and Bridging Models of Shared Representations and Representation Sharing Processes

Participant: Alain Giboin.

Context of the work: GDR CNRS Psycho Ergo, Groupe thématique Coopération homme-homme et Coopéra-
tion homme-machine. Action de recherche RefCom (Référentiel commun), co-leded with Pascal Salembier (UTT).

Sharing representations or shared representations are often claimed to be a key factor for a collaboration to succeed. The notions of shared representations and representation sharing processes are examined in the research literature from several points of view; this variety of viewpoints gave rise to different conceptualizations, which are referred to using such terms as Common Frame of Reference, Mutual Intelligibility, Shared Context, Team/Situation Awareness, etc. In 2010, in order to achieve mutual intelligibility between researchers working on such conceptualizations, we elaborated and asked participants to the RefCom joint research action to test and to apply a grid for collaboratively comparing and bridging the conceptualizations (see Edelweiss activity report 2010 [13]). This year, we analyzed and reported the results of the test and application of the grid [39]. This resulted in a revision of the grid.

6.2.7. Frameworks for taking pragmatic dimensions of ontologies into account

Participant: Alain Giboin.

Context of the work: Follow-up to the Palette European project. This work was done in collaboration with the Centre de Recherche sur l’Instrumentation, la Formation et l’Apprentissage, ULg (Belgium).

When designing ontologies, ontologists (i.e., knowledge engineers specialized in ontology engineering) most often focus on the semantic dimensions of ontologies (such as expressiveness, level of granularity, etc.). Pragmatic dimensions, i.e. dimensions related to the context of use (including the purpose) of the ontologies, are often neglected whereas they are critical to users: ontologies indeed are used in context. In brief, pragmatic dimensions are not taken seriously into account when engineering ontologies but they have to.

13 http://raweb.inria.fr/rapportsactivite/RA2010/edelweiss/
We developed a framework to analyze the way we attempted, in the context of the Palette EU project, to contextualize the ontologies underlying a set of semantic knowledge services dedicated to communities of practice. The framework was derived from the Ontology Framework elaborated by members of the Ontology Engineering community during the Ontology Summit 2007 [62]. Both frameworks define a series of “pragmatic dimensions” of ontologies. Because our derived framework did not cover all possible dimensions, we complemented it, by relying on existing work from the Ontology Engineering community in general, and from the Pragmatic Web community in particular [16].

6.2.8. Explanation of Semantic Web Query Results

Participants: Rakebul Hasan, Fabien Gandon, Olivier Corby.

This PhD thesis, directed by Fabien Gandon and Olivier Corby, aims at opening the query-solving mechanism to the users, and handling and explaining the distribution of a query over several sources on the Semantic Web. This work is part of the Kolflow ANR project.

The current Semantic Web search engines are not able to explain how a given query result is obtained or why it has failed to obtain a result. The goal of opening the query-solving mechanism is to enable the Semantic Web query engines to explain the query solving process taking into account the inferences used to obtain the results for a given query. In addition, explanation of the performance indicators of the query-solving process contributes to the understanding of the resolution process. These performance indicators can be effectively used to help in formulating queries by suggesting alternative queries based on the history of the performance of the query-solving process. Another focus of this thesis is on how the distribution of the queries can be performed over the distributed sources and how explanation can be used to better understand the queries and their results over the distributed sources.

In the early stage of this thesis, our current focus is on explaining the Semantic Web query results taking the inferences into account. We are working on justification of results for SPARQL query with RDFS entailment. Our next focus will be on the different abstractions of these justifications with different degree of details and different types of presentations depending on different level of user expertise.

6.2.9. Pervasive Sociality through Social Objects

Participants: Nicolas Marie, Fabien Gandon.

The work is related to semantic spreading activation algorithm, from idea to first results and visualization. Spreading activation is a method for searching semantic networks by labeling a set of initial nodes with weights (called activation), propagating (spreading) that activation out to other nodes linked to the source nodes and iterating propagation. Previously, at the end of 2010, we designed an ontology called OCSO 14. This ontology aims at describing in a structured format social objects (content augmented by social functionalities independently of its nature: video, place, text, etc.) and corresponding social activity. Then, the need of powerful and semantic sensitive algorithm to process such data led us to follow the track of semantic spreading activation.

Two posters were published at IC [42] and Web Science [43] presenting OCSO model and research axis about semantic spreading activation. A state of the art about exploitation of semantics in spreading activation and its position in the general context of this algorithm family was written. Then a formal proposition was made and algorithm development started leading to experimental results. The state of the art, the formal proposition and early results were published at the Social Objects workshop [41]. The end of the year was mainly focused about results visualization through Gephi and knowledge acquisition on algorithm and its behavior through multiple tests.

14 http://ns.inria.fr/ocs0/V0.2/
6. New Results

6.1. High level model for shapes

6.1.1. Constructive implicit modeling

Participants: Adrien Bernhardt, Marie-Paule Cani, Maxime Quiblier, Cédric Zanni.

Implicit surfaces are an appealing representation for free-form, volumetric shapes. In addition to being able to represent shapes of arbitrary topological genus, they have the ability to be constructed by successively blending different components, which eases interactive modeling.

Within Cédric Zanni’s PhD, we are collaborating with a researcher in formal computation, Evelyne Hubert, to improve and extend the analytical methods for computing closed form solutions for convolution surfaces. We introduced a warping method for enabling the modeling of complex helical shapes from a single implicit primitive (fig. 3), which greatly enhances efficiency [14]. We also proposed a method based on anisotropic, surface Gabor noise, for generating procedural details on skeleton-based implicit surfaces. The surfaces enhanced with details can still be smoothly blended, with a natural transition between the details they carry. A paper has been submitted for publication. We are currently developing normalized convolution surfaces, invariant through homothetic transformations, and which will provide an intuitive blending sharpness parameter, usable with simple additive blending.

Lastly, we contributed to a new blending operator, gradient blending, which enables us to blend implicit shapes not only in function of the field values but also of their gradients. This solves a number long standing problems in implicit modeling: we can generate bulge-free blending, ensure that the topological genus of the blended shape remains the one of the union of the input one, and avoid the blur of small details. A paper is currently submitted for publication.
6.1.2. Ontology-based mesh segmentation
Participant: Olivier Palombi.

The smart use of data by automated systems is now a problem having implications as various as the optimization of the functioning of web search engine or medico-surgical simulation. The program MyCF is an attractive innovation in this field, gathering an organization of the biomedical knowledge in the form of ontologies, 3D acquisition of anatomical structures and also the possibility to export these 3D structures to biomechanics simulation programs, like SOFA. Our work consisted in creating, thanks to the program Protégé, an ontology of the functions of the human body (which didn’t existed then) and to couple it with the anatomical ontology FMA (ontology we significantly reworked). The objective was to build connections introducing a link between anatomical structures and functions they bear. Once this goal reached, it became possible for a computer to link the elimination of an anatomical structure and the loss of a function, this real-time. Thus we established the foundations of an ontology which currently gather 84484 entities (4330 of which are functional entities) and 4159 relations. So, our contributions are threefold: first of all, the creation of an ontology of functions of the human body which is an original one; then the redrafting of the FMA ontology which allowed us to complete some lacks and above all to make of it a tool more oriented towards the practical applications we waited about it, and finally, the contribution which seems to us the more significant is, without any doubt, the institution of a link between these two ontologies. This work is at is moment unfinished and should be pursued in order to approach always more the reality in the field of medico-surgical simulation.

6.1.3. French translation of the Foundational Model of Anatomy (FMA) ontology
Participant: Olivier Palombi.

The goal of this study, performed in collaboration with the LITIS was to facilitate the translation of FMA vocabulary into French. We compare two types of approaches to translate the FMA terms into French. The first one is UMLS-based on the conceptual information of the UMLS metathesaurus. The second method is lexically-based on several Natural Language Processing (NLP) tools. The two approaches permitted us to semi-automatically translate 3,776 FMA terms from English into French, this was to added to the existing 10,844 French FMA terms in the HMTP (4,436 FMA French terms and 6,408 FMA terms manually translated) [9].

6.1.4. Homology computation
Participants: Dobrina Boltcheva, Jean-Claude Léon.

This work is a part of the BQR project IDEAL (see Section 8.1.2) which is performed in collaboration with Leila de Floriani from the University of Genova in Italy. The main goal of this project is to study non-manifold geometrical models and to find out features allowing to classify these models and criteria for determining their shape. We are interested in non-manifold models such as idealized industrial CAD models, since they are still ill-understood even if they are frequently used in computer graphics and many engineering applications. This year, we have worked on the computation of topological invariants on non-manifold simplicial complexes, such as the homology groups, since they play a crucial role in the field of shape description and analysis. The goal was also to acquire a better understanding of the behaviour of the homological groups on non-manifold models. A first step towards this goal has been achieved this year and we have developed an efficient method for computing the homology of a large simplicial complex from the homologies of its sub-complexes. This work has been already published in a research report in 2010 and a journal paper in the context of the international conference on Solid and Physical Modeling [2].

6.1.5. Creased paper modeling
Participants: Marie-Paule Cani, Stefanie Hahmann, Damien Rohmer.

Although very common in real life, 3D creased paper models are rarely seen in virtual scene due to the lack of available modeling tools.

---

1Laboratoire d’Informatique, de Traitement de l’Information et des Systèmes (LITIS) Institut National des Sciences Appliquées de Rouen – Université du Havre – Université de Rouen
We developed a new approach to efficiently generate a 3D model of creased paper from a boundary curve (see fig. 4). The generated surface lies on the given boundary curve while preserving the lengths with respect to the original pattern. Contrary to other approaches, this method can seamlessly handle sharp creases while automatically generating the optimal mesh to this shape. The generation is fast enough to be used interactively, and the physical properties such as developability are approximately preserved. This generation of static surface generation has been published in EG short paper [22].

6.1.6. Spline surface models for arbitrary topologies

Participant: Stefanie Hahmann.

In geometric modeling, quad meshes have always been popular, in the sense that NURBS surfaces which are composed of a tensor product network of quadrilateral patches, are the inevitable standard for describing free-form shapes. They are defined on a chess board like assembly of quadrilateral parameter domains. But when modeling shapes of arbitrary topological type with tensor product NURBS, it is necessary to overcome the restriction of the tensor product configuration where always 4 patches meet at a common corner by using singular parameterizations.

We are developing new smooth parametric surface models defined on irregular quad meshes that combine the advantages of both, the arbitrary topology of quad meshes and the smoothness of the tensor product patches. Herein, tensor product polynomial patches are assembled with tangent plane continuity, in one-to-one correspondence to the mesh faces. They are thus capable to represent manifold shapes of arbitrary topological type since no restriction on the number of patches assembled around a mesh vertex exists.

While subdivision surfaces can also produce a smooth shape with only a few subdivision steps from a coarse mesh, our parametric surfaces have the advantage to provide an explicit parameterization. Moreover, all classical modeling operations such as trimming, intersection, blending and boolean operations can be performed with parametric patches. Tensor product patches can furthermore make profit from the powerful tools of existing modeling systems for purposes of evaluation, display, interrogation and all operations cited above.

In collaboration with G.-P. Bonneau (Artis team) several parametric triangular surface models for arbitrary topologies have been published in the past (CAGD, IEEE TVCG and ACM ToG). A new tensor product spline surface model has been developed this year. It solves the problem of defining a $G^1$-continuous surface interpolating the vertices of an irregular quad mesh with low degree polynomial tensor product patches. It further aims to produce shapes of very high visual quality while reducing the number of control points. A comparison with existing methods and a journal paper are on-going work.
6.1.7. Point sampled surfaces
Participant: Stefanie Hahmann.

Point sampled geometry from scanned data exhibits very characteristic shapes, due to the presence of sharp features in most manufactured and designed objects. Therefore, reconstruction of surfaces from unorganized point sets using MLS fitting requires additional attention. In fact, it is an inherent property of MLS fitting to produce smooth surfaces, thus all sharp features in the point cloud may also be smoothed out. Instead of searching for appropriate new fitting functions our approach was to introduce a new method for selecting an appropriate local point neighborhood for the projection operator so that a standard MLS fitting automatically reproduces sharp features.

This work was part of Christopher Weber’s Ph.D. thesis, which has been co-advised by S. Hahmann and H. Hagen from TU Kaiserslautern, Germany. First part of the work on Gauss map clustering for feature point detection has been published in the SMI 2010 proceedings. The second part has been submitted to a Computer Graphics journal. The thesis has been defended in August 2011.

6.1.8. Volume preserving Free-Form Deformations
Participant: Stefanie Hahmann.

Figure 5. Comparison between standard FFD deformation (top) and our method preserving the volume (bottom).

Free Form Deformation (FFD) is a well established technique for deforming arbitrary object shapes in space. Although more recent deformation techniques have been introduced, amongst them skeleton-based deformation and cage based deformation, the simple and versatile nature of FFD is a strong advantage, and justifies its presence in nowadays leading commercial geometric modeling and animation software systems. Several authors have addressed the problem of volume preserving FFD. These previous approaches however make either use of expensive non-linear optimization techniques, or resort to first order approximation suitable only for small-scale deformations. Our approach was to take advantage from the multi-linear nature of the volume constraint in order to derive a simple, exact and explicit solution to the problem of volume preserving FFD. Two variants of the algorithm have been developed, without and with direct shape manipulation. Moreover, we showed that the linearity of our solution enables to implement it efficiently on GPU.

This work has been done in collaboration with G. Elber from TECHNION, H. Hagen from TU Kaiserslautern, G.-P. Bonneau and S. Barbier from Artis INRIA. It has been accepted for publication in the journal The Visual Computer [6].

6.2. High level models for animation

6.2.1. Geometrical methods for skinning character animations
Participants: Marie-Paule Cani, Stefanie Hahmann, Damien Rohmer.
Skeletal animation is an efficient and widely used technique in video games or movie industry due to its flexibility and simplicity. Still, the skinning method do not take into account informations about the physical properties of the underlying material. Therefore effects such as muscle bulging or fat tissue compression cannot be modeled without the addition of a tedious manual correction. Within Damien Rohmers PhD [1], an active geometry framework was proposed in order to enhance geometry information with a priori knowledge about how the underlying material can deform. For instance, bending the belly of an animal will be constraint to generate bulges that will preserve locally the volume (see fig. 6). The process can be either purely automatically generated, or it can be artistically controled.

6.2.2. Action representation, segmentation and recognition

Participant: Remi Ronfard.

Following Daniel Weinland’s PhD thesis, we published a survey of modern methods for representing, segmenting and recognizing full-body actions in video [12]. A taxonomy of methods is elaborated in that paper, where actions can be represented with local, structured or global features both in time and in space. The potential for future work in grammar-based action recognition is emphasized, with possible applications in corpus-based procedural modeling of actions.

6.2.3. Frame-based simulation of deformable solids

Participants: Guillaume Bousquet, François Faure.

We present a new type of deformable model which combines the realism of physically based continuum mechanics models and the usability of frame-based skinning methods [4]. The degrees of freedom are coordinate frames (see Figure 15). In contrast with traditional skinning, frame positions are not scripted but move in reaction to internal body forces. The displacement field is smoothly interpolated using blending techniques such as dual quaternions. The deformation gradient and its derivatives are computed at each sample point of a deformed object and used in the equations of Lagrangian mechanics to achieve physical realism. This allows easy and very intuitive definition of the degrees of freedom of the deformable object. The meshless discretization allows on-the-fly insertion of frames to create local deformations where needed. We formulate the dynamics of these models in detail and describe some pre-computations that can be used for speed. We show that our method is effective for behaviors ranging from simple unimodal deformations to complex realistic deformations comparable with Finite Element simulations.
We extend the approach (see fig. 7) to the simulation of complex, intricately material distributions using material-aware shape functions [3]. Given a volumetric map of the material properties of an object and a number of control nodes, a distribution of the nodes is computed automatically, as well as the associated shape functions. Reference frames are attached to the nodes, and deformations are applied to the object using linear blend skinning. A continuum mechanics formulation is derived from the displacements and the material properties. We introduce novel material-aware shape functions in place of the traditional radial basis functions used in meshless frameworks. These allow coarse deformation functions to very efficiently resolve non-uniform stiffnesses. Complex models can thus be simulated at high frame rates using a small number of control nodes.

To encourage its use, the software is freely available in the simulation platform SOFA.

6.3. Towards interactive digital creation media

6.3.1. Sketch-based modeling and shape editing

Participants: Adrien Bernhardt, Rémi Brouet, Marie-Paule Cani, Jean-Claude Léon, Olivier Palombi.
3D modeling from a sketch is a fast and intuitive way of creating digital content.

We developed a method based on convolution surfaces for inferring free-form shapes in 3D from arbitrary progressive sketches, without any a priori knowledge on the objects being represented (see the section describing the Aestem Studio software. We recently investigated whether 2D deformation could be a better approach than sketching for defining the 2D sketch given as input [18]. Results are very promising; We are planning both to allow such intuitive deformations, combined with sketching, within the Aestem software, and to extend them to the editing of the 3D deformed shape. This will be done in the context of Rémi Brouet’s PhD thesis, co-advised by Renaud Blanch from the I3HM/LIG team.

We also develop methods for interpreting complex sketches (contours with T-junctions) based on some a priori knowledge. Our first work on this topic used the conventions of anatomical drawing to infer the 3D geometry of vascular systems, with branching and occlusions, from a single sketch [10]. We are also investigating the design of realistic terrains from a single sketch, within the PhD thesis of Adrien Bernhardt (see fig. 8). Our first advances include a new representation and new methods for generating a high-field from user-sketched constraint curves [30], [17].

6.3.2. Free-form sculpture

Participants: Marie-Paule Cani, Lucian Stanculescu.

In the context of Lucian Stanculescu’s PhD thesis, co-adviced by Raphaëlle Chaine from LIRIS (Lyon), we developed an interactive sculpting system enabling both arbitrary deformation and topological changes of a free-form shape [11]. Our method is based on a semi-regular mesh which adaptively refines and changes its topology according to the need. See Figure 9. We are currently extending the method for handing the sculpting of composite objects made of many different components.

![Figure 9. Sculpting with topological changes made within the framework from [11].](image)

6.3.3. Hand Navigator

Participants: Jean-Rémy Chardonnet, Jean-Claude Léon.

The different deformation models we developed in the past few years open the problem of providing intuitive interaction tools for specifying the desired deformations in real-time. Therefore, our recent work focused on developing new devices for interacting with the model to deform. For the past two years, we focused on developing a peripheral device similar to a mouse, called the HandNavigator, enabling to control simultaneously ten or more degrees of freedom of a virtual hand. This device consists in a 3D mouse for the position and orientation of the hand in 3D space, enhanced with many sensors for moving and monitoring the virtual fingers. Thanks to a pre-industrialization project funded by the incubator GRAVIT, the first prototype, patented by INRIA, has been extended with the incorporation of new sensors and new shapes to improve the device efficiency. An ongoing extension of the patent and a partnership with HAPTION company are new
step toward the industrialization of this device. Dissemination to general public has been performed at the “Fête de la Science” and another exhibition. Publications will take place after setting up the patent extension. The ongoing BQR INTUACTIVE funded by Grenoble-INP will lead to further scientific topics regarding interactions during grasping as well as with deformable bodies and a partnership has been set up with GIPSA-Lab to study the muscular activity during interactions.

6.3.4. Procedural modeling

**Participants:** Marie-Paule Cani, Arnaud Emilien.

We developed a method for procedurally generating villages with the appropriate roads and streets on arbitrary terrains, in collaboration with Eric Galin from LIRIS, Lyon. This work will be continued within Arnaud Emilien’s PhD thesis towards more general models for populating terrains with houses, vegetation, and animals. We will focus on the development of intuitive ways to edit procedural models, to overcome the main drawback of these approaches.

6.3.5. Computational model of film editing

**Participant:** Remi Ronfard.

![Interface developed for automatic editing of animated movies](image)

*Figure 10. Interface developed for automatic editing of animated movies [31], [21].*

Building on Remi Ronfard’s experience leading the virtual cinematography research team at Xtranormal Technology, Montreal, we designed a novel computational model for automatic editing of animated movies (see fig. 10). A prototype has been implemented in a collaboration with the Bunraku/mimetic team, and demonstrated in poster sessions at the Symposium on Computer Animation (SCA) [31] and International Conference on Interactive Digital Storytelling (ICIDS) [21]. This early work opens new directions that will be further explored by the IMAGINE team, including corpus-based learning of cinematography and editing styles.
6. New Results

In the continuation of our previous work, in 2011 we developed our work on evaluation of ontology matching and especially in running new experiments and generating new tests (§ 6.1.1). We also continued our work on trust in semantic peer-to-peer systems (§ 6.2.2), the use of the $\mu$-calculus for evaluating RDF path queries (§ 6.2.1) and ontology matching for linking data (§ 6.1.2).

6.1. Ontology matching and alignment

We pursue our work on ontology matching and alignment support with contributions to evaluation, data interlinking and multilingual matching.

6.1.1. Evaluation

Participants: Cássia Trojahn dos Santos [Contact], Jérôme Euzenat, Jérôme David.

Evaluation of ontology matching algorithms requires to confront them with test ontologies and to compare the results. Since 2004, we run the Ontology Alignment Evaluation Initiative (OAEI) which organises evaluation campaigns for assessing the degree of achievement of actual ontology matching algorithms [4]. This year, the evaluation campaign had 16 different teams entered the evaluation which consisted of 5 different sets of tests. The participating systems and evaluation results were presented in the 6th Ontology Matching workshop, that was held in Bonn, DE [17][9].

The main activities carried out in 2011 were related to the automation and execution of the OAEI 2011 campaign, in the framework of the SEALS project (see § 8.2.1). This involved the following main tasks:

- describe evaluation processes within the early version of the SEALS platform [11];
- develop a client allowing participants to validate their wrapped tools and evaluate (offline and locally) their tools;
- develop a test generator for automatic generation of systematic benchmarks [12];
- providing participants with a better way to bundle their tools so that they can be evaluated within the SEALS platform; and
- analysis and report of the evaluation campaign results [9].

This work has been used in the OAEI 2011 evaluation campaign. More information on OAEI can be found at http://oaei.ontologymatching.org/.

6.1.2. Ontology matching for linked data

Participants: Zhengjie Fan, Jérôme Euzenat [Contact], Jérôme David.

The web of data consists of using semantic web technologies to publish data on the web in such a way that they can be interpreted and connected together. It is thus critical to be able to establish links between these data, both for the web of data and for the semantic web that it contributes to feed.

In the context of the Datalift project (see § 8.1.1), we are developing a data interlinking module. Based on our analysis of the relationships between ontology matching and data interlinking [13], our goal is to generate data interlinking scripts on from ontology alignments. For that purpose, we have integrated existing technologies within the Datalift platform: the Alignment API, for taking advantage of the EDOAL language and Silk, developed by Frei Universität Berlin, for processing linking scripts. So far we have demonstrated the ability to process simple scripts.

This work is part of the PhD of Zhengjie Fan, co-supervised with François Scharffe (LIRMM), within the Datalift project.
6.1.3. **Multilingual ontology matching**

**Participants:** Cássia Trojahn dos Santos [Contact], Jérôme David, Jérôme Euzenat, Giuseppe Pirrò.

We have participated in the creation of a benchmark for multilingual ontology matching, the MultiFarm dataset. This dataset is composed of a set of ontologies translated in different languages and the corresponding alignments between these ontologies. It is based on the OntoFarm dataset, which has been used successfully for several years in the Ontology Alignment Evaluation Initiative. By translating the ontologies of the OntoFarm set into eight different languages – Chinese, Czech, Dutch, French, German, Portuguese, Russian, and Spanish – we created a comprehensive set of realistic test cases. We plan to include this new dataset in the OAEI 2012 campaign.

Finally, in the context of the Cameleon project (see § 8.3.1) we have been working on the creation of a multilingual comparable corpora using as seed a set of multilingual aligned ontologies. These resources will be exploited in the process of populating and enriching ontologies as well as in the process of cross-lingual ontology alignment.

6.2. **Ontology networks**

Dealing with the semantic web, we are interested in ontology networks, i.e., sets of distributed ontologies that have to work together. One way for these systems to interact consists of exchanging queries and answers. For that reason, we pay particular attention to query systems.

6.2.1. **Path queries and μ-calculus**

**Participants:** Melisachew Wudage Chekol [Contact], Jérôme Euzenat, Pierre Genevès, Nabil Layaïda.

Querying the semantic web is mainly done through SPARQL [15]. One of its extensions, PSPARQL (Path SPARQL) provides queries with paths of arbitrary length. We study the static analysis of queries written in this language with techniques based on μ-calculus interpretation that have been used for XPath. We have more specifically considered PSPARQL query containment: determining whether, for any graph, the answers to a query are contained in those of another query [18][14]. To that extent, we proposed an encoding of RDF graphs as transition systems and PSPARQL queries as μ-calculus formulas. We then reduce the containment problem to testing satisfiability in the logic.

This work is part of the PhD of Melisachew Wudage Chekol, co-supervised with Nabil Layaida (WAM).

6.2.2. **Trust in peer-to-peer semantic systems**

**Participants:** Manuel Atencia [Contact], Jérôme Euzenat, Marie-Christine Rousset.

In a semantic peer-to-peer network, peers use separate ontologies and rely on alignments between their ontologies for translating queries. Nonetheless, alignments may be incorrect –unsound or incomplete– and generate flawed translations, thus leading to unsatisfactory answers. We have put forward a trust mechanism that can assist peers to select those peers in the network that are better suited to answer their queries [8]. The trust that a peer has towards another peer depends on a specific query and represents the probability that the latter peer will provide a satisfactory answer. In order to compute trust, we exploit both alignments and peers’ direct experience, and perform Bayesian inference. We have implemented our technique and conducted an evaluation. Experimental results showed that trust values converge as more queries are sent and answers received. Furthermore, the use of trust is shown to improve both precision and recall of query answers.

This work has been developed in collaboration with Marie-Christine Rousset (LIG) in the context of the DataRing project (see § 8.1.2).
6. New Results

6.1. Autonomous Development of Representations

6.1.1. Open-ended bootstrapping of new sensorimotor representations

Participants: Alexander Gepperth.

We have explored a novel approach to the open-ended development of internal representations in autonomous agents, addressing in particular the transfer of knowledge between different modalities or abstraction levels. We propose a self-organized neural learning paradigm termed PROPRE (projection-prediction) that is driven by predictability: competitive advantages are given to those feature-sensitive elements that are inferable from activity in a reference representation, which may be innate or previously formed by learning. For generating and adapting the new induced representations, PROPRE implements a bi-directional interaction of clustering (“projection”) and inference (“prediction”), the key ingredient being an easy-to-compute online measure of predictability, by which the projection step is encouraged to favor sensitivity to predictable clusters. We demonstrated the potential of this paradigm by several simulation experiments with synthetic inputs. We showed that induced representations are indeed significantly more sensitive to predictable stimuli, that they are continuously being adapted to changing input statistics and that the behavior under severe resource constraints is favorable.

6.1.2. The contribution of context information to object detection in intelligent vehicles

Participants: Alexander Gepperth, Michael Garcia Ortiz.

In this work package, we explored the potential contribution of multimodal context information to object detection in an "intelligent car". The used car platform incorporates several sophisticated processing subsystems, both for the detection of objects from local visual patterns as well as for the estimation of global scene properties, e.g., the shape of the road area, or the 3D position of the ground plane (sometimes denoted "scene context" or just "context"). Annotated data recorded on this platform is publicly available as the "HRI RoadTraffic" vehicle video dataset, which formed the basis for this investigation.

In order to quantify the contribution of context information, we investigated whether it can be used to infer object identity with little or no reference to local patterns of visual appearance. Using a challenging vehicle detection task based on the "HRI RoadTraffic" dataset, we trained selected algorithms to estimate object identity from context information alone. In the course of our performance evaluations, we also analyzed the effect of typical real-world conditions (added noise, high dimensions, environmental variation) on context model performance.

As a principal result, we showed that the learning of context models is feasible with all tested algorithms, and that object identity can be estimated from context information with similar accuracy as by relying on local pattern recognition methods. We also found that the use of basis function representations (also known as "population codes") allows the simplest (and therefore most efficient) learning methods to perform best in the benchmark, suggesting that the use of context is feasible even in systems operating under strong performance constraints.

6.1.3. Discovering object concept through developmental learning

Participants: Natalia Lyubova, David Filliat.

The goal of this work is to design a visual system for a humanoid robot. Taking inspiration from children’s perception and following the principles of developmental robotics, the robot should detect and learn objects from interactions with people and from experiments it performs with objects, avoiding the use of image databases or of a separate training phase. In our model, all knowledge is therefore iteratively acquired from low-level features and builds up hierarchical object models, which are robust to changes in the environment, background and camera motion.
In our scenario, people in front of the robot are supposed to interact with objects to encourage the robot to focus on them. We therefore assume that the robot is attracted by motion and we segment possible objects based on clustering of the optical flow. Additionally, the depth information from a Kinect is used to filter visual input, considering the constraints of the robot’s working area and to refine the object contours obtained from motion segmentation.

The appearance of objects is encoded following the Bag of Visual Words approach with incremental dictionaries. We combine several complementary features to maximize the completeness of the encoded information (SURF descriptor and superpixels with associated colors) and construct pairs and triples of these features to integrate local geometry information. These features make it possible to decide if the current view has been already seen or not. A multi-view object model is then constructed by associating recognized views and views tracked during object motion.

This system is implemented on the iCub humanoid robot, which detects objects in the visual space and characterizes their appearance, their relative position and their occurrence statistics. Ten objects were presented in the current experiment; each of them was manipulated by a person during 1-2 minutes. Once the vocabulary reached a sufficient amount of knowledge, the robot was able to reliably recognize human hands and most of objects.

6.1.4. Scaling-up Knowledge for a Cognizant Robot

Participants: Thomas Degris, Joseph Modayil.

A cognizant robot is a robot with a deep and immediately accessible understanding of its interaction with the environment—an understanding that the robot can use to flexibly adapt to novel situations. Such a robot will need a vast amount of situated, revisable, and expressive knowledge to display flexible intelligent behaviors. Instead of relying on human-provided knowledge, we consider the case where an arbitrary robot can autonomously acquire pertinent knowledge directly from everyday interaction with the environment. We study how existing ideas in reinforcement learning theory can be used to formalize knowledge and use reinforcement learning techniques to enable a robot to maintain and improve its own knowledge. We consider robot performing a continual learning process that scales-up knowledge acquisition to cover a large number of facts, skills and predictions. This knowledge has semantics that are grounded in sensorimotor experience and can then be used for more abstract process such as planning. We see the approach of developing more cognizant robots as a necessary key step towards broadly competent robots.

Paper being published: Scaling-up Knowledge for a Cognizant Robot accepted at Designing Intelligent Robots: Reintegrating AI, AAAI Spring Symposium 2012.

6.1.5. Learning parallel combinations of motor primitives from demonstration and linguistic guidance with non-negative matrix factorization

Participants: Olivier Mangin, Pierre-Yves Oudeyer.

We have elaborated and experimented a novel approach to joint language and motor learning from demonstration. It enables discovery of a dictionary of motor and linguistic primitives, that can be combined in parallel to represent training data as well as novel skills in the form of combinations of known skills. These methods and the results of our experiments participate in addressing two main issues of developmental robotics: 1) symbol grounding for language learning; 2) achieving compositionality in motor-learning from demonstration, which enables re-using knowledge and thus scaling to complex tasks. In particular, we are interested in learning motor primitives active in parallel, a less explored way of combining such primitives. To address these challenges we have explored and studied the use of nonnegative matrix factorization to discover motor primitives from histogram representations of data acquired from real demonstrations of dancing movements. Initial results were presented in [30] and further results are presented in an article under review.


6.2.1. The SAGG-RIAC algorithm: competence based active learning of motor skills

Participants: Adrien Baranes, Pierre-Yves Oudeyer.
We have continued to develop and experiment the Self-Adaptive Goal Generation - Robust Intelligent Adaptive Curiosity (SAGG-RIAC) algorithm as an intrinsically motivated goal exploration mechanism which allows high-dimensional redundant robots with various body schemas to efficiently and actively learn motor skills in their task space. The main idea is to push the robot to perform active babbling in the low-dimensional goal/task space, as opposed to motor babbling in the high-dimensional actuator space (possibly defined with motor primitives), by self-generating goals actively and adaptively in regions of the task space which provide a maximal competence improvement for reaching those goal states. Then, a lower level active motor learning algorithm is used to drive the robot to locally explore how to reach a given self-generated goal. We have conducted systematic experiments with high-dimensional continuous sensorimotor spaces related to different robotic setups such as a highly-redundant robotic arm, a quadruped, and an arm controlling a fishing rod with a flexible wire and show that 1) exploration in the task space can be a lot faster than exploration in the actuator space for learning inverse models in redundant robots; 2) selecting goals based on the maximal improvement heuristics creates developmental trajectories driving the robot to progressively focus on areas of increasing complexity and is statistically significantly more efficient than selecting goals randomly, as well as more efficient than different standard active motor babbling methods. These results were published in [13], [15], [17] and a journal publication is in preparation.

6.2.2. SGIM-D: Bootstrapping Intrinsically Motivated Learning with Human Demonstration
Participants: Mai Nguyen, Pierre-Yves Oudeyer.

We have studied the coupling of internally guided learning and social interaction, and more specifically the improvement owing to demonstrations, of the learning by intrinsic motivation. We have designed Socially Guided Intrinsic Motivation by Demonstration (SGIM-D), an algorithm for learning the mapping between high dimensions in continuous, non-preset, highly redundant environments. We have shown through a robot learning experiment involving a high-dimensional sensorimotor space related to fishing skills that SGIM-D efficiently combines the advantages of social learning and intrinsic motivation to gain a wide repertoire while being specialised in specific subspaces. An article presenting aspects of this work was awarded the second best student paper award in IEEE ICDL/Epirob 2011 [27].

6.2.3. Maturationally-Constrained Competence-Based Intrinsically Motivated Learning
Participants: Adrien Baranès, Pierre-Yves Oudeyer.

We have continued to develop computational models of the coupling of intrinsic motivations and physiological maturational constraints, showing that both mechanisms may have complex bidirectional interactions allowing the active control of the growth of complexity in motor development which directs an efficient learning and exploration process. The coupling relies on the Self-Adaptive Goal Generation - Robust Intelligent Adaptive Curiosity algorithm (SAGG-RIAC) that instantiates an intrinsically motivated goal exploration mechanism for motor learning of inverse models. Then, we have introduced a functional model of maturational constraints inspired by the myelination process in humans, and showed how it can be coupled with the SAGG-RIAC algorithm, forming a new system called McSAGG-RIAC2. We have then conducted systematic experiments to evaluate qualitative and, more importantly, quantitative properties of these systems when applied to the learning of the forward and inverse kinematic of an unknown robotic arm of up to 60 dimensions, the learning of walking in a 12DOF quadruped controlled with 24 dimensions motor synergies, and learning the control of a fishing rod involving a flexible/rope component. These results were published in [13], [15], [17] and a journal publication is in preparation.

6.2.4. Actor-Critic for Parallel Learning
Participants: Thomas Degris, Matha White, Richard Sutton.

Parallel learning is necessary for a robot to learn multiple tasks in parallel while executing a behavior in the environment not necessarily directly related to the tasks to learn. In previous existing work, an interesting class of learning algorithms for control are actor–critic. First, these algorithms can be used with high-dimensional action space. Second, they also sometimes provide computational models for biological decision-making systems. At FLOWERS, we work on new actor–critic algorithms suitable for parallel learning, with
theoretical guarantees, applicable and practical to use with robots, and formulated in the general framework of reinforcement learning.

### 6.2.5. Curiosity for Parallel Learning of Predictions and Tasks from the Continuous Interaction of a Robot with its Environment

**Participants:** Thomas Degris, Adam White, Pierre-Yves Oudeyer.

On one hand, a robot needs a wide variety of knowledge to fully interact with its environment. On the other hand, a robot, like humans or animals, can only perform one behavior at a time in the real world to learn this vast amount of knowledge. A solution to scale up learning while keeping the interaction time with the real world realistic is to learn multiple elements of knowledge simultaneously in parallel. The Horde architecture proposes a set of demons each learning about new policies (i.e. skills) and predictions about these skills (i.e. partial models) off-policy simultaneously. The number of demons learning in parallel is limited only by memory and processing power, and not by the fact that there is only one sensorimotor interaction with the environment to learn from. At FLOWERS, we investigate the question of what the behavior policy of the robot should be to speed-up learning of the demons. Our goal is to test if the Horde scales-up to complex humanoid robots and if, driven by intrinsic motivations, it can autonomously learn building blocks of knowledge for future, more complex, behaviors.

### 6.2.6. Optimal Teaching on Sequential Decision Tasks

**Participants:** Manuel Lopes, Maya Cakmak.

A helpful teacher can significantly improve the learning rate of an autonomous learning agent. Teaching algorithms have been formally studied within the field of Algorithmic Teaching. These give important insights into how a teacher can select the most informative examples while teaching a new concept. However the field has so far focused purely on classification tasks. In this paper we introduce a novel method for optimally teaching sequential decision tasks. We present an algorithm that automatically selects the set of most informative demonstrations and evaluate it on several navigation tasks. Next, we present a set of human subject studies that investigate the optimality of human teaching in these tasks. We evaluate examples naturally chosen by human teachers and found that humans are generally sub-optimal. Then based on our proposed optimal teaching algorithm we try to elicit better teaching from humans. We do this by explaining the intuition of the teaching algorithm in an informal language prior to the teaching task. We found that this improves the examples elicited from human teachers on all considered tasks. This shows that a simple modification the instructions given to human teachers, has the potential of greatly improving the performance of the agent trained by the human [41].

### 6.2.7. Inverse Coordinated Reinforcement Learning

**Participants:** Manuel Lopes, Jonathan Sprauel.

Inverse Coordinated Reinforcement Learning

We extended of inverse reinforcement learning to the multi-agent case. Under this formalism a team of agents can learn a task goal, encoded as a reward function, by observing another team executing that task. Our agents behave using local information and limited communication following the coordinated reinforcement learning framework. We show that a team behavior can be learned using this formalism and how well this mechanism can deal with changing initial conditions and number of agents [68].

### 6.3. Motor Learning and Morphological Computation

#### 6.3.1. Morphological Computation in Acroban the Humanoid: Balance Control and Dynamic Walking

**Participants:** Olivier Ly, Pierre-Yves Oudeyer, Matthieu Lapeyre, Jérome Béchu, Paul Fudal, Haylee Fogg.
We have continued to elaborate and experiment the humanoid platform Acroban and its use to study various scientific topics. Our goal was to study three main issues: 1) Compliance and semi-passive dynamics in the framework of dynamic walking in humanoid robots and more generally its impact in terms of semi-passive interactive motor primitives and their robustness to unknown external perturbations; 2) the advantage of a bio-inspired multi-articulated vertebral column in the dynamics of these motor primitives; The platform uses mechatronic components that allow us to adjust dynamically the compliance of actuators, which combines with the intrinsic mechanical compliance of the structure due to the use of elastics and springs. We have explored how these capabilities can allow us to enforce morphological computation in the design of robust dynamic locomotion. Compliance also allows us to design semi-passive motor primitives using the torso as a system of accumulation/release of potential/kinetic energy. This is made possible by the combination of adequate morphology and materials, full-body compliance, semi-passive and self-organized stable dynamics, as well as the possibility to experiment new motor primitives by trial-and-error thanks to light-weightedness. These results were presented in [25]. A dedicated web page with videos is available at: http://flowers.inria.fr/acroban.php.

6.3.2. Maturational constraints for motor learning in high-dimensions: the case of biped walking

Participants: Matthieu Lapeyre, Pierre-Yves Oudeyer, Olivier Ly.

We have elaborated and began to experiment a new developmental approach to motor learning in very high-dimensions, applied to learning biped locomotion in humanoid robots. This approach relies on the formal modeling and coupling of several advanced mechanisms inspired from human development for actively controlling the growth of complexity and harnessing the curse of dimensionality: 1) Maturational constraints for the progressive release of new degrees of freedoms and progressive increase their explorable ranges; 2) Motor synergies; 3) Morphological computation; 4) Social Guidance. An experimental setup involving a simulated version of the Acroban Humanoid robot, based on the V-REP simulator, has been elaborated, and initial encouraging results were obtained. These results are presented in [23].

6.3.3. Acroban v2: improving morphological computation with dampers

Participant: Olivier Ly.

Theoretical studies and experiments concerning in particular dynamics of passive walkers drove us to design, construct and continue to experiment a new version of Acroban. This new version has two goals both fitting in the study of the impact of morphology in the behaviour of the robot:

- experiment deep structural modifications of the morphology, in order to avoid as much as possible inelastic chocks. Indeed, during the gait, the unstability is mainly due to chock at the landing of the foot.
- improve the global ratio weight/power of the robot in order to get more dynamic movements.

Indeed, this new version uses RX-28 motors which are lighter than the RX-64 motors which are used in the first version of Acroban. The robot is smaller and lighter. First experiments show that the obtained ratio weight/power is better than the first version. Movements of the robot, and in particular amplitude of locomotion movements, are not limited by torque now. Second, we have experimented plastic materials to design the structure in order to makes it naturally flexible comparing to the metal used in the first version. This way, we improve the natural compliance of the robot. Finally, and this is probably the most important change, we used non actuated linear joints in the hip and in the spline. To control these linear joints, instead of servo-motors, we use dampers. This kind of design is new in humanoid robotic. While bringing new control problems (because of the non-controlled joints which makes the robot semi-passive), this design softens chocks in a significant maner. Experiment shows that stability of the whole structure is greatly improved especially during locomotion.

6.4. HRI and Robot Language Teaching

6.4.1. Intuitive and Robust Physical Human-Robot Interaction with Acroban

Participants: Olivier Ly, Pierre-Yves Oudeyer, Pierre Rouanet, Matthieu Lapeyre, Jérome Béchu, Paul Fudal, Haylee Fogg.
We have experimented and shown how the humanoid robot Acroban allows whole-body robust, natural and intuitive physical interaction with both adults and children. These physical human-robot interaction are made possible through the combination of several properties of Acroban: 1) it is whole-body compliant thanks to variable impedance control and also thanks to the use of elastics and springs; 2) it has a bio-inspired vertebral column allowing more flexibility in postural and equilibrium control; 3) it is light-weight; 4) it has simple low-level controllers that leverage the first three properties. Moreover, the capabilities for physical human-robot interaction that we show are not using a model of the human, and in this sense are “model free”: 1) the capability of the robot to keep its equilibrium while being manipulated or pushed by humans is a result of the intrinsic capability of the whole body to absorb unpredicted external perturbations; 2) the capability of leading Acroban by the hand is an emergent human-robot interface made possible by the self-organizing properties of the body and its low-level controllers and was observed a posteriori only after the robot was conceived and without any initial plan to make this possible. Finally, an originality of Acroban is that is is made with relatively low-cost components which lack of precision is counterbalanced with the robustness due to global geometry and compliance. These results were presented in [28]. A dedicated web page with videos is available at: http://flowers.inria.fr/acroban.php.

6.4.2. A Real World User Study of Different Interfaces for Teaching New Visually Grounded Words to a Robot
Participants: Pierre Rouanet, Pierre-Yves Oudeyer, Fabien Danieau, David Filliat.

We have continued to elaborate and experiment an integrated system based on a combination of advanced Human-Robot Interaction, visual perception and machine learning methods that allows non-expert users to intuitively and robustly teach new visually grounded words to robots. This system is based on the state-of-the-art bags of words technique but focuses on different mediator based interfaces that we can propose to the users. Indeed, we argue that by focusing on interaction we could help users to collect good learning examples and thus improve the performance of the overall learning system. We compared four different interfaces and their impact on the overall system through a real world study where we asked participants to show and teach a robot names for five different objects. Three interfaces were based on mediator objects such as an iPhone, a Wiimote and a laser pointer and provided the users with different kinds of feedback of what the robot is perceiving. The fourth interface was gesture based with a Wizard-of-Oz recognition system included in order to compare our mediator interfaces with a more natural interaction. We showed that the interface may indeed strongly impact the quality of the learning examples collected by users, especially for small objects. More precisely, we showed that interfaces such as the iPhone interface do not only give feedback about what the robot is perceiving but also drive users to pay attention to the learning examples they are collecting. Thus, this interface allows non-expert users to intuitively and easily collect almost as good learning examples as expert users trained for this task and aware of the different visual perception and machine learning issues. Finally, we showed that the mediator based interfaces were judged as easier to use than the a priori more natural gestures based interface. This work was presented in [29].

6.4.3. Language Acquisition as a Particular Case of Context-Dependant Motor Skills Acquisition
Participants: Thomas Cederborg, Pierre-Yves Oudeyer.

Imitation learning, or robot programing by demonstration, have made important advances in recent years. We have proposed to extend the usual contexts investigated to also include linguistic expressions. We have proposed a modification to existing algorithms within the imitation learning framework so that they can handle learning from the demonstration of several unlabelled tasks (or motor primitives) without having to inform the imitator of what task is being demonstrated or what the number of tasks is, which then allows directly for relatively complex language learning. A mechanism for detecting weather or not linguistic/speech input is relevant to the task has also been proposed. With these additions it becomes possible to build an imitator that bridges the gap between imitation learning and language learning by being able to learn linguistic expressions using methods from the imitation learning community. In this sense the imitator learns a word by knowing that a certain speech pattern present in the context means that a specific task is to be executed. The imitator is
however not assumed to know that speech is relevant and has to figure this out on its own by looking at the demonstrations. To demonstrate this ability to find the relevance of speech non-linguistic tasks are learnt along with linguistic tasks and the imitator has to figure out when speech is relevant (in some tasks speech should be completely ignored and in other tasks the entire policy is determined by speech). A simulated experiment demonstrates that an imitator can indeed find the number of tasks it has been demonstrated, discover what demonstrations are of what task, for which of the tasks speech is relevant and successfully reproduce those tasks. This work is presented in a publication under review.

6.4.4. Robot Learning by Imitation of Internal Cognitive Operations in the Context of Language Acquisition

Participants: Thomas Cederborg, Pierre-Yves Oudeyer.

We have examined the problem of learning socio-linguistic skills through imitation when those skills involve both observable motor patterns and internal unobservable cognitive operations. This approach is framed in a research program trying to investigate novel links between context-dependent motor learning by imitation and language acquisition. More precisely, the paper presents an algorithm for learning how to respond to communicative/linguistic actions of one human, called an interactant, by observing how another human, called a demonstrator, responds. The response of the demonstrator, which depends on the context, including the signs of the interactant, is assumed to be appropriate and the robotic imitator uses these observations to build a general policy of how to respond to interactant actions. In this paper the communicative actions of the interactant are hand signs, and the learnt behavior consists of how to respond to the hand signs of a small and simple sign language, both in terms of adequately focusing attention on the right part of the scene, and in terms of responding physically. As a response to two continuous signs of the interactant, the demonstrator focuses on one out of three objects, and then performs a movement in relation to the object focused on. An algorithm is proposed based on a similarity metric between demonstrations, and a simulated experiment is presented where the unseen “focus on object” operation and the hand movements are successfully imitated, including in situations where there are no demonstrations. This work has been published in [21].

6.4.5. Learning Simultaneously New Tasks and Feedback Models in Socially Guided Robot Learning

Participants: Manuel Lopes, Thomas Cederborg, Pierre-Yves Oudeyer.

We have developed a system that allows a robot to learn simultaneously new tasks and feedback models from ambiguous feedback in the context of robot learning by imitation. We have considered an inverse reinforcement learner that receives feedback from a user with an unknown and noisy protocol. The system needs to estimate simultaneously what the task is, and how the user is providing the feedback. We have further explored the problem of ambiguous protocols by considering that the words used by the teacher have an unknown relation with the action and meaning expected by the robot. This allows the system to start with a set of known symbols and learn the meaning of new ones. We have presented computational results that show that it is possible to learn the task under a noisy and ambiguous feedback. Using an active learning approach, the system is able to reduce the length of the training period. [24], [26].

6.5. Hardware

6.5.1. Ergo-Robots/FLOWERS Fields: Towards Large-Scale Robot Learning Experiments in the Real World

Participants: Jérome Béchu, Fabien Bénureau, Haylee Fogg, Paul Fudal, Hugo Gimbert, Matthieu Lapeyre, Olivier Ly, Olivier Mangin, Pierre Rouanet, Pierre-Yves Oudeyer.
In the context of its participation to the exhibition “Mathematics: A Beautiful Elsewhere” at Fondation Cartier pour l’Art Contemporain in Paris, starting from 19th October 2011 and to be held until 18th March 2012, the team has elaborated and experimented a robotic experimental set-up called “Ergo-Robots/FLOWERS Fields”. This set-up is not only a way to share our scientific investigations with the general public, but attacks a very important technological challenge impacting the science of developmental robotics: How to design a robot learning experiment that can run continuously and autonomously for several months? Indeed, developmental robotics takes life-long learning and development as one of its central objective and object of study, and thus shall require experimental setups that allow robots to run, learn and develop for extended periods of time. Yet, in practice, this has not been possible so far due to the unavailability of platforms adapted at the same time to learning, exploration, easy and versatile reconfiguration, and extended time of experimentation. Most experiments so far in the field have a duration ranging from a few minutes to a few hours. This is an important obstacle for the progress of developmental robotics, which would need experimental set-ups capable of running for several months. This is exactly the challenge explored by the Ergo-Robots installation, which we have approached by using new generations of affordable yet sophisticated and powerful off-the-shelf servomotors (RX Series from Robotis) combined with an adequately designed software and hardware architecture, as well as processes for streamlined maintenance. The experiment is now running for five months, six days a week, in a public exhibition which has strong constraints over periods of functioning and no continual presence of dedicated technicians/engineers on site. The experiment involved five robots, each with 6 degrees of freedoms, which are endowed with curiosity-driven learning mechanisms allowing them to explore and learn how to manipulate physical objects around them as well as to discover and explore vocal interactions with humans/the visitors. The robots are also playing language games allowing them to invent their own linguistic conventions. A battery of measures has been set up in order to study the evolution of the platform, with the aim of using the results (to be described in an article) as a reference for building future robot learning experiments on extended periods of time, both within the team and in the developmental robotics community. More information available at: http://flowers.inria.fr/ergo-robots.php and http://fondation.cartier.com/.

Figure 15. Installation of Ergo Robots at Foundation Cartier.
6. New Results

6.1. Note

Note that we do not include here the results from Souhila Kaci and Tjitze Rienstra since they joined the team in September 2011.

6.2. Ontological Query Answering with Rules

**Participants:** Jean-François Baget, Marie-Laure Mugnier, Michaël Thomazo, Michel Leclère, Eric Salvat, Mélanie König.

**In collaboration with:** Sebastian Rudolph (Karlsruhe Institute of Technology)

We have developed a framework based on rules that have the ability of generating unknown individuals, an ability sometimes called value invention in databases. These rules are of the form $\text{body} \rightarrow \text{head}$, where the body and the head are conjunctions of atoms (without function symbols except constants) and variables that occur only in the head are existentially quantified, hence their name existential rules hereafter. E.g., $\forall x (\text{Human}(x) \rightarrow \exists y (\text{isParent}(y, x) \land \text{Human}(y)))$. These rules can be seen as the logical translation of conceptual graph rules, historically a main focus of the team [70], [55]. Existential rules have the same logical form as the well-known Tuple-Generating Dependencies (TGDs) in databases [45]. TGDs have been extensively used as a high-level generalization of different kinds of constraints, e.g., for data exchange [57]. Recently, there has been renewed interest for TGDs seen as rules in the context of ontological query answering. Indeed, the value invention feature has been recognized as crucial in an open-world perspective, where it cannot be assumed that all individuals are known in advance. The deductive database language Datalog allows to express some ontological knowledge but it does not allow for value invention. This motivated the recent extension of Datalog to TGDs (i.e., existential rules), which gave rise to the Datalog +/- family [52], [53], [54]. In KRR and in the Semantic Web, ontological knowledge is often represented with formalisms based on description logics (DLs). However, DLs traditionally focused on reasoning tasks about the ontology itself (the so-called TBox), for instance classifying concepts; querying tasks were restricted to ground atom entailment. Conjunctive query answering with classical DLs has appeared to be extremely complex, hence less expressive DLs more adapted to conjunctive query answering on large amounts of data have been designed recently, namely DL-Lite [51], $\mathcal{EL}$ [41], [63], and more generally Horn DLs (see e.g., [60]), cf. also the tractable profiles of the Semantic Web language OWL2. Existential rules cover the core of lightweight DLs dedicated to query answering, while being more powerful and flexible [53], [44], [21]. In particular, they have unrestricted predicate arity (while DLs consider unary and binary predicates only), which allows for a natural coupling with database schemas, in which relations may have any arity; moreover, adding pieces of information, for instance to take contextual knowledge into account, is made easy by the unrestricted predicate arity, since these pieces can be added as new predicate arguments.

Building on our previous work on conceptual graphs, while meeting this new trend, we have developed a knowledge representation framework centered on existential rules, which can be seen both as logic-based and graph-based.

Entailment, hence query answering, with existential rules is not decidable, thus finding decidable classes of rules as expressive as possible is a crucial issue. We have pursued our previous work on better understanding the border between decidability and undecidability. We have also extended rule dependency to k-dependency, which takes into account sequences of rule applications.

▷ Results published in Artificial Intelligence Journal [13] (extending the work in [3], [44]); keynote talk synthesizing this work at RR'2011 [20]; extension to k-dependency at RR'2011 [22]
For newly exhibited decidable classes (namely, “frontier-one”, “frontier-guarded” and “weakly-frontier-guarded” rules), the problem complexity was unknown, moreover there was no algorithm for computing entailment. First, we have classified these classes with respect to combined complexity (i.e., usual complexity) with both unbounded and bounded predicate arity, and data complexity (i.e., restricting the input of the decision problem to the facts). An interesting result is that some of the new classes (namely frontier-one and frontier-guarded rules) have a polynomial time data complexity. Secondly, we have provided a generic algorithm for query entailment with a large class of rules including these classes, which is worst-case optimal for combined complexity (with or without bounded predicate arity) as well as for data complexity.

▷ Results partially published at IJCAI’2011 [21]. Long paper in preparation with extended complexity results and all proofs, for submission to a major artificial intelligence journal.

6.3. Processing Conjunctive Queries with Negation

Participants: Marie-Laure Mugnier, Michel Leclère, Khalil Ben Mohamed, Michaël Thomazo.

Conjunctive queries have long been recognized as the basic queries in database and knowledge-based systems. The fundamental decision problems on these queries, namely query inclusion checking (given two queries \( q_1 \) and \( q_2 \), is \( q_1 \) included in \( q_2 \), i.e., is the set of answers to \( q_1 \) included in the set of answers to \( q_2 \) for all databases) and query entailment (is a given query entailed by the database) are NP-complete. When atomic negation is added to queries and databases, these problems become \( \Pi^P_2 \)-complete (with the open world assumption for the query entailment problem). Note that these problems can be recast as entailment in the FOL fragment of existentially closed conjunction of literals (without function symbols except constants). On the one hand, we have led a theoretical complexity study: we have investigated the role of pairs of literals called “exchangeable” (which generalizes the notion of unifiable literals) in the complexity increase. The main results are that when the number of exchangeable pairs is bounded, say by \( k \), then the complexity falls from \( \Pi^P_2 \)-complete to \( P^{NP} \)-complete for any \( k \geq 3 \), and is NP-complete for \( k \leq 1 \) (the case \( k = 2 \) being open).

In collaboration with: Geneviève Simonet (LIRMM Algeco team)

▷ Research Report [65]. Long version accepted at Information and Computation.

On the other hand, we have proposed, refined and compared experimentally several algorithms. This study follows first results of us in [61] and is the core of Khalil Ben Mohamed’s PhD thesis defended in December 2010 [64].

▷ Results published DEXA 2011 [24] (extending our work in RFIA 2010 [48], DEXA 2010 [46], AIMSA 2010 [47]).

Let us point out that both theoretical and practical results still hold when the predicates are preordered, which allows to take very light ontologies into account, i.e., where concepts and relations are organized in a specialization preorder.

6.4. Argumentation Systems for Decision Making

Participants: Rallou Thomopoulos, Madalina Croitoru, Jérôme Fortin, Marie-Laure Mugnier.

In collaboration with: Joël Abecassis (IATE/INRA), Jean-Rémi Bourguet (UM3), Patrice Buche (IATE/INRA), Sébastien Destercke (IATE/CIRAD) Nir Oren (Univ. of Aberdeen, Scotland)

Scientific investigations in this axis are guided by applications of our partners in agronomy (IATE laboratory). Substantial part of the work has consisted of analyzing the proposed applications and the techniques they require in order to select appropriate applications with respect to our team project.
Argumentation is a reasoning model based on the construction and the evaluation of arguments. In his seminal paper, Dung has proposed an abstract argumentation framework [56]. In that framework, arguments are assumed to have the same strength. This assumption is unfortunately strong and often unsatisfied. Consequently, several generalizations of the framework have been proposed in the literature. In [49] and [50], we have led a comparative study of these generalizations. It clearly shows under which conditions two proposals are equivalent. We have also integrated those generalizations into a common more expressive framework.

An instantiation of Dung’s abstract framework with the conceptual graph framework has been proposed. This representation uses default conceptual graph rules, an extension of classical conceptual graph rules (equivalent to existential rules, see Axis 1) with Reiter’s defaults [67] allowing for non-monotonic reasoning, that we developed independently of the argumentation framework [42], [43]. In the conceptual graph representation, arguments are represented as nested graphs, attacks between arguments can be computed from the structure of arguments and default rules allow to compute several kinds of extensions (i.e., maximal sets of arguments jointly acceptable according to a given semantics).

This approach has been applied to agrifood chain analysis, which is a highly complex procedure since it relies on numerous criteria of various types: environmental, economical, functional, sanitary, etc. Quality objectives imply different stakeholders, technicians, managers, professional organizations, end-users, public organizations, etc. Since the goals of the implied stakeholders may be divergent, decision-making raises arbitration issues. Arbitration can be done through a compromise—a solution that satisfies, at least partially, all the actors—or favor some of the actors, depending on the decision-maker’s priorities. We have analyzed a case study concerning risks/benefits within the wheat-to-bread chain. It concerns the controversy about the possible change in the ash content of the flour used for commonly used French bread. Several stakeholders of the chain are concerned, in particular the Ministry of Health through its recommendations in a national nutrition and health program, millers, bakers and consumers.

As already pointed out, the proposed approach is novel both for theoretical and application aspects.

▷ Results presented in [30], [28].

Let us mention additional results related to the applications in agronomy on decision making combining machine learning based on decision trees and ontologies [58], [30], as well as results obtained by our collaborators on semi-automatic data extraction from web data (tables), data reliability, and the representation and flexible querying of imprecise data with fuzzy sets [16], [14], [17], [26], [31], [25], [27], [33], [34]. These investigations are complementary to the above mentioned results on argumentation and generally relate to other aspects in the same applicative projects.

6.5. Semantic Data Integration

Participants: Michel Leclère, Michel Chein, Madalina Croitoru, Rallou Thomopoulos, Léa Guizol.

It often happens that different references (i.e. data descriptions), possibly coming from heterogeneous data sources, concern the same real world entity. In such cases, it is necessary: (i) to detect whether different data descriptions really refer to the same real world entity and (ii) to fuse them into a unique representation. Since the seminal paper [66], this issue has been been studied under various names: “record linking”, “entity resolution”, “reference resolution”, “de-duplication”, “object identification”, “data reconciliation”, etc., mostly in databases (cf. the bibliography by William E. Winckler 1). It has become one of the major challenges in the Web of Data, where the objective is to link data published on the web and to process them as a single distributed database. Most entity resolution methods are based on classification techniques; Fatiha Saïs, Nathalie Pernelle and Marie-Christine Rousset proposed the first logical approach [68]. Many experiments on public data are underway, in France (cf. DataLift2 and ISIDORE3 projects) or internationally (e.g., VIAF project4 led by OCLC5, whose aim is to interconnect authority files coming from 18 national organizations).

2 DataLift, http://datalift.org/
3 ISIDORE, http://www.rechercheisidore.fr/
4 The Virtual International Authority File, http://www.oclc.org/research/activities/viaf/
Two years ago, we began a collaboration with ABES (National Bibliographic Agency for Universities, which takes part in the VIAF project). The aim of this collaboration is to enable the publication of ABES metadata bases on the Web of Data and to provide an identification service dedicated to bibliographic notices. ABES bibliographic bases, and more generally document metadata bases, appear to be a privileged application domain for the representation and reasoning formalisms developed by the team. This work has an interdisciplinary dimension, as it also requires experts in the Library and Information Science domain. We think that a logical approach is able to provide a generic solution for entity resolution in document metadata bases, even though it is generally admitted in Library and Information Science that “there is no single paradigmatic author name disambiguation task—each bibliographic database, each digital library, and each collection of publications, has its own unique set of problems and issues” [69].

6.5.1. SUDOC Metadata Formalization

The first step of collaboration with ABES was to formalize the SUDOC catalogue, which contains all French academic libraries bibliographic notices, into a knowledge base using a suitable knowledge representation and reasoning language. This required to first analyze SUDOC content, as well as document description standards (CRM-CIDOC, FRBR, Dublin Core). We then designed an ontology expressed in the Semantic Web languages RDFS + OWL, compatible with document description standards, as well as translations from any SUDOC set of notices into a set of RDF facts according to this ontology. These translations have been implemented, which allows to export SUDOC bases into Semantic Web formats. Moreover, using the RDFS to CG second translation mentioned above, we are now able to import SUDOC bases into our tools CoGUI + CoGITaNT.

▷ Technical report [40].

6.5.2. Implementation of an Entity Identification Service

In order to perform entity resolution (for entities restricted to "authors" for now), we have defined a set of rules allowing to enrich Sudoc descriptions; then, using enriched descriptions, authors can be classified according to a proximity criterion. A prototype providing this service has been implemented on top of Cogui. Experiments are currently led in the context of the SudocAd project jointly conducted by ABES and GraphIK. SudocAd aims at enriching the author field of a bibliographic record describing a document with links to Sudoc authorities referring to the authors of the target document. A general description of the implemented approach, an analysis of this approach on a representative sample of bibliographic records and first results on 13400 bibliographic records extracted from a corpus independent from Sudoc catalog are presented in the final report of SudocAd.

▷ Link to SudocAd Final Report: http://www.abes.fr/Media/Fichiers/Sudoc-Fichiers/SudocAD_rapportFinal

Finally, we have defined an extension of our own logical framework (existential rules, constraints, homomorphism-based mechanisms) based on Hector J. Levesque and Gerhard Lakemeyer’s Standard Names [62], and the notion of knowledge base faithfulness with respect to the entity resolution problem (intuitively, the fact that the knowledge base is non-ambiguous). This is still ongoing work.

▷ Research Report [38].
6. New Results

6.1. High speed autonomous driving on roads

Participants: Fawzi Nashashibi, Laurent Bouraoui, Paulo Lopes Resende.

In this exceptional year, IMARA-INRIA had a very busy agenda with several scheduled demonstrations in the frameworks of several European projects, especially HAVEit, CityMobil and CityNetMobil. The HAVEit project final event took place at Boras (Sweden) in June. IMARA was tightly involved in the development of the Joint Demonstrator called FASCar-II. In this demonstrator, IMARA was responsible for the development of the Co-Pilot system. This is the main decisional system that handles the manoeuvres to be executed taking into account the multisensor data fusion sub-system, the driver monitoring system and the infrastructure (data provided by the infrastructure-to-vehicle telecommunications). The system also generates the trajectories to be executed as well the control-command laws to be send to the actuators [40] and provides passive or active assistance to the driver according to the active automation level.

In parallel, IMARA has also developed the “Wireless Infrastructure to Vehicle communication system”. Here, a specific hardware and software systems have been developed to allow V2I and V2V applications integrated on the Joint System demonstrator. This architecture – called the 4GCube – is based on the wireless communications devices and standards (802.11 a/b/g); they are IPv6 / CALM compliant architectures that have been tested with 2.4 GHz and 5 GHz bands, allowing multiple services handling.

The integration and the validation of the Co-pilot system as well as the communication device were done during the Final Event of HAVEit with a tremendous success and with high professional standards. The FASCar was able to demonstrate high driving autonomous skills at high speeds (up to 120 km/h) and was able to demonstrate new advanced features like overtaking mobile and static obstacles.

6.2. New urban transportation platforms: INRIA’s Cybus

Participants: Laurent Bouraoui, François Charlot, Carlos Holguin, Fawzi Nashashibi, Tony Noël, Michel Parent, Paulo Lopes Resende, Jianping Xie, Armand Yvet.
In order to achieve autonomous driving, autonomous systems (robots, intelligent vehicles, UAV’s, UGV’s,...) must have a decisional system that integrates an advanced perception system that performs sensors data fusion and environment modeling. From perception to control, task planning and path planning algorithms have to plan safe and optimized itineraries while processing sensory data. Motion control is the last link of the processing chain where itineraries and dynamic trajectories are executed by the low level control system. IMARA works on each of the topics mentioned above.

With the European projects CityMobil and CityNet Mobil, IMARA had the opportunity to validate the autonomous driving architecture developed through the projects showcases held in European cities. In 2011, several cities hosted IMARA team in charge of demonstrating autonomous driving and autonomous sensor-based navigation using sensory data (laser scanners) and GNSS (GPS). These events were a total success and were the opportunity to deploy and test the large-scale SLAMMOT system used for environment mapping, vehicle localization and mobile obstacles detection but also the new VMS (Vehicle Management System) developed in order to coordinate the mobility and navigation of several Cybercars (INRIA’s Cybus platforms).

For this purpose, the perception sub-system was based on the generic SLAM-based system that was already presented in detail last year [43]. The proposed localization architecture has been implemented in two different vehicle platforms. AGV is a fully autonomous vehicle equipped with two IBEO Alasca-XT laser scanners (left and right front corner). Cycab is a prototype of smart car mounted with a single IBEO-ML laser scanner in front.

This year, because of delivery problems related to an industrial partner, IMARA had to design and develop its own cybercars. Thus the Cybus are the newest prototyping and demonstration platforms designed at INRIA. Apart from the chassis and engines, the whole hardware and software systems were developed thanks to IMARA’s researchers and engineers talents. These electric vehicles are based on a Yamaha chassis but the embedded intelligence is the result of this year’s IMARA developments. Much of the perception and control software is now registered.

The system developed can be seen as an experimental platform for a new public mobility transportation system operating in mixed environments. This electric vehicle is a “clean” transportation mean that is capable to achieve the well known last-mile itinerary in urban areas where classical transportation means are inefficient or simply non profitable. In order to demonstrate the feasibility of the system, a 3-months service has been programmed in the City of La Rochelle. The Mayor of the City authorized the Cybus to operate during 3-months providing free transport service to the inhabitants of the city. For that purpose, a 2 kilometers route has been defined on which 5 calling stations where installed. The users were able to call the Cybus from any station to reach any other station. The evaluation of the system has shown an acceptability of more than 95% among population; very low failures or technical leakages were reported. Following this successful operation,
the European Commission asked INRIA to extend this operation as well as its technical capabilities in order to achieve a simultaneous multiple Cybus navigation. This was achieved last December and was a real success.

Following this success, a new proposal of CityMobil-2 project is under submission with the objective this time to extend real operational mobility services to 6-12 months in selected European cities!

6.3. Communications Management in Cooperative Intelligent Transportation Systems

Participants: Thierry Ernst, Manabu Tsukada.

Cooperative Intelligent Transportation Systems (Cooperative ITS) are systems where the vehicles, the roadside infrastructure, central control centers and other entities exchange information in order to achieve better road safety, traffic efficiency and comfort of the road users. This exchange of information must rely on a common communication architecture. The ITS Station reference architecture has thus been specified in ISO and ETSI. It allows vehicles and roadside ITS stations to organize themselves into Vehicular Ad-hoc Network (VANET), presumably though IPv6 GeoNetworking using IEEE802.11p and to connect seamlessly to the
Internet though any available access technology. Several paths may thus be available at a given vehicle ITS station to communicate with other ITS stations. Paths are of three types: direct path, optimize path and anchor path. The objective of the study is to optimize the communication between ITS Stations by selecting the best available communication path. This requires first to gather information available locally at the ITS station (position, speed, application requirements, media characteristics, capabilities, path status, ...) and collected from neighbors ITS stations (position, speed, services, ...) and then to process this information through a decision-making algorithm. First, we define a network module allowing the combination of IPv6 together with GeoNet-working. Second, we propose a cross-layer path selection management module. Our contributions are mapped to the ITS station reference architecture by defining the relation between the ITS station network and transport layer (which hosts our IPv6 GeoNetworking contribution) and the vertical ITS station cross-layer entity (which hosts the path decision-making algorithm). We specify the functions allowing the exchange of parameters through the Service Access Point (SAP) between the network layer and the management entity (MN-SAP). The parameters used at the cross layer ITS station management entity are abstracted in a way so that they are agnostic to the protocols used at the ITS station network and transport layer, therefrom allowing easy replacement of protocol elements (e.g. replacing NEMO by other mobility support protocol) or permutation of the network stack (IPv6 or GeoNetworking, a combination of both or other network stack).

6.4. Managing the system (via probabilistic modeling)

6.4.1. Belief propagation inference for traffic prediction

Participants: Cyril Furtlehner, Jean-Marc Lasgouttes, Arnaud Lewden, Victorin Martin.

This work [41] deals with real-time prediction of traffic conditions in a setting where the only available information is floating car data (FCD) sent by probe vehicles. The main focus is on finding a good way to encode some coarse information (typically whether traffic on a segment is fluid or congested), and to decode it in the form of real-time traffic reconstruction and prediction. Our approach relies in particular on the belief propagation algorithm.

These studies are done in particular in the framework of the projects TRAVESTI and Pumas.

This year’s highlights are

- A particular effort has been done this year in studying the theoretical aspects of the ways to encode real valued variable into an binary Ising model. A publication on the subject is in preparation.
- A review of our work on road traffic inference using methods from statistical physics has been published [21].
- The investigation of the effect of various types of normalization in the belief propagation algorithm has lead to a technical report [38].
- Arnaud Lewden has specified and implemented the new software BPstruction, which is our contribution to the Pumas project. Besides implementing traffic reconstruction from FCD, it is intended as a testbench for our research on inference using Belief Propagation.
- Victorin Martin has given a talk at the “Séminaire de Modélisation des Réseaux de Transport” at IFSTTAR. He presented there our method for real-time traffic reconstruction and prediction.
- Jean-Marc Lasgouttes also presented this work at the Xerox Research Centre Europe seminar.

6.4.2. Evaluation of dual mode transport system by event-driven simulation

Participants: Arnaud de La Fortelle, Sami Mahari.

The European project CATS — City Alternative Transport System — is developing and evaluating a new vehicle system using a single type of vehicle for two different usages: individual use or collective transport. Real experiments will necessarily take place with a limited number of vehicles and stations. Hence there is a need for evaluation using simulations. We have been developing a discrete events simulator for that purpose, based on a previous work done for collective taxis [42].
Our model relies on an adapted events/decision graph that extends previous graphs. The new feature of this model is the way we deal with two modes that can be extended to many other modes. This work therefore shows on a concrete example a method to efficiently merge multiple modes into one model.

- This year has seen the design and first implementation of the simulator.
- The results have been presented at a conference [29].

6.4.3. Multi-speed exclusion processes

Participants: Cyril Furtlehner, Jean-Marc Lasgouttes, Maxim Samsonov.

The slow-to-start mechanism is known to play an important role in the particular shape of the Fundamental diagram of traffic and to be associated to hysteresis effects of traffic flow. We study this question in the context of stochastic processes, namely exclusion and queueing processes, by including explicitly an asymmetry between deceleration and acceleration in their formulation. Spatial condensation phenomena and metastability are observed, depending on the level of the aforementioned asymmetry. The relationship between these 2 families of models is analyzed on the ring geometry, to yield a large deviation formulation of the fundamental diagram (FD)

This work has been presented at the TGF’11 conference [22], and a more extensive article is in preparation for a journal.

6.4.4. Dynamics of points of interest in a social game

Participants: Guy Fayolle, Jean-Marc Lasgouttes.

Ma Micro Planète is a geolocalized video game which entices players to use sustainable means of transport. At the heart of the game are community-driven points of interest (POI’s), or sites, which have a score that depends on the players activity. The aim of this work is to understand the dynamics of the underlying stochastic process.

We examine the system in the thermodynamic limit, as the number of players tends to infinity, the existence of which is proved under general conditions, where the probability of increasing the score of a visited POI is a function of the state of the system. Concerning the existence of a stationary regime, some complete answers are given for particular values of the parameters, and the existence of possible phase transition phenomena is enlightened.

A publication on the subject is in preparation.

6.4.5. Random walks in the quarter plane

Participant: Guy Fayolle.

In collaboration with K. Raschel (CNRS, Université F. Rabelais à Tours), we pursued the works initiated in 2010 in two main directions.

6.4.5.1. The zero drift case

In several recent studies on random walks with small jumps in the quarter plane, it has been noticed that the so-called group of the walk governs the behavior of a number of quantities, in particular through its order. In the article [11], when the drift of the random walk is equal to 0, we provide an effective criterion giving the order of this group. More generally, we also show that in all cases where the genus of the algebraic curve defined by the kernel is 0, the group is infinite, except precisely for the zero drift case, where finiteness is quite possible.

6.4.5.2. Counting and asymptotics

The enumeration of planar lattice walks, is a classical topic in combinatorics. For a given set $S$ of allowed unit jumps (or steps), it is a matter of counting the number of paths starting from some point and ending at some arbitrary point in a given time, and possibly restricted to some regions of the plane.

Like in the probabilistic context, a common way of attacking these problems relies on the following analytic approach. Let $f(i, j, k)$ denote the number of paths in $\mathbb{Z}^2_+$ starting from $(0, 0)$ and ending at $(i, j)$ at time $k$.

Then the corresponding CGF
\[ F(x, y, z) = \sum_{i,j,k \geq 0} f(i, j, k)x^i y^j z^k \]
satisfies the functional equation

\[ K(x, y)F(x, y, z) = c(x)F(x, 0, z) + c(y)F(0, y, z) + c_0(x, y), \]

where \( x, y, z \) are complex variables, although the time variable \( z \) plays somehow the role of a parameter. The question of the type of the associated counting generating functions, that is rational, algebraic, holonomic (solution of a linear differential equation with polynomial coefficients), was solved whenever the group is finite (see RA 2010). When the group is infinite, the problem is still largely.

It turns out that the nature of singularities play a deep important role in this classification. Making use of the general and powerful approach proposed in the book [3], a paper entitled Some exact asymptotics in the counting of walks in the quarter-plane has been submitted to AofA (International Conference on Analysis of Algorithms, Montreal, June 2012), in which a new approach is proposed to obtain some exact asymptotics for walks confined to the quarter plane.

### 6.4.6. Statistical physics and hydrodynamic limits

**Participants:** Guy Fayolle, Cyril Furtlehner.

Having in mind a global project concerning the analysis of complex systems, we first focus on the interplay between discrete and continuous description: in some cases, this recurrent question can be addressed quite rigorously via probabilistic methods.

To attack this class of problems, in touch with many applications domains (e.g. biology, telecommunications, transportation systems), we started from paradigmatic elements, namely the discrete curves subjected to stochastic deformations, as those mentioned for instance in [39]. After convenient mappings, it appears that most models can be set in terms of interacting exclusion processes, the ultimate goal being to derive hydrodynamic limits for these systems after proper scalings. We extend the key ideas of [39], where the basic ASEP system on the torus was the toy model. The usual sequence of empirical measures, converges in probability to a deterministic measure, which is the unique weak solution of a Cauchy problem.

The Gordian knot is the analysis of a family of specific partial differential operators in infinite dimension. Indeed, the values of functions at given points play here the role of usual variables, their number becoming infinite. The method presents some new theoretical features, involving promeasures (as introduced by Bourbaki), variational calculus, functional integration, and the construction of generalized measures. In [20], we present a detailed analysis of the ASEP system on the torus \( \mathbb{Z}/N\mathbb{Z} \). Then we claim that most of the arguments a priori work in higher dimensions (ABC, multi-type exclusion processes, etc), leading to systems of coupled partial differential equations of Burgers’ type. In the course of the study, several fascinating multi-scale problems emerge quite naturally, bringing to light some connections with the so-called renormalization in theoretical physics.
6. New Results

6.1. Feature space modelling

6.1.1. A novel shape boundary based description for leaf identification

**Participants:** Itheri Yahiaoui, Olfa Mzoughi, Nozha Boujemaa.

The problem of automatic leaf identification is particularly difficult for two main reasons: (i) the first is the enormous number of leaf species and (ii) the second, which is relevant for some special species but more complex, is the high inter-species and the low intra-species similarity.

Our research has focused on analysing leaf morphology in order to determine a numeric key description for leaf species robust to all the above mentioned constraints. The approach that we propose is a shape boundary description that combines two complementary information: (i) the first one outlines local variations of the leaf margin. This is performed using the Directional Fragment Histogram (DFH), introduced in [25], which encodes the relative frequency distribution of groups of contour points with uniform orientation, (ii) the second property emphasizes the spatial distribution of contour points (in terms of distances). This is done by comparing the shape to standard geometric ones (such as circle, rectangle, ellipse, convex hull, etc.).

This descriptor was evaluated within the framework of ImageCLEF 2011 plant task where a crowd-sourced database, called Pl@ntLeaves [13], was used and a high number of image retrieval techniques was tested (a total of 8 groups from all around the world that have submitted 20 runs [19]). Our descriptor brought the best rate for scan-like pictures and was close to the best rate for scan pictures. Besides to the accuracy, this descriptor requires very low computational time, which accomplishes a basic condition for real world application.

6.1.2. Visual-based plant species identification from crowdsourced data

**Participants:** Hervé Goëau, Alexis Joly, Souheil Selmi, Laurent Joyeux.

Inspired by citizen sciences, the main goal of this work is to speed up the collection and integration of raw botanical observation data, while providing to potential users an easy and efficient access to this botanical knowledge. We therefore designed and developed an original crowdsourcing web application dedicated to the access of botanical knowledge through automated identification of plant species by visual content.

Technically, the first side of the application deals with content-based identification of plant leaves. Whereas state-of-the-art methods addressing this objective are mostly based on leaf segmentation and boundary shape features, we developed a new approach based on local features and large-scale matching. This approach obtained the best results within one sub-task of ImageCLEF 2011 plant identification benchmark [19]. The second side of the application deals with interactive tagging and allow any user to validate or correct the automatic determinations returned by the system.

Overall, this collaborative system allows to enrich automatically and continuously the visual botanical knowledge and therefore to increase progressively the accuracy of the automated identification. A demo of the developed application was presented at ACM Multimedia conference [13]. This work was done in collaboration with INRIA team ZENITH and with the botanists of the AMAP UMR team (CIRAD). It is also closely related to a citizen science project around plant’s identification that we developed with the support of the TelaBotanica social network inside the Pl@ntNet project.

6.1.3. Spatial relations between salient points on a leaf

**Participants:** Sofiène Mouine, Itheri Yahiaoui, Anne Verroust-Blondet.
Figure 1. Normalized classification scores for scan images (up) and scan-like images (down). Our approach bin is in red, other ImageCLEF methods in blue.
In the scope of the Pl@ntNet project, our recent work has consisted in finding spatial relationships between salient points on a leaf. As a first step, classic detectors were used to find significant points in the leaf area and then the Shape context descriptor, originally applied on contour points, was introduced to measure a spatial relation between interest points. We have tested different configurations by varying the set of voting points. First results confirm that including spatial relations enriches the local description of each point. We are currently improving a veins and landmark extraction approach [12] in order to include also veins points in the voting set.

6.1.4. 3D mesh segmentation by example

Participants: Esma Elghoul, Anne Verroust-Blondet.

In recent years, there has been an increasing interest for automatic 3D segmentation. Indeed, segmentation of 3D objects is an important step in many applications such as part indexing of 3D objects, pattern recognition, compression, morphing, texture mapping and simplification. It refers to the process of partitioning 3D shapes into multiple parts, based on semantic criteria and/or geometric criteria.

Our work consists in introducing an approach to segment a 3D object class referring to a given segmented object from this class (we called it segmentation by example). The considered segmentation method is not automatic: we want to use interactive tools that proved advantageous to segment a 3D shape into relevant parts.

As a first task, we reviewed the state of the art in 3D segmentation techniques recently proposed in the literature. The different techniques were evaluated and classified for the purpose of choosing the more appropriate one for our work. We opted for extending the technique of random walks [24] to build an interactive tool of 3D segmentation.

For the second task, we had to solve a basic problem: that of similarly direct objects belonging to the same class. Indeed, each 3D model is provided in a random orientation in the space. In order to align objects of a same class, we used the alignment approach developed in [22] which computes 3 alignment axes. To properly orient our objects between them, we had to develop an additional process to the last one. It combines a 2D ICP and a 3D ICP approaches to give the best orientation among 48 possibilities and pair each two objects meshes.

So having a user-supplied already segmented model (model (1)) and a second model (model (2)) belonging to the same class (not segmented but similarly oriented), we developed a method to put into correspondence segmented parts of model (1) with faces of model (2). Then we computed a segmentation of model (2) using a derivative approach of the random walks. We applied this technique as well to segment all the objects that belong to the class of model (1). Our approach provides good results for manufactured object classes such as chairs and tables.

6.2. Feature space structuring

6.2.1. Random Maximum Margin Hashing

Participant: Alexis Joly.

Following the success of hashing methods for multidimensional indexing, more and more works are interested in embedding visual feature space in compact hash codes. Such approaches are not an alternative to using index structures but a complementary way to reduce both the memory usage and the distance computation cost. Several data dependent hash functions have notably been proposed to closely fit data distribution and provide better selectivity than usual random projections such as LSH. However, improvements occur only for relatively small hash code sizes up to 64 or 128 bits due to the lack of independence between the produced hash functions. In this work, we introduced a new hash function family that attempts to solve this issue in any kernel space. Rather than boosting the collision probability of close points, this method focus on data scattering. By training purely random splits of the data, regardless the closeness of the training samples, it is indeed possible to generate consistently more independent hash functions. On the other side, the use of large margin classifiers allows to maintain good generalization performances. Experiments did show that our
new Random Maximum Margin Hashing scheme (RMMH) outperforms four state-of-the-art hashing methods, notably in kernel spaces. Overall, this new concept of randomly trained classifiers opens the door to many other problems including large-scale learning, visual vocabulary construction or distributed content-based retrieval methods. A paper describing RMMH was published in the proceedings of CVPR 2011 [14].

6.2.2. Scalable information retrieval in distributed architectures

**Participants:** Mohamed Riadh Trad, Alexis Joly, Nozha Boujemaa.

Organizing media according to the occurrence of real-life events is attracting increasing interest in the multimedia community. However, whereas text based methods are now mature enough to deal with huge datasets, there are still some challenging issues managing multimedia contents. This becomes even more challenging in the context of User Generated Contents. Low-level visual metadata are indeed not simple textual or scalar values, their management requires efficient similarity search in high dimensional spaces.

Similarity search in high dimensional spaces has been the focus of many works in the database community in the recent years. State-of-the-art methods focus mainly on space partitioning techniques and more recently on hash-based probabilistic algorithms.

Although, hash-based approaches proved to be scalable, the computational cost is still too high for some real world applications and K-Nearest Neighbours Graph constructed can be more desirable than the costly online K-NN search. In fact, the basic LSH algorithm partitions the space uniformly and thus it does not exploit the clustering property of the data, which may result in slow query response and wasted space with additional hash tables. These limitations were pointed out with our scalable prototype for large scale event matching [18].

Scaling up LSH-based techniques and applications is then closely related to buckets occupations and objects distribution within the index structure. Recent works achieve better data distribution over the buckets with guarantees on occupation. As one result, we easily bound the similarity join size and evaluate bound algorithms complexity.

Based on these works, we designed and implemented a scalable prototype for distributed similarity search and K-NN graph construction. We have made several experiments querying real world large datasets. The prototype proved to be efficient for both search and K-NN graph construction.

Ongoing experiments process a 1.2 million images dataset. Results will be submitted for publication.

6.2.3. Visual similarity sensitive hashing methods for semantic image search in very large collections of images

**Participants:** Saloua Ouertani-Litayem, Alexis Joly, Nozha Boujemaa.

With the rapid development of information acquisition technology, we have witnessed an explosive growth in the scale of shared data collections. Then, it is now possible to tackle fundamental problems with very large datasets’ context. Especially those addressing challenging tasks in machine learning for developing large scale approaches for multimedia retrieval and mining. Computer Vision is experiencing this paradigm shift, with large annotated image and video datasets becoming available. Indeed, various benchmark datasets for image classification have been released such as image-net and LabelMe. Therefore, a key challenge is taken up through out the Phd aiming to build efficient methods for training and matching efficiently very large collections of images.

We proposed several SVM-based strategies to build new supervised hash function families from large annotated collections of features. We indeed investigated with an approach consisting in benefitting from different embedding approaches in order to build compact codes indexed with efficient similarity search structures. Therefore, we have extended a kernelized hashing method [14] with multi-class SVM to solve a K-class classification problem by choosing the maximum applied to the outputs of K SVMs. We indeed proposed hashing methods based on the multi-class SVM classification strategies: One vs One (OVO) And One vs All (OVA). An important task during this process was to experimentally evaluate the quality lost induced by such representations with respect to the efficiency gains. We then compared multi-class SVM strategies with different underlying kernels.
Inspired by state-of-the-art hashing in kernel space methods, we investigated an approach consisting of benefiting from both semantic hashing-like techniques and kernel embedding approach in order to build compact category-aware codes indexed with efficient similarity. Experiments are performed on the ImageNet ILSVRC 2010 dataset [23]. Results will be submitted for publication.

6.3. Pattern recognition and statistical learning

6.3.1. Machine identification of biological shapes

Participants: Asma Rejeb Sfar, Donald Geman, Nozha Boujemaa.

Stored images of biological objects are accumulating at a staggering rate due to new sensor technologies, expanding use in medical diagnostics, web-based search engines and growing demands for web-based services in traditional sciences such as botany. These developments have been accompanied by an increasing demand for the automated analysis of these data, such as counting cell types, detecting lesions and other abnormalities in medical images, and identifying botanical shapes.

All these tasks have one feature in common: massive diversity among the shapes. Indeed, such shapes display enormous within-class variation and are generally highly deformable. Also, they often exhibit a hierarchical organization resulting from evolutionary processes.

There is currently no existing methodology in image analysis and computer vision which can be applied to a multi-class shape recognition problem of this complexity. Consequently, there is a need for a new, generic methodology for categorizing hierarchically-structured families of deformable shapes, particularly when both the number of categories and the within-category variation are very large.

We proposed a coarse-to-fine (CTF) approach in both shape representation and image parsing. The representation is hierarchical in both class and pose.

We focused on botanical shapes, specifically categorizing simple leaves according to species. So, we determined a suitable representation for the pose of a simple leaf and designed and tested a two-stage pose detector. Then, we constructed classifiers based on the plant taxonomy.

Results will be submitted for publication.

6.4. Interactive search and personalisation

6.4.1. Database denoising and multi visual queries

Participant: Sébastien Poullot.

One of IMEDIA's tasks inside the SCARFACE project is to introduce and develop a character retrieval system. For this purpose, we take as entries the tracking of the persons in video sequences computed by Thalès and construct a database of the profiles. A profile is a 3D frame, a bounding box that changes along the timeline. Two original works have been proposed for searching in the profile database.

The first one consists in analysing features of each profile with respect to all the profiles in order to extract relevant features from it, and construct more representative databases.

The second one is to be able to search inside the database with a set of queries (pictures of the same person). An a priori work can be done on this set of queries in order to extract the relevant features (and remove the irrelevant ones). On the other side an a posteriori work can be done on late merging depending on the specificities of each sub query.

TRECVID Instance Search 2011

Before starting the developments for SCARFACE, we tested various algorithms in TRECVID 2011 INS (instance search) task. This task is close to the SCARFACE one: from a set of captures of one object, one should find its occurrences in a set of video sequences. This work has been done during the stay of Sébastien Poullot at NII (the Japan National Institute of Informatics) in July and August 2011.
The differences with SCARFACE are:

- a high diversity in the type of the objects (people but also, places, vehicle, animals, etc),
- the location of the object in the database is not given.

Our approach obtains good results (above the median scores of all teams) and works in a very short time (and without indexing system for speeding up the process) [15]. The choice for SCARFACE’s method partially depends on these results. We still continue on the INS task in order to achieve better scores (various descriptors and various post and late fusion between sub queries).

**Query generative models**

Moreover, in order to enhance visual query results, we want to create some visual query generative models. It is directly linked to SCARFACE (a priori processes) and TRECVID works: given a set of images (considered as queries), we extract what gather them and what separate them in order to construct artificial relevant queries. For now we essentially work on some logo databases.

6.4.2. **Object-based Visual Query Suggestion**

**Participants:** Amel Hamzaoui, Pierre Letessier, Alexis Joly, Nozha Boujemaa.

After our work on the shared neighbours clustering methods in multi-sources case published in [10], we are interested now to the case of a bipartite graph that we apply to object-based visual query suggestion using the visual words mining technique [16]. In fact, state-of-the-art visual search systems allow to retrieve efficiently small rigid objects in very large datasets. They are usually based on the query-by-window paradigm: a user selects any image region containing an object of interest and the system returns a ranked list of images that are likely to contain other instances of the query object. User’s perception of these tools is however affected by the fact that many submitted queries actually return nothing or only junk results (complex non-rigid objects, higher-level visual concepts, etc.). We address the problem of suggesting only the object’s queries that actually contain relevant matches in the dataset. This requires to first discover accurate object’s clusters in the dataset (as an off-line process); and then to select the most relevant objects according to user’s intent (as an online process). We therefore introduce a new object’s instances clustering framework based on two main contributions: efficient object’s seeds discovery with adaptive weighted sampling and bipartite shared-neighbours clustering. Experiments show that this new method outperforms state-of-the-art object mining and retrieval results on OxfordBuilding dataset. We finally describe two object-based visual query suggestion scenarios using the proposed framework and show examples of suggested object queries.

6.4.3. **Interpretable Visual Models for Human Perception-based Object Retrieval**

**Participants:** Ahmed Rebai, Alexis Joly, Nozha Boujemaa.

Understanding the results returned by automatic visual concept detectors is often a tricky task making users uncomfortable with these technologies. In this work we attempt to build humanly interpretable visual models, allowing the user to visually understand the underlying semantic. We therefore proposed a supervised multiple instance learning algorithm that selects as few as possible discriminant local features for a given object category. The method finds its roots in the lasso theory where a L1-regularization term is introduced in order to constraint the loss function, and subsequently produce sparser solutions. Efficient resolution of the lasso path is achieved through a boosting-like procedure inspired by Blasso algorithm. Quantitatively, the method achieved similar performance as current state-of-the-art, and qualitatively, it allows users to construct their own model from the original set of patches learned, thus allowing for more compound semantic queries. This work is part of the PhD of Ahmed Rebai [8] and it was published in ICMR 2011 proceedings [17]. This work was then extended to using geometrically checked feature sets rather than using single local features to describe the content of visual patches. We did show that this allows drastically reducing the number of the selected visual words while improving their interpretability. A publication was submitted to pattern recognition journal [11].
6.4.4. Relevance feedback on local features: Application to plants annotations and identification

Participants: Wajih Ouertani, Michel Crucianu, Nozha Boujemaa.

As biological image databases are increasing rapidly, automated species identification based on digital data is of great interest for accelerating biodiversity assessment, researches and monitoring. In this context, our work falls within an investigation of computer vision techniques or more precisely: object recognition and content based image retrieval techniques to help botanist identifying and organizing his digital images’ collections. Under believe that perception, recognition and decision are parts of human skills, this work focus on an interactive mechanism which tries to extract useful information from the user and gives him help to deal with large data amount. We adopted an explicit relevance feedback (RF) schema and we worked on extending it to deal with local intention through local features (LF) description. This mechanism helps discovering and dynamically defining new concept and interesting plant parts and feed identification ways interactively. Moreover since it relies to the content rather than labels one direct application is to fill the initially sparse annotation space with right annotations and in a reasonable time and with the introduce of one or many expertises. We recently explored and tested images local features matching involving high order features and non-rigid adaptation tentative to structure database with a patterns’ discovery stage. Using those type of methods we expect to introduce a high level appearance information that tends to go beyond classical bag of features and histogram based distances at least from semantic gap and interpretation point of view. We argue our exploration way with the fact that initial search space can be exceedingly rich. By pre-structuring it we can hope to obtain a smaller search space together with more reliable inference. Also learning parts interactively with localized local features may require a lot of interaction since it requires a considerable number of examples. We experienced the design of combined machine learning and prior mining of matches which we are actually improving.

6.5. Software

6.5.1. IKONA/MAESTRO software


This year, IKONA has been extended in the context of Pl@ntNet, Glocal, I-SEARCH and R2I projects. For the Pl@ntNet project, along the continuing improvements in the MAESTRO software, a number of new features were added. Namely the support for the automatic image segmentation and subsequent use of segmented regions; descriptors with the various shape’s geometric parameters; use of multiple orientations for Harris points; run-time additions to the external database and immediate availability of the new images for the search; descriptors to facilitate external data usage; colour SIFT and Affine Covariance descriptors; integration of the thesis work of Ahmed Rebai for objects retrieval; and tools for statistical tests.

In addition, a number of new web services were developed and deployed: the dynamic indexation system of the on-line pad (“carnet-en-ligne” of Tela Botanica)images; the search with multiple views; the update of Pl@ntNet internal demonstration allowing to present features such as visual similarity search, textual search, filtering (pre- and post-filtering), and different methods of research; the implementation of the organ prediction web service and other web services of botanical information statistics; the administration of the indexation system and the experimentation of new research methods (GPS spatial and temporal search).

For the Glocal project, an interface was developed for the demonstrations of a search engine in large scale events database (the queries are event images and the result is a list of the closest events in terms of time alignment and image content), and new web services were developed and updated according to the data exchange format and the middle-ware of the project - among others: fraud detection, import media from the web, associate media with existing event in the repository, and event matching web services. The queries are composed of either a medium link (an external image) or an event link (set of external images).
For the I-SEARCH project, an integration was performed to provide global and local 2D image low level descriptors. For videos, an automated extraction of visual words tool was integrated to show to users image patches which are the most meaningful.

For the R2I project, a detailed technical documentation of the procedure of maestro’s installation, web services and tomcat server were provided to Exalead partner.
6. New Results

6.1. Interaction Techniques

Participants: Caroline Appert, Michel Beaudouin-Lafon [correspondant], David Bonnet, Anastasia Bezerianos, Olivier Chapuis, Guillaume Faure, Emilien Ghomi, Stéphane Huot, Mathieu Nancel, Wendy Mackay, Cyprien Pindat, Emmanuel Pietriga, Theophanis Tsandilas, Julie Wagner.

Acquiring a target, such as pointing to an icon, a button or a landmark on a digital map, is the most common action in today’s graphical user interfaces. We have continued our work to better understand this seemingly simple action and make it faster and more reliable. This year we have conducted theoretical work on small target acquisition using a mouse and investigate three potential sources of problems: motor accuracy, legibility, and quantization. We find that quantization is not a problem, but both motor and visual sizes are limiting factors. This suggests that small targets should be magnified in both motor and visual space to facilitate pointing. Since performance degrades exponentially as targets get very small, we further advocate the exploration of uniform, target-agnostic magnification strategies. We also confirm Welford’s 1969 proposal that motor inaccuracy can be modeled by subtracting a “tremor constant” from target size. We argue for the adoption of this model, rather than Fitts’ law, when reflecting on small target acquisition.

With TorusDesktop, we revisited a pointing technique that allows to wrap the mouse cursor around screen edges in conventional desktop environments. Allowing the cursor to jump from one edge of the screen to the opposite one (i.e., turning the desktop into a torus) was already explored, but never studied empirically nor designed for everyday desktop usage. We have introduced a dead zone and an off-screen cursor feedback that ease the use of this technique. We also conducted three controlled experiments to refine the design and evaluate its performance. Our results suggest clear benefits in several conditions, but also some potential limitations due to users’ over-estimation of cursor wrapping advantages. An implementation of TorusDesktop for the Mac OS X desktop can be downloaded for free at http://insitu.lri.fr/TorusDesktop.

We continued our work on wall-sized displays, focusing on the study of high-level tasks such as pan-zoom navigation (Figure 10), that have received little attention. Indeed, while pointing on this type of display has been studied extensively, it remains unclear which techniques are best suited to perform multiscale navigation in these environments. Building upon empirical data gathered from studies of pan-and-zoom on desktop computers and studies of remote pointing, we identified three key factors for the design of mid-air pan-and-zoom techniques: uni- vs. bimanual interaction, linear vs. circular movements, and level of guidance to accomplish the gestures in mid-air. After an extensive phase of iterative design and pilot testing, we ran a controlled experiment aimed at better understanding the influence of these factors on task performance. This work received a best paper award at CHI 2011.

On the opposite side, we have studied small displays such as the ones smartphones are equipped with. One major challenge with this type of device is to make the user able to interact in parallel with both the device and other artefacts in his environment (e.g., giving a phone call while holding a paper document). The Swiss Army Menu (SAM) is a radial menu that enables a very large number of functions accessible via small thumb movements. The design of SAM relies on four different kinds of items, support for navigating in hierarchies of items and a control that requires only the thumb of the hand that holds the device. SAM can offer a set of functions so large that it would typically have required a number of widgets that could not have been displayed in a single viewport at the same time.
The different interaction techniques presented above are aimed at improving the control within a given representation. As a complement, we have also worked on improving the user’s experience by providing him with two (or more) representations of the data he is interacting with and ways to transition between these representations.

Gliimpse [17] is a quick preview technique that smoothly transitions between document markup code (HTML, wiki markup or LaTeX) and its visual rendering (see Figure 11). It allows users to regularly check the code they are editing in-place, without leaving the text editor. Gliimpse can complement classical preview windows by offering rapid overviews of code-to-document mappings and leaving more screen real-estate. As the technique smoothly show the links between the code and the rendered result, it can also help to learn how complex markup code will result in the final document (e.g., HTML tables or LaTeX formulae).

In collaboration with University of Toronto and OCAD University, we designed a novel visualization technique called ChronoLenses [14], aimed at supporting users in time-series visual exploration tasks. ChronoLenses perform on-the-fly transformation of the data points in their focus area, tightly integrating visual analysis with user actions, and enabling the progressive construction of advanced visual analysis pipelines, supporting tasks that require visualizing derived values, identifying correlations, or discovering anomalies beyond obvious outliers.

We further explored user understanding of data presented in Dual-Scale data charts, charts that incorporate two different data resolutions into one chart in order to emphasize data in regions of interest (focus) or to enable the comparison of data from distant regions (context) [13]. In collaboration with researchers from inria AVIZ, we presented a unified description of different Dual-Scale data charts, and we compared them in terms of user understanding using elementary graphical perception tasks, such as comparing lengths and distances. Our study suggests that cut-out charts which include collocated full context and focus are the best alternative, and that superimposed charts in which focus and context overlap on top of each other should be avoided.

6.2. Research Methods
Perception, Cognition, Interaction - New Results - Project-Team IN-SITU

Participants: Caroline Appert, Michel Beaudouin-Lafon, Anastasia Bezerianos, Olivier Chapuis, Jérémie Garcia, Stéphane Huot, Ilaria Liccardi, Wendy Mackay [correspondant], Emmanuel Pietriga.

InkSplore [19]. We conducted three studies with contemporary music composers at IRCAM. We found that even highly computer-literate composers use an iterative process that begins with expressing musical ideas on paper, followed by active parallel exploration on paper and in software, prior to final execution of their ideas as an original score. We conducted a participatory design study that focused on the creative exploration phase, to design tools that help composers better integrate their paper-based and electronic activities. We then developed InkSplorer as a technology probe that connects users? hand-written gestures on paper to Max/MSP and OpenMusic. Composers appropriated InkSplorer according to their preferred composition styles, emphasizing its ability to help them quickly explore musical ideas on paper as they interact with the computer. We conclude with recommendations for designing interactive paper tools that support the creative process, letting users explore musical ideas both on paper and electronically.

Wikibook: [22]. With the Wikibook project [22] we investigate how Wikibooks authors collaborate to create high-quality books. We combined Information Retrieval and statistical techniques to examine the complete multi-year lifecycle of over 50 high-quality Wikibooks. We found that: 1. The presence of redundant material is negatively correlated with collaboration mechanisms; 2. For most books, over 50% of the content is written by a small core of authors; and 3. Use of collaborative tools (predicted pages and talk pages) is significantly correlated with patterns of redundancy. Non-redundant books are well-planned from the beginning and require fewer talk pages to reach high-quality status. Initially redundant books begin with high redundancy, which drops as soon as authors use coordination tools to restructure the content. Suddenly redundant books display sudden bursts of redundancy that must be resolved, requiring significantly more discussion to reach high-quality status. These findings suggest that providing core authors with effective tools for visualizing and removing redundant material may increase writing speed and improve the book’s ultimate quality.

6.3. Engineering of interactive systems

Participants: Caroline Appert, Michel Beaudouin-Lafon, Olivier Chapuis, James Eagan, Tony Gjerlufsen, Stéphane Huot, Wendy Mackay, Clemens Nylandsted Klokmose, Emmanuel Pietriga [correspondant], Clément Pillias, Romain Primet.

We started working on jBricks [24], a Java toolkit enabling the exploratory prototyping of interaction techniques and rapid development of post-WIMP applications running on cluster-driven interactive visualization platforms such as the WILD wall display (Section 7.1). Research on cluster-driven wall displays has mostly focused on techniques for parallel rendering of complex 3D models. There has been comparatively little research effort dedicated to other types of graphics and to the software engineering issues that arise when prototyping novel interaction techniques or developing full-featured applications for such displays. To fill this gap, jBricks integrates a high-quality 2D graphics rendering engine and a versatile input configuration module into a coherent framework, hiding low-level details from the developer. The goal of this framework is to ease the development, testing and debugging of interactive visualization applications for wall-sized displays. It also offers an environment for the rapid prototyping of novel interaction techniques and their evaluation through controlled experiments, such as the one we recently conducted about mid-air pan-and-zoom techniques for wall-sized displays (see Section 6.1).

We developed the Shared Substance framework for multisurface interaction [20]. It is based on Substance, which implements a novel programming model called data orientation that separates functionality from data. Shared Substance extends Substance to distributed environments. It makes distribution explicit so that the programmer can dynamically add, reconfigure and remove components at runtime. An application built with Shared Substance is a collection of processes called environments that run on different machines. Environments are discovered dynamically and can appear and disappear at any time. Each environment contains a hierarchical data structure that can be shared, in whole or in part, with other environments. Sharing can be done through replication or mounting, which entail different performance trade-offs. Shared Substance also includes the Instrumental Interaction Kit (IIK) to facilitate the development of instruments in a multisurface environment [34], [42]. We have used Shared Substance to develop several applications.
for our WILD multisurface environment: *Substance Canvas* manages a virtual canvas that can span multiple interactive surfaces managed by different computers, such as the tiled display, interactive table and users’ laptops of the WILD room; Content can be added to the canvas from various an extensible set of sources, including live applications using Scotty (see below). *Substance Grise* wraps an existing application for displaying 3D brain scans into a Shared Substance environment; This allows us to run 64 copies of the application, each showing a different brain scan, and synchronize the 3D orientation of the scans using a brain prop that the user turns in his hand.

We explored the notion of user interface programming at run-time to create more malleable software [18]. Rather than creating a new user interface toolkit or supporting the customization of an interactive application from outside, we explored how well-defined hooks and a few high-level constructs could allow a programmer to modify an application “from inside”, i.e. using code that is dynamically loaded by the application at run-time. Compared with existing approaches, this supports deep customization that involve the behavior of the application, not just the surface of its user interface. The Scotty prototype implements run-time interface programming in the Mac OS X environment for any application written with the native Cocoa framework. We have used Scotty to distribute the user interface of an application over multiple devices, e.g. to move the Print button of an application onto an iPhone so the user can safely print while physically close to the printer; to replace a tool palette in an application with a toolglass; to check for the presence of attachments in an email application before sending an email; to add subtitles to a video viewer that does not have this functionality. We have also used Scotty in connection with the work on Shared Substance (see above) to teleport a live vector-based representation of a running applications to the WILD wall display. The advantage of this approach over, e.g., VNC, is that the content is properly scaled, taking advantage of the full resolution of the wall.
6. New Results

6.1. Modeling

6.1.1. Reassembly

Participants: Nicolas Mellado, Patrick Reuter, Gaël Guennebaud, Pascal Barla, Christophe Schlick.

In the context of cultural heritage, 3D laser scanning and photogrammetric 3D acquisition of broken content is becoming increasingly popular, resulting in large collections of detailed fractured archaeological 3D objects that have to be reassembled virtually. We recently investigated a semi-automatic reassembly approach for pairwise matching of digital fragments, that makes it possible to take into account both the archeologist’s expertise, as well as the power of automatic geometry-driven matching algorithms. In order to increase matching efficiency and robustness, we currently focus on shape analysis with higher level representation to guide ICP-like algorithms.

6.2. 3D Data Rendering and Visualization

6.2.1. Soft shadows

Participant: Gaël Guennebaud.

Shadows are a fundamental visual effect which both increase the level of realism of a 3D scene, and help to identify spatial relationships between objects. This latter observation makes them particularly important in the context of interactive 3D applications. Generating high quality soft shadows in real-time is still an open challenge. In the continuity of our previous collaboration with the State Key Lab of CAD&CG of Zhejiang University (China) [39], we developed a perceptually based metric dedicated to the prediction of ideal shadow map resolutions [16]. This metric allows us to adaptively generate shadow map tiles. As a result, we managed to render wide and complex exterior scenes with high quality while maintaining high performance (see figure 2).

![Figure 2. Our soft shadow rendering system generates adaptive shadow map tiles and can therefore render softer shadows (right @ 25 fps) faster than their hard shadow counter part (left @ 15 fps).](image)

6.2.2. Synthesis and control of breaking waves

Participants: Nicolas Maréchal, Pascal Barla, Gaël Guennebaud, Patrick Reuter.
Modeling complex breaking waves over arbitrary bathymetry is a tedious problem. Currently, most of the existing methods are based on physical simulations by solving the navier-stokes equation. Controlling the shape of breaking waves is almost impossible with such approaches, and the simulation does not run in real-time. In order to overcome these limitations, we propose a phenomenological approach based on a real-time simulation using airy's wave theory. Our system handles phenomena such as shoaling, refraction and grouping (see figure 3), and the rendering style can be adapted by the user.

Figure 3. Breaking waves generated using our system over an arbitrary bathymetry. 2D wave shape profiles are shown for better legibility.

6.2.3. Analysis and visualization of surface relief

Participants: Lucas Ammann, Pascal Barla, Gaël Guennebaud, Xavier Granier, Patrick Reuter.

Given a base surface with relief, we developed an analysis technique that leverages the complexity found in detailed 3D models for illustrative shading purposes. The key originality of our approach is to extract the relief features such as concavities, convexities and inflections at multiple scales and directions using local cubic-polynomial fitting. We use this information to guide a variety of shading techniques. Our approach is parametrization-free and meshless, allowing for a wide variety of applications ranging from scientific visualization to special effects for the movie industry.

6.3. Expressive Rendering

Figure 4. Conveying visual information through expressive rendering
6.3.1. Line-based Rendering

Participants: Pascal Barla, Jiazhou Chen, Xavier Granier, Christophe Schlick.

We have introduced [18] a new technique called Implicit Brushes to render animated 3D scenes with stylized lines in real-time with temporal coherence. An Implicit Brush is defined at a given pixel by the convolution of a brush footprint along a feature skeleton; the skeleton itself is obtained by locating surface features in the pixel neighborhood. Features are identified via image-space fitting techniques that not only extract their location, but also their profile, which permits to distinguish between sharp and smooth features. Profile parameters are then mapped to stylistic parameters such as brush orientation, size or opacity to give rise to a wide range of line-based styles. This work has won the 3rd best paper award at Eurographics annual conference.

6.3.2. Shape Depiction through Shading

Participants: Pascal Barla, Xavier Granier, Christophe Schlick.

Recently, a number of techniques have been proposed to exaggerate the depiction of shape through the shading of 3D objects. However, existing methods are limited to a single type of material, simple light sources, and they give a fake percept where 3D shape seems to be flattened or embossed, or produce temporal artifacts. We have recently shown that adjusting lighting amplitude for each direction (Radiance Scaling [38], selected as a best paper at I3D 2010 and extended as a TVCG journal paper [17]) may enhance the shape depiction. The technique has been ported to Meshlab (http://meshlab.sourceforge.net/).

6.3.3. Dynamic Expressive Shading Primitives

Participant: Pascal Barla.

Shading appearance in illustrations, comics and graphic novels is designed to convey illumination, material and surface shape characteristics at once. Moreover, shading may vary depending on different configurations of surface distance, lighting, character expressions, timing of the action, to articulate storytelling or draw attention to a part of an object. We have developed [31] a method that imitates such expressive stylized shading techniques in dynamic 3D scenes, and which offers a simple and flexible means for artists to design and tweak the shading appearance and its dynamic behavior. The key contribution of our approach is to seamlessly vary appearance by using a combination of shading primitives that take into account lighting direction, material characteristics and surface features.

6.3.4. Non-Uniform Compositing of Styles

Participants: Jiazhou Chen, Xavier Granier.

In order to investigate how the composition of different styles may help in directing user attention, we have developed [22] a non-uniform composition that integrates multiple rendering styles in a picture driven by an importance map. This map, either issued from salience estimation or designed by a user, is introduced both in the creation of the multiple styles and in the final composition. Our approach accommodates a variety of stylization techniques, such as color desaturation, line drawing, blurring, edge-preserving smoothing and enhancement.

6.4. Interaction

6.4.1. Toucheo: Multitouch + Stereo

Participants: Martin Hachet, Benoit Bossavit, Aurélie Cohé.
We propose a new system that efficiently combines direct multitouch interaction and 3D stereoscopic visualization (see Figure 5). In our approach, the users interact by means of simple 2D gestures on a monoscopic touchscreen, while visualizing occlusion-free 3D stereoscopic objects floating above the surface at an optically correct distance. By coinciding the 3D virtual space with the physical space, we produce a rich seamless workspace where both the advantages of direct and indirect interaction are jointly exploited. In addition to standard multitouch gestures and controls (e.g. pan, zoom, and standard 2D widgets) from which we take advantage, we have designed a dedicated multitouch 3D transformation widget. This widget allows the near-direct control of rotations, scaling, and translations of the manipulated objects. To illustrate the power of our setup, we have designed a demo scenario where participants reassemble 3D virtual fragments. This scenario, as many others, takes benefit of our proposal, where the strength of both multitouch interaction and stereoscopic visualization are unified in an innovative and relevant workspace [19] [25]. See highlights.

Figure 5. Toucheo

6.4.2. Touch-based interaction
Participants: Jérémy Laviole, Aurélie Cohé, Martin Hachet.

We have continued exploring 3D User Interfaces for [multi-]touch screens. In particular, in [26], we conducted a user study to better understand the impact of directness on user performance for a RST docking task, for both 2D and 3D visualization conditions. We have also designed a new 3D transformation widget, called tBox, that can be operated easily and efficiently from simple gestures on touch-screens. In our approach, users apply rotations by means of physically plausible gestures, and we have extended successful 2D tactile principles to the context of 3D interaction [23].

6.4.3. Immersive environments
Participants: Martin Hachet.

We have continued working on immersive environments. In particular, with the "Digital Sound" group of LaBRI, we have studied how sound processes should be visualized in immersive setups [13]. Another collaboration is with the REVES Inria project-team, where we have explored how to design 3D pieces of architecture in a CAVE [21].

6.4.4. Brain-Computer Interaction
Participants: Fabien Lotte.

With Fabien Lotte joining the IPARLA team as a research scientist in January 2011, a new research topic related to interaction is being explored: Brain-Computer Interfaces (BCI). BCI are communication systems that enable its users to send commands to the computer by means of brain activity only, this activity being generally measured using ElectroEncephaloGraphy (EEG). This is therefore a new way to interact with computers and interactive 3D applications. In this area, we have explored new techniques to analyze and process EEG signals in order to identify the mental state of the user [14] [20]. This has led to improved robustness and mental state recognition performances. Another challenge in BCI is that it requires the collection of several examples
of EEG signals from the user in order to calibrate the system. This makes the calibration step long and inconvenient. In order to alleviate this problem, we have proposed to generate artificial EEG signals from a few EEG signals already available. Our evaluations have shown that it can indeed significantly reduce the calibration time [28]. Together with the Inria VR4I-team, we have also explored the use of a new mental state to drive a BCI: attention and relaxation states. We have shown that it is indeed possible to identify relaxation and concentration in EEG signals, and that it can be used to drive a BCI [24]. Finally, we have critically analyzed the usefulness and potential of BCI for 3D video games [27].

6.4.5. Tangible user interfaces

Participant: Patrick Reuter.

Tangible user interfaces have proven to be useful for the manipulation of 3D objects, such as for selection and navigation tasks, and even for deformation tasks. Deforming 3D models realistically is a crucial task when it comes to study the physical behavior of 3D objects, for example in engineering, in sculpting applications, and in other domains. With recent progress in physical deformation models and the increasing computing power, physically-realistic deformation simulations can now be driven at interactive rates. Consequently, there is an increasing demand for efficient and user-friendly user interfaces for physically-realistic deformation in real-time. We designed a general concept for designing physically-realistic deformations of 3D models with a tangible user interface, and instantiated our concept with a concrete prototype using a passive tangible user interface that incarnates the 3D model and that runs in real-time [32].
6. New Results

6.1. Visual tracking

6.1.1. 3D model-based tracking

Participants: Antoine Petit, Eric Marchand.

Our 3D model-based tracking algorithm [3] was used in various contexts. First, it has been studied and tested on a mock-up of a telecommunication satellite using a 6-DOF robotic arm, with satisfactory results in terms of accuracy of the pose estimation and computational costs [41], [42]. A potential application would be the final phase of space rendezvous mission using visual navigation. Then, it has been considered for designing a visual servoing scheme able to control the walking of a humanoid robot [29].

6.1.2. Omnidirectional stereovision

Participants: Guillaume Caron, El Mustapha Mouaddib, Eric Marchand.

Omnidirectional cameras allow direct tracking and motion estimation of planar regions in images during a long period of time. However, using only one sensor leads to plane and trajectory reconstruction up to a scale factor. We proposed to develop dense plane tracking based on omnidirectional stereovision to answer this issue. The method estimates simultaneously the parameters of several 3D planes along with the camera motion using a spherical projection model formulation [20].

6.1.3. Motion estimation using mutual information

Participant: Eric Marchand.

Our work with Amaury Dame related to template tracking using mutual information as registration criterion has been extended to motion estimation applications. It has been applied to mosaicing from an image sequence [28]. The main advantage is that this approach is robust to noise, lighting variations and does not require a statistically robust estimation process.

6.1.4. Augmented reality

Participants: Pierre Martin, Hideaki Uchiyama, Eric Marchand.

We developed an approach for detecting and tracking various types of planar objects with geometrical features [45]. We combine traditional keypoint detectors with Locally Likely Arrangement Hashing (LLAH) for keypoint matching. In order to produce robustness to scale changes, we build a non-uniform image pyramid according to keypoint distribution at each scale. It demonstrates that it is possible to detect and track different types of textures including colorful pictures, binary fiducial markers and handwritings. This approach was extended to consider non-rigidly deformable markers [46].

6.2. Visual servoing

6.2.1. Micro-manipulation

Participant: Eric Marchand.

We developed an accurate nanopositioning system based on direct visual servoing [43], [17]. This technique relies only on the pure image signal to design the control law, by using the pixel intensity of each pixel as visual features. The proposed approach has been tested in terms of accuracy and robustness in several experimental conditions. The obtained results have demonstrated a good behavior of the control law and very good positioning accuracy: 89 nm, 14 nm, and 0.001 degrees in the $x$, $y$ and $\theta_z$ axes of a positioning platform, respectively.
6.2.2. Multi sensor-based control  
**Participants:** Olivier Kermorgant, François Chaumette.

We have designed a generic sensor-based control approach to automatically tune the weights related to the features involved as inputs of a control scheme, allowing to take constraints into account. This scheme has been applied to several configurations, such as fusing the data provided by an eye-in-hand camera and an eye-to-hand camera, ensuring the visibility constraint, and avoiding the robot joint limits [30], [31], [32], [11].

6.2.3. Visual navigation of mobile robots  
**Participants:** Eric Marchand, Andrea Cherubini, Fabien Spindler, François Chaumette.

We have developed a visual servoing scheme based on the mutual information between the images acquired by an onboard camera and a visual memory to control the orientation of a vehicle during its navigation [27]. We have also fused the data provided by a pan-tilt camera and a laser range sensor for the autonomous navigation of a mobile vehicle while avoiding obstacles [23], [22]. Real experiments with our Cycab (see Section 5.4) have been conducted on Place de Jaude in Clermont-Ferrand in the scope of the ANR Tosa CityVIP project (See Section 8.2.1).

6.2.4. Visual servoing for aircrafts  
**Participants:** Céline Teulière, Eric Marchand, Laurent Coutard, François Chaumette.

A dynamic controller has been designed for the homing of a quadri-rotor aerial vehicle [39]. A color-based tracking algorithm has also been designed and combined with an image-based visual servoing for chasing a moving target from a flying UAV [44]. Finally, a method has been developed to detect and localize an aircraft carrier in an image sequence, from which visual servoing control laws have been designed for the automatic landing [25], [26].

6.3. Medical robotics  

6.3.1. Visual servoing based on ultrasound images  
**Participants:** Caroline Nadeau, Alexandre Krupa.

We developed a new approach of ultrasound image based visual servoing that directly uses the intensities of the ultrasound image pixels as visual features. This method that spares any segmentation or image processing time consuming step was initially proposed to control the 6 DOF of a conventional 2D probe for positioning and tracking tasks [38], [48]. To increase the tracking performance we also adapted this method by considering a predictive control law based on the periodicity of physiological motions [36]. Rigid motion compensation experiments were conducted in the context of the ANR USComp project (See Section 8.2.3). The method was also improved by estimating on-line the image 3D gradient required for the positioning task and extended for the use of a bi-plane ultrasound probe [37]. Finally, the use of a 3D motorized probe was also considered to compute directly the image 3D gradient and a comparison of the results obtained with the different probes (2D, bi-plan, 3D) was performed [12].

6.3.2. Autonomous control modes for ultrasound probe guidance  
**Participants:** Tao Li, Alexandre Krupa.
In the context of the ANR Prosit (See Section 8.2.2), we developed several autonomous control modes in order to assist a doctor during a robotized and teleoperated ultrasound examination (tele-echography). The robotic tasks we proposed concern: an automatic scanning of the patient by a 2D probe, a shared control mode that maintains the visibility of an anatomic element of interest while the doctor teleoperates the slave robot holding the 2D probe, an automatic positioning task that allows the doctor to retrieve a desired anatomic section that was previously captured by the doctor. The two latter modes are based on visual servoing schemes that use as input image moments extracted from the observed 2D ultrasound image. This extraction is performed thanks to an active contour (snake) based on Fourier descriptors that we developed and implemented on GPU in order to provide real-time performance [34], [47]. The proposed autonomous control modes were experimentally validated on the Lagadic medical robotics platform (see Section 5.3) and are now in the process of being integrated on the Prosit robot platform.

6.3.3. Real-time 3D ultrasound image reconstruction and 3D deformation tracking

Participants: Deukhee Lee, Alexandre Krupa.

We developed and implemented on GPU an algorithm that reconstructs in real-time a sequence of dense ultrasound volumes from a set of pre-scan 2D ultrasound images provided online by a motorized ultrasound probe [33]. Then we proposed a dense ultrasound tracking algorithm that estimates in real time both rigid and non-rigid motions of a region of interest observed in the sequence of reconstructed ultrasound volumes [33]. The algorithm consists in estimating in real-time, from intensity-value changes between successive 3D ultrasound images, motions of a set of 3D control points that describe the evolution of 3D Thin-Plate Splines (TPS) modeling the deformation. The estimated rigid motion was then used in a pose-based control scheme to automatically displace the probe held by a robot for soft tissue motion compensation. These works were conducted in the context of the ANR USComp project (See Section 8.2.3).
6. New Results

6.1. Large-scale image search

6.1.1. Aggregating local image descriptors into compact codes

Participants: Matthijs Douze, Hervé Jégou [INRIA Rennes], Patrick Pérez [Technicolor], Florent Perronnin [Xerox RCE], Jorge Sánchez [Xerox RCE], Cordelia Schmid.

In [5] we consolidate and extend earlier results for large-scale image search. Different ways of aggregating local image descriptors into a vector are compared. The Fisher vector, see Figure 1, is shown to achieve better performance than the reference bag-of-visual words approach for any given vector dimension. Furthermore, we jointly optimize dimensionality reduction and indexing in order to obtain a precise vector comparison as well as a compact representation. The evaluation shows that the image representation can be reduced to a few dozen bytes with good search accuracy. Given such small codes, searching a 100 million image dataset takes about 250 ms on one processor core.

6.1.2. Searching in one billion vectors: re-rank with source coding

Participants: Laurent Amsaleg [CNRS, IRISA], Matthijs Douze, Hervé Jégou [INRIA Rennes], Romain Tavenard [University Rennes I].

In this work [13] we extend our earlier work [4]. An additional level of processing is added to the product quantizer to refine the estimated distances. It consists in quantizing the difference vector between a point and the corresponding centroid. When combined with an inverted file, this gives three levels of quantization. Experiments performed on SIFT and GIST image descriptors show excellent search accuracy outperforming three state-of-the-art approaches.

6.1.3. Combining attributes and Fisher vectors for efficient image retrieval

Participants: Matthijs Douze, Arnau Ramisa, Cordelia Schmid.

Attributes were recently shown to give excellent results for category recognition. In [9] we demonstrate their performance in the context of image retrieval. We show that combining attributes with Fisher vectors improves performance for retrieval of particular objects as well as categories. Furthermore, we implement an efficient coding technique for compressing the combined descriptor to very small codes. Experimental results show that our approach significantly outperforms the state of the art, even for a very compact representation of 16 bytes per image. We show that attribute features combined with Fisher vectors improve the retrieval of image categories and that those features can supplement text features.

6.1.4. Bag-of-colors for improved image search

Participants: Matthijs Douze, Hervé Jégou [INRIA Rennes], Christian Wengert [Kooaba].
In [19] we investigate the use of color information when used within a state-of-the-art large scale image search system. We introduce a simple color signature generation procedure, used either to produce global or local descriptors. As a global descriptor, it outperforms several state-of-the-art color description methods, in particular the bag-of-words method based on color SIFT. As a local descriptor, our signature is used jointly with SIFT descriptors (no color) to provide complementary information.

6.2. Learning and structuring of visual models

6.2.1. Learning to rank and quadratic assignment

Participants: Thomas Mensink, Jakob Verbeek, Tiberio Caetano [NICTA Canberra].

In [16] we show that the optimization of several ranking-based performance measures, such as precision-at-k and average-precision, is intimately related to the solution of quadratic assignment problems, especially when the score function allows for pairwise label dependencies. Both the task of test-time prediction of the best ranking and the task of constraint generation in estimators based on structured support vector machines can all be seen as special cases of quadratic assignment problems. Although such problems are in general NP-hard, we identify a polynomially-solvable subclass (for both inference and learning) that still enables the modeling of a substantial number of pairwise rank interactions. We show preliminary results on a public benchmark image annotation data set, which indicates that this model can deliver higher performance over ranking models without pairwise rank dependencies. This work was performed during a visit to NICTA Canberra by T. Mensink (March – June, ’11) and J. Verbeek (May ’11).

![Figure 2. An automatically obtained dependency tree over 96 labels, that contains 3 labels per node.](image)

6.2.2. Learning structured prediction models for interactive image labeling

Participants: Thomas Mensink, Jakob Verbeek, Gabriela Csurka [Xerox RCE].

In [25] we propose structured models for image labeling that take into account the dependencies among the image labels explicitly. These models are more expressive than independent label predictors, and lead to more accurate predictions. While the improvement is modest for fully-automatic image labeling, the gain is significant in an interactive scenario where a user provides the value of some of the image labels. Such an interactive scenario offers an interesting trade-off between accuracy and manual labeling effort. The structured models are used to decide which labels should be set by the user, and transfer the user input to more accurate
predictions on other image labels. Experimental results on three publicly available benchmark data sets show that in all scenarios our structured models lead to more accurate predictions, and leverage user input much more effectively than state-of-the-art independent models. See Figure 2.

6.2.3. Modeling spatial layout with Fisher vectors for image categorization

Participants: Frédéric Jurie [University of Caen], Josip Krapac, Jakob Verbeek.

In [15] we introduce an extension of bag-of-words image representations to encode spatial layout. Using the Fisher kernel framework we derive a representation that encodes the spatial mean and the variance of image regions associated with visual words. We extend this representation by using a Gaussian mixture model to encode spatial layout, and show that this model is related to a soft-assign version of the spatial pyramid representation. We also combine our representation of spatial layout with the use of Fisher kernels to encode the appearance of local features. Through an extensive experimental evaluation, we show that our representation yields state-of-the-art image categorization results, while being more compact than spatial pyramid representations. In particular, using Fisher kernels to encode both appearance and spatial layout results in an image representation that is computationally efficient, compact, and yields excellent performance while using linear classifiers.

6.2.4. Unsupervised metric learning for face identification in TV video

Participants: Ramazan Cinbis, Jakob Verbeek, Cordelia Schmid.

The goal of face identification is to decide whether two faces depict the same person or not. In [8] we address the identification problem for face-tracks that are automatically collected from uncontrolled TV video data. Face-track identification is an important component in systems that automatically label characters in TV series or movies based on subtitles and/or scripts: it enables effective transfer of the sparse text-based supervision to other faces. We show that, without manually labeling any examples, metric learning can be effectively used to address this problem. This is possible by using pairs of faces within a track as positive examples, while negative training examples can be generated from pairs of face tracks of different people that appear together in a video frame. In this manner we can learn a cast-specific metric, adapted to the people appearing in a particular video, without using any supervision. Identification performance can be further improved using semi-supervised learning where we also include labels for some of the face tracks. We show that our cast-specific metrics not only improve identification, but also recognition and clustering. See Figure 3.

Figure 3. Projections of face signatures projected to two dimensions, using (a) a metric trained on faces detected in still images, (b) using hand labeled faces detected in videos, (c) a metric trained from face tracking results (no manual labeling). Face signatures of different people are color coded. A good face metric can be learned directly from face tracking results, without using any hand labeled examples.
6.2.5. Large-scale image classification

Participants: Miro Dudik [Yahoo! Research], Zaid Harchaoui, Jerome Malick [INRIA Grenoble, BIPOP Team].

We introduced in [10] a new scalable learning algorithm for large-scale multi-class image classification, based on the multinomial logistic loss and the trace-norm regularization penalty. Reframing the challenging non-smooth optimization problem into a surrogate infinite-dimensional optimization problem with regular $\ell_1$-regularization penalty, we propose a simple and provably efficient coordinate descent algorithm. Furthermore, we showed how to perform efficient matrix computations in the compressed domain for quantized dense visual features, scaling up to 100,000s examples, 1,000s-dimensional features, and 100s of categories. Promising experimental results on the “Fungus”, “Ungulate”, and “Vehicles” subsets of ImageNet were obtained, where our approach performed significantly better than state-of-the-art approaches for Fisher vectors with 16 Gaussians.

6.3. Human action recognition

6.3.1. Action recognition by dense trajectories


Feature trajectories have shown to be efficient for representing videos. Typically, they are extracted using the KLT tracker or matching SIFT descriptors between frames. However, the quality as well as quantity of these trajectories is often not sufficient. Inspired by the recent success of dense sampling in image classification, in [18] we propose an approach to describe videos by dense trajectories. An overview of our framework is shown in Figure 4. We sample dense points from each frame and track them based on dense optical flow. Our trajectories are robust to fast irregular motions as well as shot boundaries. Additionally, dense trajectories cover the motion information in videos well. We also investigate how to design descriptors to encode the trajectory information. We introduce a novel descriptor based on motion boundary histograms, which is robust to camera motion. This descriptor consistently outperforms other state-of-the-art descriptors, in particular in uncontrolled realistic videos. We evaluate our video description in the context of action classification with a bag-of-features approach. Experimental results show a significant improvement over the state of the art on four datasets of varying difficulty, e.g., KTH, YouTube, Hollywood2 and UCF sports.

![Figure 4. Illustration of dense trajectories extraction and description. Left: dense sampling of feature points at multiple scales; middle: tracking feature points with a dense optical flow field; right: descriptors are computed along the trajectory.](image-url)
6.3.2. Weakly supervised learning of interactions between humans and objects

**Participants:** Vittorio Ferrari [ETH Zürich], Alessandro Prest, Cordelia Schmid.

In [7] we introduced a weakly supervised approach for learning human actions modeled as interactions between humans and objects. Our approach is human-centric: we first localize a human in the image and then determine the object relevant for the action and its spatial relation with the human. The model is learned automatically from a set of still images annotated only with the action label. Our approach relies on a human detector to initialize the model learning. For robustness to various degrees of visibility, we build a detector that learns to combine a set of existing part detectors. Starting from humans detected in a set of images depicting the action, our approach determines the action object and its spatial relation to the human. Its final output is a probabilistic model of the human-object interaction, i.e., the spatial relation between the human and the object. We present an extensive experimental evaluation on the sports action dataset from Gupta et al., the PASCAL 2010 action dataset, and a new human-object interaction dataset. In the PASCAL visual object classes challenge 2011 our approach achieved best results on three out of ten action classes and the best result on average over all classes.

6.3.3. Explicit modeling of human-object interactions in realistic videos

**Participants:** Vittorio Ferrari [ETH Zürich], Alessandro Prest, Cordelia Schmid.

In [26] we introduced an approach for learning human actions as interactions between persons and objects in realistic videos. Previous work typically represents actions with low-level features such as image gradients or optical flow. In contrast, we explicitly localize in space and track over time both the object and the person, and represent an action as the trajectory of the object wrt to the person position. Our approach relies on state-of-the-art approaches for human and object detection as well as tracking. We show that this results in human and object tracks of sufficient quality to model and localize human-object interactions in realistic videos. Our human-object interaction features capture relative trajectory of the object wrt the human. Experimental results on the Coffee & Cigarettes dataset show that (i) our explicit human-object model is an informative cue for action recognition; (ii) it is complementary to traditional low-level descriptors such as 3D-HOG extracted over human tracks. When combining our human-object interaction features with 3D-HOG features, we show to improve over their separate performance as well as over the state of the art. See Figure 5.

6.3.4. Actom sequence models for efficient action detection

**Participants:** Adrien Gaidon, Zaid Harchaoui, Cordelia Schmid.

In [12] we address the problem of detecting actions, such as drinking or opening a door, in hours of challenging video data. We propose a model based on a sequence of atomic action units, termed "actoms", that are characteristic for the action. Our model represents the temporal structure of actions as a sequence of histograms of actom-anchored visual features. Our representation, which can be seen as a temporally structured extension of the bag-of-features, is flexible, sparse and discriminative. We refer to our model as Actom Sequence Model (ASM). Training requires the annotation of actoms for action clips. At test time, actoms are detected automatically, based on a non-parametric model of the distribution of actoms, which also acts as a prior on an action’s temporal structure. We present experimental results on two recent benchmarks for temporal action detection. We show that our ASM method outperforms the current state of the art in temporal action detection.

6.3.5. A time series kernel for action recognition

**Participants:** Adrien Gaidon, Zaid Harchaoui, Cordelia Schmid.

In [11] we address the problem of action recognition by describing actions as time series of frames and introduce a new kernel to compare their dynamic aspects. Action recognition in realistic videos has been successfully addressed using kernel methods like SVMs. Most existing approaches average local features over video volumes and compare the resulting vectors using kernels on bags of features. In contrast, we model actions as time series of per-frame representations and propose a kernel specifically tailored for the purpose of action recognition. Our main contributions are the following: (i) we provide a new principled way to compare
the dynamics and temporal structure of actions by computing the distance between their auto-correlations, (ii) we derive a practical formulation to compute this distance in any feature space deriving from a base kernel between frames, and (iii) we report experimental results on recent action recognition datasets showing that it provides useful complementary information to the average distribution of frames, as used in state-of-the-art models based on bag-of-features.
LEO Team

6. New Results

6.1. Efficient XML and RDF data management

**Participants:** Mohamed Amine Baazizi, Nicole Bidoit, Dario Colazzo, François Goasdoué, Konstantinos Karanasos, Asterios Katsifodimos, Julien Leblay, Noor Malla, Ioana Manolescu, Alexandra Roatis, Marina Sahakyan, Federico Ulliana.

6.1.1. Materialized views for XML

We have continued our work on optimizing XML queries through materialized view-based rewriting, implemented within the ViP2P system. We published in IEEE ICDE 2011 an algorithm for rewriting XQuery queries using materialized XQuery view, which improves the state of the art in terms of expressive power of the supported XQuery subset, in collaboration with V. Vassalos (AUEB, Greece) [41]. Two follow-up works concern: efficient algebraic algorithms for incrementally maintaining the materialized views when the underlying documents change, in collaboration with A. Bonifati (CNR, Italy) [35], and algorithms for automatically recommending views to materialize for a given XML query workload, with V. Vassalos.

6.1.2. Type-based Update Optimization for XML

XML projection is a well-known optimization technique for reducing memory consumption for XQuery in-memory engines in order to overcome the main-memory limitations of these systems (Galax, Saxon, QizX, and eXist). One of our main research line focuses on a schema-based projection technique for for update optimization. The update language considered is XQuery Update Facility (XUF). The main idea behind this technique is: given a query $q$ over an XML document $t$, instead of evaluating $q$ over $t$, the query $q$ is evaluated on a smaller document $t'$ obtained from $t$ by pruning out, at loading time, parts of $t$ that are irrelevant for $q$. The queried document $t'$, a projection of the original one, is often much smaller than $t$ due to selectivity of queries.

The scenario and type-based projection proposed for XML queries, cannot be applied directly for updates. We have proposed a new scenario which is composed of four steps:

1. from the update $U$ and the DTD $D$, a type projector $P$ is inferred;
2. the document $t$, valid wrt $D$, is projected following $P$ in a streaming manner, at loading time;
3. $U$ is evaluated over the projection $P(t)$ and produces a partial result $U(P(t))$;
4. the initial document $t$ is merged with $U(P(t))$, in a streaming manner, at writing (serializing) time in order to produce the final result $U(t)$.

The scenario has been first studied and implemented for a kind of type projector which is a good compromise between simplicity and effectiveness, and corresponding results have been published in [31]. Subsequently, we have improved this technique by designing: (i) a new kind of type projector that minimizes the amount of data kept in the projection, and (ii) a new merge algorithm using the improved type projector. This analysis is complicated by the strong interconnection between the two tasks: while minimizing the projection we need to ensure a fast and correct merge process [17].

These results have also been presented in [34], providing an overview on the use of types and constraints from relational to XML data, and in the tutorial [63] focusing on schema-based techniques for safe and efficient XML processing.
6.1.3. XML query-update independence

A query and an update are independent when the query result is not affected by update execution, on any possible input database. Detecting query-update independence is of crucial importance in many contexts: view maintenance, concurrency, access control policies etc. Benefits are amplified when query-update independence can be checked statically. We propose a novel schema-based approach for detecting XML query-update independence. Differently from traditional schema-based analysis for XQuery, our system infers sequence of labels, called chains, that are vertically navigated in each schema instance by query and update paths. More precisely, for each node that can be selected by a query/update path in a schema instance, the system infers a chain recording: a) all labels that are encountered from the root to the selected node, and b) the order in which these labels are traversed. The contextual and ordering information provided these chains is at the basis of an extremely precise static independence analysis. We have devised a quite precise chain-inference system, and devised techniques for efficient implementation of the chain-based independence analysis. Results and experiments concerning this line of research have been recently submitted to an international conference.

6.1.4. Precision and complexity of XQuery type inference

A key feature of XQuery is its type system. Any language expression is statically typed and its type is used during program type-checking. In XQuery, types of input data and functions are defined in terms of regular expression types, but it is quite easy to write queries that generate non-regular languages. As a consequence, any type system for XQuery has to rely on a type inference process that approximates the (possibly non-regular) output type of a query with a regular type. This approximation process, while mandatory and unavoidable, may significantly decrease the precision of the inferred types. In [37] we study and compare in terms of precision and computational complexity two main existing XQuery type systems.

6.1.5. Managing temporal XML documents

The management of temporal data is a crucial issue in many database applications. We are currently investigating efficient storage and update methods for temporal XML documents, with a focus on compactness of the representation. One of the method developed relies on the type based optimization method developed for updates [31]. First results about this line of research are included in recent publications [29], [30]. Next research activities focus on the design of expressive temporal query and update languages, and on the use of techniques proposed in [29] for query and update optimization.

6.1.6. Materialized view selection for RDF

Syntactically, RDF, the data format of the Semantic Web, resembles relational data. However, RDF query processing is significantly complicated by the irregular nature of RDF data and by its simplistic data model, which leads to syntactically complex queries (involving many joins over the whole triple data set). When a query workload is known, the performance of the workload can be significantly improved by materializing access support data structures such as materialized views. Our efficient algorithms for selecting a set of views to materialize in order to speed up the processing of a set of RDF queries are described in a recent publication [21]. The prototype implementing them has been demonstrated at [53].

6.1.7. Hybrid models for XML and RDF

We have obtained interesting results in the area of jointly managing XML and RDF data. A first direction of work in this area was to support annotated documents, that is, XML documents where individual nodes or fragments could be annotated with RDF triples. This model allows to capture, for instance, blog comments, user ratings on social sites etc. We have proposed a general model based on W3C standards for modeling such data [39].

6.1.8. RDF query answering

The current trend for efficiently querying RDF datasets consists of delegating query evaluation to a scalable RDBMS. However, RDF query answering requires in addition to handle – outside the RDBMS – the RDF semantics. We have introduced the database (DB) fragment of RDF, encompassing the popular Description
Logic (DL) one with essential RDF features like modeling incomplete information, for which we have devised novel saturation- and reformulation-based techniques for answering the Basic Graph Pattern (BGP) queries of SPARQL. This extends the state of the art on pushing RDF query processing within RDBMSs.

6.2. Models for Web data management

Participants: Serge Abiteboul, Emilien Antoine, Meghyn Bienvenu, Alban Galland.

A book on Web Data Management and Distribution [54] was published this year.

6.2.1. A rule-based language for Web data management

We recently proposed [26] a Datalog-style rule-based language (called Webdamlog) for web data management. A novel feature of our language is delegation, that is, the possibility of installing a rule at another peer. In its simplest form, delegation is essentially a remote materialized view. In its general form, it allows peers to exchange rules, i.e., knowledge beyond simple facts, and thereby provides the means for a peer to delegate work to other peers.

A key contribution of our work is a study of the impact on expressiveness of delegations and explicit timestamps. We showed that both strictly augment the power of the language. In order to validate the semantics of our model, we demonstrated that under certain natural conditions, our semantics converges to the same semantics as the centralized system with the same rules.

6.2.2. Web information management with access control

We investigated the problem of sharing private information on the Web, where the information is hosted on different machines that may use different access control and distribution schemes. Based upon our work on Webdamlog, we introduced a distributed knowledge-base model, termed WebdamExchange, that comprises logical statements for specifying data, access control, distribution and knowledge about other peers. In a demo at ICDE [28], we showed how the model can be used in a social-network context to help users keep control on their data on the web. In particular, we demonstrated how users within very different schemes of data distribution (centralized, DHT, unstructured P2P, etc.) can still transparently collaborate while keeping a good control over their own data.

6.3. Ontology-based data and document Management

Participants: Meghyn Bienvenu, François Goasdoué, Yassine Mrabet, Nathalie Pernelle, Gianluca Quercini, Chantal Reynaud, Brigitte Safar, Fabian Suchanek.

6.3.1. Semantic Annotation

We have started a work on semantic annotation of public administration data in the setting of the project DataBridges, an ICT Labs activity. We considered public data represented in tables. The tables that we studied were tables created and published by INSEE. They are spreadsheets filled with statistics about geographic locations and are usually composed of multiple columns, of which one, that we term the subject column, contains a list of textual references to geographic entities, or toponyms, while the others contain numeric attributes. We proposed an approach and an algorithm that assigns a type, or header to the subject column of an INSEE table and identifies the geographic entities referred to by the toponyms in the column [64]. An external resource, DBpedia, is used to help to disambiguate the entities mentioned in the tables and a domain ontology ensures that the types are relative to the geographic domain. This work is continued in the setting of a post-doctoral work granted by the ANR project DataBridges. The aim of the project being to enrich a data warehouse, a first work is to automatically build an initial RDF data warehouse from data collected from the web.

6.3.2. Adaptive Ontologies for Information Retrieval

We published the approach supported by the TARGET framework for Web Information Retrieval in the International Journal of Web Portals (IJWP) [22]. This approach was the core of the PhD of Cédric Pruski defended in April 2009.
6.3.3. Querying ontology-based annotations

We have pursued our work on integrating knowledge bases and semantic annotations made on more or less structured tagged documents. We have defined an approach where RDF named graphs are used to distinguish uncertain semantic annotations from rdf triples that are provided by the populated ontology. A user domain query is then reformulated to obtain answers that are ranked according to their provenance (knowledge bases or annotations) [61].

6.3.4. Watermarking for ontologies

Ontologies are usually available under some type of license. The large ontologies of the Semantic Web, e.g., are commonly licensed under a Creative Commons License or a GNU license. These licenses require giving credit to the authors of the ontology if the ontology is ever used somewhere else. However, it can be hard to prove whether an ontology is used somewhere else, because ontologies contain world knowledge. If someone “steals” an ontology and uses it somewhere else, he can always claim that he collected the data by himself from real-world sources. To tackle this problem, we have studied approaches that watermark an ontology [43]. If a watermarked ontology is used somewhere else, the mark proves that the ontology has been stolen. Existing approaches have mainly modified the facts in the ontology to create a mark. This, however, compromises the precision of the ontology. Therefore, we have developed an approach that does not modify, but remove certain facts. Thereby, the precision of the ontology is not affected. We show that only a handful of facts have to be removed from an ontology to protect it against theft.

6.3.5. Consistent query answering in DL-Lite

An important problem which arises in ontology-based data access is how to handle inconsistencies. In the database community, the related problem of querying databases which violate integrity constraints has been extensively studied under the name of consistent query answering. The standard approach is based on the notion of a repair, which is a database which satisfies the integrity constraints and is as similar as possible to the original database. Consistent answers are defined as those answers which hold in all repairs. A similar strategy can be used for description logics by replacing the integrity constraints with the ontology. Unfortunately, recent work on consistent query answering in description logics has shown this problem to be co-NP-hard in data complexity, even for instance queries and the simplest DL-Lite dialect. In light of this negative result, we considered the problem of identifying cases where consistent query answering is feasible, and in particular, can be done using query rewriting, with the aim of better understanding the cases in which query rewriting can be profitably used. In [51], we make some first steps towards this goal by formulating general conditions which can be used to prove that a consistent rewriting does or does not exist for a given DL-Lite TBox and instance query.

6.3.6. Module-based data management in DL-Lite

The current trend for building an ontology-based data management system (DMS) is to capitalize on efforts made to design a preexisting well-established DMS (a reference system). The method amounts to extract from the reference DMS a piece of schema relevant to the new application needs – a module –, possibly to personalize it with extra-constraints w.r.t. the application under construction, and then to manage a dataset using the resulting schema. We have revisited the reuse of a reference ontology-based DMS in order to build a new DMS with specific needs. We go one step further by not only considering the design of a module-based DMS (i.e., how to extract a module from a ontological schema): we also study how a module-based DMS can benefit from the reference DMS to enhance its own data management skills. We consider the setting of the DL-LiteA dialect of DL-Lite, which encompasses the foundations of the QL profile of OWL2 (i.e., DL-LiteR): the W3C recommendation for managing efficiently large datasets. We introduce and study novel properties of robustness for modules that provide means for checking easily that a robust module-based DMS evolves safely w.r.t. both the schema and the data of the reference DMS. From a module robust to consistency checking, for any data update in a corresponding module-based DMS, we show how to query the reference DMS for checking whether the local update does not bring any inconsistency with the data and the constraints of the reference DMS. From a module robust to query answering, for any query asked to a module-based
DMS, we show how to query the reference DMS for obtaining additional answers by also exploiting the data stored in the reference DMS.

6.4. Data and Knowledge Integration

Participants: Julio Cesar Dos Reis, Fayçal Hamdi, Rania Khefifi, Yassine Mrabet, Nathalie Pernelle, Chantal Reynaud, Fatiha Saïs, Brigitte Safar, Fabian Suchanek, Danai Symeonidou.

6.4.1. Reference Reconciliation

The reference reconciliation problem consists in deciding whether different data descriptions refer to the same real world entity (same person, same conference etc.) Some of existing approaches, such as LN2R, are declarative and knowledge-based. Different kinds of knowledge can be declared in a domain ontology, like disjointness between classes or key constraints. This knowledge can be exploited to infer reconciliation and non-reconciliation decisions.

Our reference reconciliation work pursues three directions:

- develop an automatic approach of key constraint discovery. We have proposed in [46] KD2R, a method which allows automatic discovery of key constraints associated to OWL2 classes. These keys are discovered from RDF data which can be incomplete. The proposed algorithm allows this discovery without having to scan all the data. KD2R has been tested on data sets of the international contest OAEI and obtains promising results.

- develop a reference reconciliation method for detecting redundant data in case of web data tables that are semantically annotated by an ontology. Each table cell values consists in numerical fuzzy set (NFS) or in symbolic fuzzy set (SFS). We have developed a method which uses ontology knowledge and computes similarity scores to decide the data redundancy. We have also proposed two similarity measures for numerical fuzzy set as well as symbolic fuzzy set. The proposed measures are more flexible than existing ones. This approach has been published in [36], [58]. We are working on its extension to be able to distinguish redundant data from similar ones by using provenance information.

- develop a new approach which addresses the problem of resource discovery in the Linked Open Data cloud (LOD) where data described by different schemas is not always linked. We have proposed an automatic approach in [42], [58] that allows discovery of new links between data. These links can help to match schemas that are conceptually relevant with respect to a given application domain. Furthermore, these links can be exploited during the querying process in order to combine data coming from different sources. In this approach we exploit the semantic knowledge declared in different schemas in order to model: (i) the influences between concept similarities, (ii) the influences between data similarities, and (iii) the influences between data and concept similarities. The similarity scores are computed by an iterative resolution of two non linear equation systems that express the concept similarity computation and the data similarity computation.

6.4.2. Context-aware Personal Information Management

Personal information management (PIM) is the practice and analysis of the activities performed by people to acquire, organize, maintain, and retrieve information for everyday use. PIM is a growing area of interest because everyone is looking for better use of our limited personal resources of time, money and energy. Several research on the topic is being done in different disciplines, including human-computer interaction, database management, information retrieval and artificial intelligence.

The increasingly big amount of personal information (e.g., mails, contacts, appointments) managed by a user is characterized by their heterogeneity, their dispersion and their redundancy. The general goal of this work consists in designing a system, which allows providing the end-users personal data access with services that are relevant to his/her needs, and to access personal data both by mobile devices (smartphone) and Internet-connected Personal Computers. More specifically, we focus here on the problem of defining a common
meta-model for a flexible and homogeneous personal information management. The meta-model that we propose allows users creating personal information and organizing them according to different points of view (ontologies) and different contexts. Contextual queries are defined to allow users to retrieve its personal information using the geographical contexts. The semantic Web languages (OWL, RDF and SPARQL) are used to implement the approach.

6.4.3. Mapping between ontologies

We pursue our work on ontology alignment in the setting of the ANR GeOnto project by aiming to provide full life-cycle support for ontologies.

We investigated how alignment results generated by our alignment tool, TaxoMap, can be used to enrich one ontology with another. We shown that the enrichment process depends on characteristics of the ontology used for enrichment. Three enrichment contexts identified in the setting of the ANR project GeOnto have been studied and enrichment treatments performed. A first context considers ontologies of the same application domain and of a reasonable size. A second context considers small ontologies previously extracted from a generalist one. A third context considers enrichment from a huge, generalist ontology, such as Yago. Early results obtained in the setting of the ANR project GeOnto in the topographic domain have been published in [50], [25].

The module supporting our enrichment approach has been implemented in TaxoMap Framework using patterns. Initially, TaxoMap Framework was composed of our alignment tool, TaxoMap, we are working on for several years in the team and of a mapping refinement module. We extended it in order to obtain a broader framework and an interactive environment by including TaxoPart, a partitioning tool we developed to split two huge ontologies which could not be aligned into two sets of blocks of a limited size, and a module specific to ontology enrichment. Moreover, we re-implemented TaxoMap, our alignment tool, as a web service to make it easily accessible at: http://taxomap.lri.fr:8000/axis2/services/TaxoMapService?wsdl.

We also started a PhD work, joined with CRP Henri Tudor in Luxembourg, to investigate issues dealing with medical knowledge organizing systems evolution. We will define a formal framework to support medical knowledge organizing systems evolution in a consistent way and also to support the maintenance of mappings directly impacted by knowledge organizing systems local evolution.

On a related topic, we have developed a probabilistic framework, PARIS (Probabilistic Alignment of Relations, Instances and Schema), for matching ontologies holistically, thereby exploiting synergies between matches on the instance level and matches on the schema level [57]. The framework is parameter-free and does not require resource-specific tuning. PARIS is fully implemented and has been shown to match some of the largest ontologies on the Semantic Web with a precision of around 90%.

6.4.4. Integration of Web resources

We have pursued our work on integration of resources available on the Web in Adaptive Hypermedia Systems (AHS), allowing creators to define their own adaptation strategies based on their own domain models. The approach is based on a set of 22 adaptation patterns, independent of any application domain and independent of any adaptation engine, published in [59], [47]. These elementary adaptation patterns are organized in a typology in order to facilitate their understanding and their use in the EAP framework to define complex strategies. In [24], we described the whole process to generate complex adaptation strategies and how the generated strategies can be integrated into existing AHSs. The results of an experiment conducted in the e-learning domain is presented. It showed that the pattern-based approach for defining adaptation strategies is more suitable than those based on “traditional” AH languages.

We also pursued our work on the integration of the EAP framework and other AHSs. Our collaboration with A. Cristea from the University of Warwick (UK) led us to a very detailed study of adaptation languages. The first flexible generic adaptation language is the LAG adaptation language. We studied the expressivity of this initial adaptation language in comparison with our newly proposed language, in the EAP framework, and the pros and cons of various decisions in terms of the ideal way of defining an adaptation language. We proposed a unified vision of adaptation and adaptation language. The unified vision is not limited to the
two languages analyzed, and can be used to compare and extend other approaches in the future. Beside this theoretical qualitative study, we also made experimental evaluation and comparison of the two languages, and an article is currently being evaluated.

We have also investigated integration of Web services. The Search Computing project ("SeCo") at the Polytechnic University of Milan aims to orchestrate Web services to answer user queries. Currently, the project represents Web services by so-called Service Marts. These are frame-like representations of the services, which follow the slot-value paradigm. This representation faces several challenges if more Web services get added to the system, because it is hard to ensure that Web services added by different users can still be joined. Therefore, we have explored a more ontological representation of Web services. In our proposal [55], Web services are represented as sub-graphs of an ontology. This allows users to add new Web services that re-use the vocabulary of existing Web services.

On related topics, together with researchers from the Max-Planck Institute in Saarbrucken, we have worked on extending the YAGO ontology. YAGO already contains dozens of millions of facts. With the present work, we aim to give these facts a temporal and a spatial dimension. For every event and every entity, we want to know where and when these objects existed. For this purpose, we have developed a methodology that extracts these types of facts from Wikipedia. We have also developed a logical reasoning framework that allows propagating these time and space annotations from some facts to others. This has grown YAGO to 80 million facts in total, making it an ontology that is anchored in time and space (Best demo award at the WWW 2011 conference [40]).

6.5. Reasoning over Distributed Systems


6.5.1. Distributed Diagnosis Problems

We pursued the work on distributed algorithms for diagnosing distributed systems. The general framework is consistency-based diagnosis for propositional-logic theories in a P2P setting with privacy constraints. It boils down to distributed implicant finding and is thus in some sense dual to the problem of consequence finding described in next paragraph. Vincent Armant is finishing his PhD and has extended his previous work on more general topics, i.e. focusing on the construction of a good decomposition of the network that will ensure an efficient reasoning mechanism. An important effort has been put in the design of a real-world sized experimentation on distributed systems.

Lina Ye defended her PhD on diagnosability analysis of distributed discrete-event systems, modeled as synchronized labeled automata. The aim of diagnosability is to ensure that a given partially observable system has the property that any fault (taken from a set of faults given a priori) will be detectable and identifiable without ambiguity in a finite time after its occurrence. Distributed diagnosability analysis is optimized by abstracting necessary and sufficient information from local objects to achieve global decision. After having addressed the distribution of the system’s model into local models, we focus in 2011 on the extension to systems where the observable information itself is distributed instead of centralized. Joint diagnosability definition has been provided and undecidability of deciding it has been proved in the general case where communication events are not observable, before proposing an algorithm to test its sufficient condition. In addition, decidability result and algorithm have been given when communications are observable.

Michel Batteux defended also his PhD (led in the framework of a CIFRE thesis with Sherpa Engineering) about diagnosability and diagnosis of technological systems. The work was led in the centralized case, focusing on defining, implementing, testing and validating on a real case study (a fuel cell system) an all-in-one tool to design a diagnosis system for technological systems by integrating representation of the system and its potential faults, off-line diagnosability analysis and automatic generation of the on-line embedded diagnoser.
6.5.2. Distributed Consequence Finding

A major reengineering of the SOMEWHERE platform, for decentralized consequence finding, has been initiated within the DISQUE project. Current efforts have focused on the rewriting of the communication layer, that now relies on the JXTA middleware. A new tool is also being developed, in order to facilitate large scale experimentations on a grid (Grid5000). This tool is designed in a fairly generic way, in order to be reusable for similar projects that require deploying sets of collaborating reasoners in a decentralized setting, and automating collaborative problem solving on various instances. We also expect this tool to be used for automating integration tests during further developments.

6.5.3. Towards distributed architectures for Modern SAT Solvers

If we aim at proposing a new architecture for distributed SAT Solvers, we pursued this year the improvements of Glucose, our centralized SAT solver. Glucose 2 won 3 medals at the SAT 2011 Competition, and one in the category Application SAT+UNSAT. We target to make a massively distributed version of Glucose, for very hard SAT problems.
6. New Results

6.1. Scene and camera reconstruction

Participants: Marie-Odile Berger, Srikrishna Bhat, Nicolas Noury, Gilles Simon, Frédéric Sur.

6.1.1. Image point correspondences and repeated patterns

Matching or tracking interest points between several views is one of the keystones of many computer vision applications, especially when considering structure and motion estimation. The procedure generally consists in several independent steps: interest point extraction, then interest point matching by keeping only the “best correspondences” with respect to the similarity between some local descriptors, and final correspondence pruning to keep those that are consistent with a realistic camera motion (here, consistent with epipolar constraints or homography transformation.) Each step in itself is a delicate task which may endanger the whole process. In particular, repeated patterns give rise to lots of false correspondences in descriptor-based matching. Actual correspondences are thus hardly, if ever, recovered by the final pruning step. Dealing with repeated patterns is of crucial importance in man-made environments. Starting from a statistical model by Moisan and Stival [25], we have proposed a one-stage approach for matching interest points based on simultaneous descriptor similarity and geometric constraint. The resulting algorithm has adaptive matching thresholds and is able to pick up point correspondences beyond the nearest neighbour. We have also shown how to improve ASIFT [26], an effective point matching algorithm to make it more robust to the presence of repeated patterns [5], [23], [8].

6.1.2. Visual words for pose computation

Visual vocabularies are standard tools in the object/image classification literature, and are emerging as a new tool for building point correspondences for pose estimation. Within S. Bhat’s PhD thesis, we have proposed several methods for visual word construction dedicated to point matching, with structure from motion and pose estimation applications in view. The three dimensional geometry of a scene is first extracted with bundle adjustment techniques based on keypoint correspondences. These correspondences are obtained by grouping the set of all SIFT descriptors from the training images into visual words using transitive closure (TC) techniques. We obtain a more accurate 3D geometry than with classical image-to-image point matching. In a second on-line step, these visual words serve as 3D point descriptors that are robust to viewpoint change, and are used for building 2D-3D correspondences on-line during application, yielding the pose of the camera by solving the PnP problem. Several visual word formation techniques have been compared with respect to robustness to viewpoint change between the learning and the test images. Our experiments showed that the adaptive TC visual words are better in many ways when compared to other classical techniques such as K-means [12].

6.1.3. Tracking by synthesis using point features and pyramidal blurring

Tracking-by-synthesis is a promising method for markerless vision-based camera tracking, particularly suitable for Augmented Reality applications. In particular, it is drift-free, viewpoint invariant and easy-to-combine with physical sensors such as GPS and inertial sensors. While edge features have been used successfully within the tracking-by-synthesis framework, point features have, to our knowledge, still never been used. This is probably due to the fact that real-time corner detectors are weakly repeatable between a camera image and a rendered texture.
We compared the repeatability of commonly used FAST, Harris and SURF interest point detectors across view synthesis [17]. We showed that adding depth blur to the rendered texture can drastically improve the repeatability of FAST and Harris corner detectors (up to 100% in our experiments), which can be very helpful, e.g., to make tracking-by-synthesis running on mobile phones. We proposed a method for simulating depth blur on the rendered images using a pre-calibrated depth response curve. In order to fulfill the performance requirements, a pyramidal approach was used based on the well-known MIP mapping technique. We also proposed an original method for calibrating the depth response curve, which is suitable for any kind of focus lenses and comes for free in terms of programming effort, once the tracking-by-synthesis algorithm has been implemented.

6.1.4. Acquisition of 3D calibrated data

Christel Leonet joined the team in October 2010 as an INRIA assistant engineer with the aim to build an integrated 3D acquisition system. More specifically, the objective of her work is to combine an IMU (Inertial Measurement Unit), a GPS receiver, a laser rangefinder and a video camera for ground truth data acquisitions of camera movements and scene structures. These data will be useful to validate several algorithms developed in our team. This year she dealt with the hand-eye coordination between the different devices. Moreover, a 3D laser pointer has been built, which allows to acquire textured 3D polygons by pointing them with the laser attached to the camera and the IMU put on a tripod.

6.2. Medical imaging


6.2.1. Vessel reconstruction with implicit surfaces

Our research activity is led in collaboration with Shacra project-team from INRIA Lille-Nord Europe and the Department of Interventional Neuroradiology from Nancy University Hospital. It was pursued this year in the context of the SOFA-InterMedS INRIA Large-Scale Initiative.

Our objective is the implicit modeling of blood vessels from 3DRA data, with the aim to use these models for real-time simulation of interventional procedures. Within A. Yureidini’s PhD thesis, a new model was developed consisting of a tree of local implicit blobby models. This model was implemented in Sofa simulation platform, enabling interactive simulation time (60 fps) and thereby showing an impressive realism during tool navigation [20]. We focused this year on the extensive validation of our RANSAC-based vessel tracking algorithm, by comparison with state-of-the-art Multiple Hypothesis Testing [24] on 10 patient data [18]. Our initial mechanism to fit the implicit model to patient data relies on the minimization of a multi-termed energy. This energy was put under scrutiny, assessing the contribution of each energy term [19]. Our current goal is to reintroduce the raw image data for a more accurate energy computation, with the aim to design a blobby deformable model.

6.2.2. A variational framework for automatic modeling of the vocal tract

Segmenting the vocal tract in MRI is difficult especially because the tongue may move near other edges in the oral cavity, such as the palate or the teeth, which may disturb the segmentation process. The idea explored in our past work was to guide the segmentation with shape priors learnt on a reference speaker within a shape-based variational framework.

Shape priors were incorporated into segmentation via a PCA model with a relatively large number of components to enable the adaptation of the model to strong morphological differences. During this year, this work was continued with the aim to detect tongue contours in physical correspondences, thus allowing us to build a model of the vocal tract. An automatic method for the identification of the end points as well as an improved variational framework to obtain curves in physical correspondences was described in [15]. Second, we extensively assessed the segmentation process. We experimentally showed that the reference model is able to cope with strong morphological differences between speakers with a limited numbers of modes.
6.2.3. Medical simulators based on task analysis

We present here two works done within a collaboration with Imperial College of London. In order to validate a virtual reality ultrasound-guided targeted liver biopsy procedure simulators previously designed [22], we have worked on task analysis to deconstruct individual procedural tasks followed by metric definition and critical performance indicator identification. Consultant and trainee scores on the performance metrics were compared. Independent t-tests revealed significant differences between trainees and consultants on 3 performance metrics: targeting, probe usage time and mean needle length in beam. ANOVA reported significant differences across years of experience on seven performance metrics: no-go area touched, targeting, length of session, probe usage time, total needle distance moved, number of skin contacts, total time in no-go area. More experienced participants consistently received better performance scores on all 19 performance metrics [9].

We used the same task analysis technique to design an inguinal hernia repair simulator [16]. The task analysis allowed to break down the complex operation into sub-tasks and it also provided the foundation for useful and productive discussions between clinical staff and developers. We deployed our system as an e-learning application, allowing surgeons to easily access the application.

6.3. National Initiatives

- SOFA-InterMedS (2009–) Participants: R. Anxionnat, M.O. Berger, E. Kerrien, A. Yureidini. The SOFA-InterMedS large-scale INRIA initiative is a research-oriented collaboration across several INRIA project-teams, international research groups and clinical partners. Its main objective is to leverage specific competences available in each team to further develop the multidisciplinary field of Medical Simulation research. Our action within the initiative takes place in close collaboration with both Shacra INRIA project-team in Lille and the Department of diagnostic and therapeutic interventional neuroradiology of Nancy University Hospital. We aim at providing in-vivo models of the patient’s organs, and in particular a precise geometric model of the arterial wall. Such a model is used by Shacra team to simulate the coil deployment within an intracranial aneurysm. The associated medical team in Nancy, and in particular our external collaborator René Anxionnat, is in charge of validating our results.

- ANR ARTIS (2009-2012) Participants: M.O. Berger, A. Eryildirim, E. Kerrien. The main objective of this fundamental research project is to develop inversion tools and to design and implement methods that allow for the production of augmented speech from the speech sound signal alone or with video images of the speaker’s face. The Magrit team is especially concerned with the development of procedures allowing for the automatic construction of a speaker’s model from various imaging modalities.

- ANR Visac (2009-2012) Participants: M.O. Berger, B. Wrobel-Dautcourt. The ANR Visac is about acoustic-visual speech synthesis by bimodal concatenation. The major challenge of this project is to perform speech synthesis with its acoustic and visible components simultaneously. Within this project, the role of the Magrit team is twofold. One of them is to build a stereovision system able to record synchronized audio-visual sequences at a high frame rate. Second, a highly realistic dense animation of the head must be produced.

6.4. European Initiatives

6.4.1. Major European Organizations with which you have followed Collaborations

Partner 1: Imperial College, London.
Pierre-Frédéric Villard has a Honorary Research Fellow contract with Imperial College. The research focusing on medical simulators based on task analysis have been done within this link. The collaboration has involved 2 research visits in London to mainly incorporate work done in Lorraine both at the LORIA and with Nancy University intern students. There was also a participation as an activity leader in a one-week summer school on Haptic Technology (to give the basics of computer haptics, including visual and haptics rendering, force feedback, haptic interfaces, collision detection, collision response and deformation modelling).
5. New Results

5.1. Decision Making

5.1.1. Optimizing Automated Service Discovery

**Participant:** Jörg Hoffmann.

*Michael Stollberg (SAP Research, Germany) and Dieter Fensel (University of Innsbruck, Austria) are external collaborators.*

We completed earlier work, done while all authors were employed at the University of Innsbruck, and published it in the International Journal of Semantic Computing [10]. In a nutshell, the work proposes to use first-order logic for annotating web services to accomplish better precision and recall in service discovery; its core contribution is a technique making such discovery more effective – discovery here involves first-order logical reasoning – by designing a caching technique storing known relationships between available services and possible discovery queries.

5.1.2. Overview of Semantic Web Service Technologies

**Participant:** Jörg Hoffmann.

*Stijn Heymans (SemanticBits, USA), Annapaola Marconi (Fondazione Bruno Kessler, Trento, Italy), Joshua Phillips (SemanticBits, USA), and Ingo Weber (University of New South Wales, Sydney, Australia) are external collaborators.*

We were invited to write a book chapter about the basic AI technologies underlying semantic Web service discovery and composition. The chapter has been published as part of a book entitled “Handbook of Service Description – USDL and its Methods” in Springer-Verlag [46].

5.1.3. Analyzing Planning Domains to Predict Heuristic Function Quality

**Participant:** Jörg Hoffmann.

The heuristic search approach to planning (cf. the above) rises and falls with the quality of the heuristic estimates. The dominant method, especially in satisficing (non-optimal) planning, is to approximate a heuristic function called $h^+$ – this is used in almost every state of the art satisficing planning system. In earlier work, Jörg Hoffmann showed that $h^+$ has some amazing qualities, in many traditional planning benchmarks, in particular pertaining to the complete absence of local minima. [62] His proofs of this are hand-made, raising the question whether such proofs can be lead automatically by domain analysis techniques. The possible uses of such analysis are manifold, e.g., for automatic configuration of hybrid planners or for giving hints how to improve the domain design. The question has been open since 2002. A serious attempt of Jörg Hoffmann resulted in disappointing results – his analysis method has exponential runtime and succeeds only in two extremely simple benchmark domains. In contrast to this, in our work here we answer the question in the affirmative. We establish connections between certain easily testable syntactical structures, called “causal graphs”, and $h^+$ topology. This results in low-order polynomial time analysis methods, implemented in the Torchlight tool, cf. Section 4.2. Of the 12 domains where Hoffmann proved the absence of local minima, TorchLight gives strong success guarantees in 8 domains. Empirically, its analysis exhibits strong performance in a further 2 of these domains, plus in 4 more domains where local minima may exist but are rare. We show that, in this way, TorchLight can distinguish Hoffmann’s “easy” domains from the “hard” ones. By summarizing structural reasons for analysis failure, TorchLight also provides diagnostic output pin-pointing potentially problematic aspects of the domain. A conference paper on this work was published at ICAPS 2011 [25], and nominated for the best paper award there. A journal paper was published in the Journal of AI Research (JAIR) [9].
5.1.4. Relaxing Bisimulation for State Aggregation in the Computation of Lower Bounds

Participant: Jörg Hoffmann.

Raz Nissim (Ben-Gurion University, Beer-Sheva, Israel) and Malte Helmert (University of Freiburg, Germany) are external collaborators.

Like the previous line of work, this addresses planning as heuristic search, specifically the automatic generation of heuristic estimates. This is also the core question investigated in the BARQ project, see below. In preparation of this project, we are conducting this line of research, which explores some of the most basic ideas behind BARQ. The basic technique under consideration was developed in prior work outside INRIA. [61] The heuristic estimates are lower bounds generated from a quotient graph in which sets of states are aggregated into equivalence classes. A major difficulty in designing such classes is that there are exponentially many states. Despite this, our technique allows explicit selection of individual states to aggregate, via an incremental process interleaving it with state space re-construction steps. We have shown previously that, if the aggregation decisions are perfect, then this technique dominates the other known related techniques, and sometimes produces perfect estimates in polynomial time. But how to take these decisions? Little is known about this as yet. In the present work, we start from the notion of a “bisimulation”, which is a well-known criterion from model checking implying that the quotient system is behaviorally indistinguishable from the original system – in particular, the cost estimates based on a bisimulation are perfect. However, bisimulations are exponential even in trivial planning benchmarks. We observe that bisimulation can be relaxed without losing any information as far as the cost estimates are concerned. Namely, we can ignore the “content of the messages sent”, i.e., the state transition labels. Such relaxed bisimulations are often exponentially smaller than the original ones. We show to what extent such relaxation can be applied also within our incremental construction process. As a result, in several benchmarks we obtain perfect estimates in polynomial time, and we significantly increase the set of benchmark instances that can be solved with this approach. Indeed, the approach obtained a 2nd place in the optimal track of the 2011 International Planning Competition, and was part of the 1st-prize winning portfolio. A conference paper was published at IJCAI 2011 [28], and a journal paper is under preparation for submission to the Journal of the ACM.

5.1.5. Relaxing Bisimulation by Choosing Transition Subsets

Participants: Michael Katz, Jörg Hoffmann.

Malte Helmert (University of Freiburg, Germany) is an external collaborator.

This line of work builds on the previous one by designing new methods for relaxing bisimulations. The key idea is to apply the bisimulation property to only a subset of the transitions in the system under consideration. We showed that one can ignore large subsets of transitions without losing any information, i.e., while still guaranteeing to obtain a perfect heuristic. At the same time, such a relaxed bisimulation makes less distinctions and may thus be exponentially smaller. For practical purposes, we designed several approximate strategies relaxing more, obtaining smaller abstractions at the expense of information loss. The techniques are currently being evaluated empirically, and a paper submission is in preparation for ICAPS’12.

5.1.6. Improving $h^+$ by Taking Into Account (Some) Negative Effects

Participants: Emil Keyder, Jörg Hoffmann.

Patrik Haslum (NICTA, Australia) is an external collaborator.

Like the previous lines, this is on planning as heuristic search. As mentioned above in Section 5.1.3, approximating the $h^+$ heuristic is the dominant approach to obtain estimates in satisficing (non-optimal) planning. That notwithstanding, $h^+$ is obtained by ignoring all negative effects, which of course leads to very bad estimates in domains where these domains play a key role, for example puzzle-like domains, e.g. Rubic’s cube, where actions interfere intensively with each other. It has long (for almost 10 years) been an active research issue how to take at least some of the negative effects into account when computing $h^+$. All attempts, however, remained at rather ad-hoc methods, like, counting the number of violated binary constraints (pairs of facts that cannot be true at the same time) within the relaxed plan underlying the estimate. In the present work,
for the first time we provide a well-founded formal approach to the issue. As was suggested in prior work, \[ 60 \], we design a compiled planning task which introduces constructs allowing \( h^+ \) to correctly handle a subset \( C \) of fact conjunctions. Whereas this prior work requires a compilation exponential in \( |C| \) – and thus allows only to introduce very few conjunctions – in our work we designed a compilation that is linear in \( |C| \). We proved that one can always choose \( C \) so that \( h^+ \) in the compiled task is a perfect heuristic. Of course, in general \( C \) might have to be exponentially large to achieve this. We designed practical methods selecting \( C \) in a way so that the overhead (the size of \( C \)) is kept at bay, while the quality of the heuristic is sufficiently improved to boost search performance. The techniques are currently being evaluated empirically, and a paper submission is in preparation for ICAPS’12.

5.1.7. Accounting for Uncertainty in Penetration Testing

**Participants:** Olivier Buffet, Jörg Hoffmann.

Carlos Sarraute (Core Security Technologies) is an external collaborator.

Core Security Technologies is an U.S.-American/Argentinian company providing, amongst other things, tools for (semi-)automated security checking of computer networks against outside hacking attacks. For automation of such checks, a module is needed that automatically generates potential attack paths. Since the application domain is highly dynamic, a module allowing to declaratively specify the environment (the network and its configuration) is highly advantageous. For that reason, Core Security Technologies have been looking into using AI Planning techniques for this purpose. After consulting by Jörg Hoffmann (see also Section 6.1.1 below), they are now using a variant of Jörg Hoffmann’s FF planner (cf. Section 4.1) in their product. While that solution is satisfactory in many respects, it also has weaknesses. The main weakness is that it does not handle the incomplete knowledge in this domain – figuratively speaking, the attacker is assumed to have perfect information about the network. This results in high costs in terms of runtime and network traffic, for extensive scanning activities prior to planning. We are currently working with Core Security’s research department to overcome this issue, by modeling and solving the attack planning problem as a POMDP instead. A workshop paper detailing the POMDP model has been published at SecArt’11 [29]. While such a model yields much higher quality attacks, solving an entire network as a POMDP is not feasible. We have designed a decomposition method making use of network structure and approximations to overcome this problem, by using the POMDP model only to find good-quality attacks on single machines, and propagating the results through the network in an appropriate manner. A conference paper is in preparation for submission to ICAPS’12.

5.1.8. Searching for Information with MDPs

**Participants:** Mauricio Araya, Olivier Buffet, Vincent Thomas, François Charpillet.

In the context of Mauricio Araya’s PhD, we are working on how MDPs—or related models—can search for information. This has led to various research directions that we describe now.

A POMDP Extension with Belief-dependent Rewards — A limitation of Partially Observable Markov Decision Processes (POMDPs) is that they only model problems where the performance criterion depends on the state-action history. This excludes for example scenarios where one wants to maximize the knowledge with respect to some random variables.

To overcome this limitation, we have proposed \( \rho \)-POMDPs, an extension of POMDPs in which the reward function depends on the belief state rather than on the state. In this framework, and under the hypothesis that the reward function is convex, we have proved that:

- the value function itself is convex; and
- if the reward function is \( \alpha \)-Hölder, then the value function can be approximated arbitrarily well with a piecewise linear and convex function.

These results allow for adapting a number of solution algorithms relying on approximating the value function.
This theoretical work has been first published in an international conference in December 2010, then in [36], where it has received a best paper award.

We are currently pursuing experimental work about the proposed algorithm.

Active Learning of MDP Models — Reinforcement Learning is about learning how to perform a task by trial and error (no model of the system to control being available). Model-based Bayesian RL (BRL) consists in all RL algorithms that maintain a belief (in the Bayesian sense) about the model of the system to control. In fact, this is a way to turn an RL problem into a POMDP—the unknown model becoming an unobservable part of the state—, thus replacing the exploration-exploitation dilemma by the definition of a prior belief over possible models.

A particular BRL task we have been considering is to actively learn the dynamical model itself, i.e., to act so as to improve the knowledge about the transition function. In a way this means solving a $\rho$-POMDP since the reward depends on a belief, not on a state. To that end, we have proposed several optimization criteria, and derived the corresponding reward functions, making sure that their computational complexity allows for their use in a BRL algorithm. We have also proved that a non-optimistic BRL algorithm—EXPLOIT—could be used in this particular case.

This work, along with experiments, has been published in [36] and [35] (French version).

PAC-BAMDP Algorithms — Exact or approximate solutions to Model-based Bayesian RL are impractical, so that a number of heuristic approaches have been considered, most of them relying on the principle of “optimism in the face of uncertainty”. Some of these algorithms have properties that guarantee the quality of their outcome, inspired by the PAC-learning (Probably Approximately Correct) framework. For example, some algorithms provably make in most cases the same decision as would be made if the true model were known (PAC-MDP property).

We have proposed a novel optimistic algorithm, BOUH, that is

- appealing in that it is (i) optimistic about the uncertainty in the model and (ii) deterministic (thus easier to study); and
- provably PAC-BAMDP, i.e., makes in most cases the same decision as a perfect BRL algorithm would.

First results about this algorithm are currently under review.

5.1.9. Scheduling for Probabilistic Realtime Systems

Participant: Olivier Buffet.

Maxim Dorin, Luca Santinelli, Liliana Cucu-Grosjean (INRIA, TRIO team), and Rob Davies (U. of York) are external collaborators.

In this collaborative research work (mainly with the TRIO team), we look at the problem of scheduling periodic tasks on a single processor, in the case where each task’s period is a (known) random variable. In this setting, some job will necessarily be missed, so that one will try to satisfy some criteria depending on the number of deadline misses.

We have proposed three criteria: (1) satisfying pre-defined deadline miss ratios, (2) minimizing the worst deadline miss ratio, and (3) minimizing the average deadline miss ratio. For each criterion we propose an algorithm that computes a provably optimal fixed priority assignment, i.e., a solution obtained by assigning priorities to tasks and executing jobs by order of priority.

This work has been presented in [26].

We also collaborate on other topics linked to real-time scheduling, as (i) on search algorithms for deterministic, but multiprocessor, problems [38], and (ii) on the problem of which jobs to drop (on-going work).

5.1.10. Adaptive Management with POMDPs

Participant: Olivier Buffet.
Iadine Chadès, Josie Carwardine, Tara G. Martin (CSIRO), Samuel Nicol (U. of Alaska Fairbanks) and Régis Sabbadin (INRA) are external collaborators.

In the field of conservation biology, adaptive management is about managing a system, e.g., performing actions so as to protect some endangered species, while learning how it behaves. This is a typical reinforcement learning task that could for example be addressed through BRL.

Here, we consider that a number of experts provide us with one possible model each, assuming that one of them is the true model. This allows making decisions by solving a mixed observability MDP (MOMDPs), where the hidden part of the state corresponds to the model (in cases where all other variables are fully observable).

We have conducted preliminary studies of this approach, using the scenario of the protection of the Gouldian finch, and focusing on the particular characteristics that could be exploited to more efficiently solve this problem. First results have been presented in [ 39 ].

5.1.11. Information Gathering with Sensor Systems

Participant: Olivier Buffet.

Elodie Chanthery, Matthieu Godichaud (LAAS-CNRS) and Marc Contat (EADS) are external collaborators.

The DOPEC project was a DGA PEA (upstream studies project) on the optimization of the use of sensor systems. In collaboration with EADS (project leader) and the LAAS, we have worked on autonomous sequential decision making problems. We were more particularly interested, on the one hand, in multi-agent problems and, on the other hand, in taking uncertainties into account.

The overall architecture that has been developed in the context of this project was presented in a national and an international conference [ 40 ], [ 23 ].

5.1.12. How do real rats solve non-stationary (PO)MDPs?

Participant: Alain Dutech.

Etienne Coutureau and Alain Marchand (Centre de Neurosciences Intégratives et Cognitives (CNIC), UMR 5228, Bordeaux) are external collaborators.

For a living entity, using simultaneously various ways for learning models or representations of its environment can be very useful to adapt itself to non-stationary environments in a Reinforcement Learning setting. In the rats and in the monkey, two different action control systems lie in specific regions of the prefrontal cortex. Neurobiologists and computer scientists find here a common ground to identify and model these systems and the selection mechanisms between them, selection that could depend on uncertainty or error signals. Using real data collected on rats with or without prefrontal lesions, reinforcement learning models are used and evaluated in order to better understand this behavioral flexibility. MAIA is more particularly involved as a reinforcement learning expert in order to suggest and build models of the various learning mechanisms. In particular, we have used an on-policy learning scheme (SARSA) to investigate how well the use of simple or complex representations (with or without memory of the immediate past) can best model the learning behavior of rats in instrumental contingency degradation tasks [ 7 ].

This work has led us to investigate in more details the relations between the prefrontal cortex and the basal ganglia and their respective role when rats learn to solve non-stationnary tasks. The research is conducted through the PEPII project IMAVO (see 7.2.7 ).

5.1.13. Developmental Reinforcement Learning

Participants: Alain Dutech, Olivier Buffet.

Luc Sarzyniec and Joël Legrand (M2R Student of UHP Nancy 1) are external collaborators.
The goal of this work is to investigate how reinforcement learning can benefit from a developmental approach in the field of robotics. Instead of having a robot directly learn a difficult task using appropriate but rich (in the number of dimensions) sensory and motor spaces, we have followed an incremental approach. Both the number of perception and action dimensions increase only when the performance of the learned behavior increases. At the core of the algorithm lies a neuronal approximator used to compute the value function of the current policy of the robot. When the perception or action space grow, neurons or networks, initialized from existing neurons and networks, are added to the control architecture.

Thus far, our research focused on the approximation architecture used to evaluate the $Q$-function. In simple robotic tasks, we investigated the use of Multi-Layer Perceptrons, either one approximation for every possible action ([41]) or one unique global approximator with as many outputs as the number of actions (Master Thesis of Joël Legrand). Currently, a reservoir computing architecture is under study as depicted in [16].

5.1.14. Classification-based Policy Iteration with a Critic

**Participant:** Bruno Scherrer.

Victor Gabillon, Alessandro Lazaric and Mohammad Ghavamzadeh (from Sequel INRIA-Lille) are external collaborators.

We study the effect of adding a value function approximation component (critic) to rollout classification-based policy iteration (RCPI) algorithms. The idea is to use the critic to approximate the return after we truncate the rollout trajectories. This allows us to control the bias and variance of the rollout estimates of the action-value function that are strongly related to the length of the rollout trajectories. Therefore, the introduction of a critic can improve the accuracy of the rollout estimates, and as a result, enhance the performance of the RCPI algorithm. We present in [49], [20] a new RCPI algorithm, called direct policy iteration with critic (DPI-Critic), and provide its finite-sample analysis when the critic is based on LSTD and BRM methods. We empirically evaluate the performance of DPI-Critic and compare it with DPI and LSPI in two benchmark reinforcement learning problems.

5.1.15. Linear Approximation of Value Functions

**Participant:** Bruno Scherrer.

Matthieu Geist (Supélec, Metz) is an external collaborator.

In the framework of Markov Decision Processes, we consider the problem of learning a linear approximation of the value function of some fixed policy from one trajectory possibly generated by some other policy.

In [30], [42], [51], we describe a systematic approach for adapting on-policy learning least squares algorithms of the literature (LSTD, LSPE, FPKF and GPTD/KTD) to off-policy learning with eligibility traces. This leads to two known algorithms, LSTD($\lambda$)/LSPE($\lambda$) and suggests new extensions of FPKF and GPTD/KTD. We describe their recursive implementation, discuss their convergence properties, and illustrate their behavior experimentally. Overall, our study suggests that the state-of-art LSTD($\lambda$) remains the best least-squares algorithm.

We also consider the task of feature selection. A promising approach consists in combining the Least-Squares Temporal Difference (LSTD) algorithm with $\ell_1$-regularization, which has proven to be effective in the supervised learning community. This has been done recently with the LARS-TD algorithm, which replaces the projection operator of LSTD with an $\ell_1$-penalized projection and solves the corresponding fixed-point problem. However, this approach is not guaranteed to be correct in the general off-policy setting. In [21], we take a different route by adding an $\ell_1$-penalty term to the projected Bellman residual, which requires weaker assumptions while offering a comparable performance. This comes at the cost of a higher computational complexity if only a part of the regularization path is computed. Nevertheless, our approach ends up to a supervised learning problem, which let envision easy extensions to other penalties.
5.2. Understanding and mastering complex systems

5.2.1. Complex systems: simulation, control and definition

5.2.1.1. Adaptive control of a complex system based on its multi-agent model

Participants: Vincent Chevrier, Tomas Navarrete.

Laurent Ciarletta (Madynes team, LORIA) is an external collaborator.

We are interested in how to build a control mechanism for a complex/dynamic system. Specifically, we want to evaluate the effectiveness of creating a control mechanism based on a multi-agent model of the system[12]. Multi-agent models can be adapted to that purpose since usual approaches using analytical models as basis can be untractable when dealing with such systems; and because if we consider that the available control actions are meant to be applied locally, a multi-agent model is necessary. We are currently working on a case study within the dynamic networks domain, namely the free-riding phenomenon present in peer-to-peer networks.

We propose an architecture that gathers information from the system and uses it to parametrize and tune a set of multi-agent models. The outcome of simulations is used to decide which control actions have to be applied to the system, in order to achieve a predefined control objective. We consider that we do not have complete information to characterize the state of the system and hence would like to focus on the following two issues of the control problem that we have identified:

1. How to build a multi-agent model that represents the evolution of a dynamic network. That is, what to do when the information given by the simulation of the multi-agent is in contradiction with the information gathered from the system
2. How to build an adaptive control mechanism based on the multi-agent model of a dynamic network. That is, how to use the information given by the multi-agent model to achieve the control objective.

The architecture we proposed, is designed as a control loop composed of the following steps: estimate the state of the system and instantiate multi-agent models accordingly, simulate different control actions, choose a control action and apply it. From one cycle to another of the control loop, each step can be tuned (in terms of model parameters, control action selection process, sampling strategy, etc.) to overcome the previously mentioned issues of the control problem.

The architecture is currently specified in terms of a formal notation. We have already implemented the architecture within the context of the free-riding problem where we use the PeerSim simulator as the target system to control.

Within our case study, we have conducted two different sets of experiments to investigate under which conditions our control architecture can achieve its goal and to investigate the efficiency of different sampling methods to estimate the state of the network. We have effectively managed to drive the system to a state where the majority of the peers share, when the initial conditions, without intervention from our architecture, would drive the system to a state where no peer would share.

The elements of the architecture having an impact on the performance of the control obtained have been identified. These are: the initialization of the parameters of the models used to estimate the state of the system, predict the evolution of the system and test the possible control actions, as well as the strategy used to observe the system and the different time horizons to consider within the architecture.

The next steps are to better identify the advantages and limits of the proposed architecture and to widen the problem family in the free riding problem.

5.2.1.2. Multi Modeling and multi-simulation

Participants: Vincent Chevrier, Julien Siebert.

This work is undertaken in a joint PhD Thesis between MAIA and Madynes Team. Laurent Ciarletta (Madynes team, LORIA) is co-advisor of this PhD.
Complex systems generally require to use different points of view (abstraction levels) at the same time on the system in order to capture and to understand all the dynamics and the complexity. Being made of different interacting parts, a model of a complex system also requires simultaneously modeling and simulation (M&S) tools from different scientific fields.

Building a model and a simulation of a complex system from the interaction of the different existing M&S tools present in each scientific field involved, is also a complex task. To represent a complex system, we need to couple several models (multi-modeling) that each represents a part of the whole system. Each model could have been designed by and for a specific scientific domain. Making different models interact raises hard issues on model interoperability (semantic coherence, formalism compatibility). As many simulators exist in the scientific fields involved, a possible approach to make a simulation of a complex system is to reuse and to make interact these existing simulators. Since each simulator has been developed for specific purposes, making them interact (multi-simulation) raises simulation issues (interoperability, synchronization).

The multi-agent paradigm is an homogeneous solution both for multi-modeling and multi-simulation of complex systems. On the one hand, a multi-agent model per se is a multi-model: a multi-agent model is made of interacting agent models and environment models. On the other hand, agent oriented software engineering (AOSE) allows designers to create complex softwares as a set of autonomous, heterogeneous and interacting softwares (i.e. as a multiagent system). Robustness, scalability, openness, modularity and interoperability are some of the properties that AOSE allows to achieve.

This work explores the contribution of multiagent paradigm to the fields of multi-modeling and multi-simulation of complex systems.

The first contribution of this work is to propose an homogeneous multiagent meta-model (called AA4MM[4]) that provides solutions both for multi-modeling and multi-simulation of complex systems by reusing existing and heterogeneous M&S tools. The core idea in AA4MM is to build a society of models, simulators and simulation softwares that solves the core challenges of multimodelling and simulation coupling in an homogeneous perspective. AA4MM has been implemented and used both for proof of concept and for a real case study. A proof of concept has been made by coupling different models together to develop a multi-model of a prey-predator model. This has permitted us to show both conceptual and operational properties of AA4MM such as interaction of heterogeneous models, modularity, interoperability.

This multiagent meta-model has been applied to model complex systems that are ubiquitous networks. Ubiquitous networks are highly dynamic computer networks that are composed of a great number of interacting and sometimes mobile nodes which can join or leave the system, interact together and where the environment plays a significant role either on radio communications or on the behavior of users. Modeling and simulation is the approach to evaluate these technologies or to build new ones.

5.2.1.3. Robustness of Cellular Automata and Reactive Multi-Agent Systems

Participants: Olivier Bouré, Vincent Chevrier, Nazim Fatès.

Our research on emergent collective behaviours focuses on robustness analysis, that is the behavioural resistance to perturbations in collective systems. We progressed in the knowledge of how to tackle this issue in the case of cellular automata (CA) and multi-agent systems (MAS).

We focused on the specific case of a perturbation of the updating scheme in CA, that is, changing the way cells are updated. Using similar ideas to the Influence-Reaction principle developed to resolve conflicts related to simultaneous actions, we created a new type of asynchronism, called beta-synchronism, which aims at disrupting the transmission of information about states between cells. We found out that the different types of asynchronism may induce radical change of behaviour for particular a value range of the synchrony rate [15].

More recently, our interest focused on a bio-inspired discrete dynamical system. Using the formalism of a subclass of cellular automata, lattice-gas CA, we study a model of swarming which displays qualitatively different behaviours under certain experimental conditions. We discussed these observations by relating them to the potential links with certain attributes of the model [48].
We studied a phase transition that occurs in the Greenberg-Hastings CA reaction-diffusion \[ 5 \]. The density classification problem was taken as a typical framework for studying how decentralised computations can be carried out with simple cells. Although it is known that this problem can not be solved perfectly, we showed that using randomness provides a solution with an arbitrarily high success rate \[ 17 \]. We also studied how to extend this result to the infinite-space case citerefPapier ???.

We studied the behaviour of the amoebae aggregation model \[ 33 \] and applied the aggregation scheme on a robotic case (ALICE robots and Khepera III with Romea interactive table).

5.2.1.4. Ant algorithms for multi-agent patrolling

**Participants:** Olivier Simonin, François Charpillet, Olivier Buffet, Arnaud Glad.

We proposed in 2007 an ant algorithm, called EVAP, to deal with multi-agent patrolling, which is based on the marking and the evaporation of a digital pheromone. During the simulations carried out to measure the performances of EVAP, we identified that the system can self-organize towards stationary cycles (a periodic attractor). These cycles correspond to an Hamiltonian or quasi-Hamiltonian covering of the environment, which is an optimal or quasi-optimal solution to the multi-agent patrolling problem. We then established the mathematical proof that the system can stabilize only in cycles, one per agent, having the same length (cf. publication in ECAI’2008). Moreover, we introduced new heuristics in the agent behavior that improve dramatically the time for convergence, and we proved that under deterministic hypotheses the system always converges to stable cycles (these results have been published in SASO 2009, AAMAS’10). Results of 2011 are:

- Defense of Arnaud Glad’s PhD. thesis (November 15th) synthesising theoretical and experimental studies of the EVAP algorithm. The writing of a journal article is also in progress.
- EVAP has been adapted to continuous space in the context of the SUSIE project, which consider the surveillance of an area with a set of autonomous aerial robots.

5.2.2. Multi-robot systems : swarm intelligence, cooperation, navigation

5.2.2.1. Multi-robot exploration and mapping

**Participants:** Olivier Simonin, François Charpillet, Antoine Bautin.

In the context of the ANR Cartomatic project, introduced in Sec. 7.2.3, we study multi-agent models for multi-robot deployment and mapping. This work is in line with the PhD thesis of Antoine Bautin, started in November 2009. New results of 2011 are:

- A new frontier assignation algorithm for multi-robot exploration has been proposed. It relies on counting the number of robots towards a frontier rather than considering only distances between robots and frontiers. We measured on benchmarks that the approach outperforms the two classical algorithms closest frontier and Greedy approach. Results are presented in \[ 37 \], \[ 43 \] and are submitted to ICRA’2012.
- We implemented and experimented the approach with autonomous mobile robots in the context of the ANR Carotte challenge (June 2011, Bourges). Our team “cartomatic” obtained one of the best map of the contest, while deploying several robots.

5.2.2.2. New experimental device: the Interactive Table

**Participants:** Olivier Simonin, François Charpillet, Nicolas Beaufort.

Olivier Rochel (INRIA research engineer, SED Nancy) is an external collaborator.

During 2010 we developed with the Nancy INRIA SED\(^6\) (Olivier Rochel) a new experimental device dedicated to swarm robotics study. It is composed of two independent components : an interactive table able to display and to compute any active environment and a set of autonomous mobile robots able to read and write information on the environment.

\(^6\)Service d’Expérimentation et Développement
Studies using the Table in 2011 are:

- We revisited the Drogoul & Ferber Foraging model, inspired by ants and also called “robot dockers” as the agents exchange the transported resources when they meet. From this simulated model we examined how it can be implemented with real mobile robots on an interactive environment, by considering that robots drop pheromones as ants. We defined a model extending the docker model with the robots on the Table, and studied its robustness to perception failure/mistakes. This work, done with Thomas Huraux (Master 2 Recherche internship), has been published in ICTAI 2011 Int. Conference [31].
- Several students (from Science Cog. Nancy 2 Master) implemented and explored pheromone-based foraging behaviors and flocking-based navigation models (supervised by François Charpillet and Christine Bourjot).

5.2.2.3. Local control based platooning

**Participants:** Alexis Scheuer, Olivier Simonin, François Charpillet, Jano Yazbeck.

We consider decentralised control methods to operate autonomous vehicles at close spacings to form a platoon. We study models inspired by the flocking approach, where each vehicle computes its control from its local perceptions. We investigate different decentralised models in order to provide robust and scalable solutions. Open questions concern collision avoidance, stability and multi-platoon navigation.

- **Coupling lateral and longitudinal controls.** A first work [67] focused on longitudinal control, which aims at computing velocities to avoid collision when all the vehicles are moving along a fixed path. When vehicles move in a two dimensional space, a lateral controller is needed to steer the vehicles. While lateral and longitudinal controls can be considered separately, the longitudinal control should be done after the lateral control: while turning, a higher inter-vehicle distance is needed to avoid collisions.

An innovative approach to improve the quality of lateral control has been proposed during Jano Yazbeck’s internship at LORIA (03/10–07/10), entitled “Decentralised local approach for lateral control of platoons” and supervised by A. Scheuer and O. Simonin. This allows to reduce the distance between each vehicle’s path and the path of the previous vehicle, by using only embedded sensors such as a laser rangefinder. It relies on memorizing and computing in real time the previous vehicle relative trajectory. This work has been published in 2011 IEEE-RSJ International Conference on Intelligent Robots and Systems (IROS’2011) [34].

- **Finding an efficient lateral control.** To obtain an even better lateral control, and to drive each vehicle exactly in the trace of the previous one, we are developing a more efficient lateral control law. This law is defined in order to reduce exponentially the tracking error (which is more or less the distance between each vehicle’s path and the path of the previous vehicle). Once again, as for the longitudinal control [67], the formula of the control law is obtained through the proof of its property: necessary conditions are simplified in order to get the final result.

5.2.2.4. Adaptation of autonomous vehicle traffic to perturbations

**Participants:** Mohamed Tlig, Olivier Simonin, Olivier Buffet.

In the context of the European InTraDE project, one problem is to handle the displacements of numerous IAVs7 in a seaport. Here we assume a supervisor planning the routes of the vehicles in the port. However, in such a large and complex system, different unexpected events can arise and damage the traffic: failure of a vehicle, human mistake while driving, obstacle on roads, local re-planning, and so on.

We started focusing on a first important sub-problem of space resource sharing among multiple agents: how to ensure the crossing of two opposed flows of vehicles on a road when one of the two paths is blocked by an obstacle, e.g., a disabled vehicle. To overcome this problem, blocked vehicles have to coordinate with vehicles of the other side to share the road and manage delays. The objective is to improve traffic flow and reduce the emergence of traffic jam.

---

7 Intelligent Autonomous Vehicle
Solving this problem with reactive coordination methods is a major challenge of the PhD thesis of Mohamed Tlig (started in December 2010).

- We started by formalizing the problem and the possible actions of agents (vehicles) following a STRIPS formalism. We adapted this model dedicated to planning to the description of local rules in reactive coordination.
- We then defined and studied in simulation two decision rules that produce two different strategies: the first one alternates between two vehicles from each side of the road, and the second one gives priority to the vehicle with the highest delay. We are preparing a publication of these first results.

5.2.3. Ambient intelligence and Actimetry

5.2.3.1. Robotics and spatial computing : the iTiles - intelligent tiles - model

Participants: Olivier Simonin, François Charpillet, Lionel Havet.

Olivier Rochel (INRIA research engineer, SED Nancy) is an external collaborator.

In the context of intelligent home and assistant robots, we explore the definition and use of an active floor based on a cellular network approach. We aim at exploring spatial calculus models when considering physical cells augmented with sensors where robots and humans can evolve. Since 2009, we study a model consisting in paving the floor with interconnected tiles. Each tile can communicate with its neighbors and can sense the presence of a robot or a human. A first Tile model has been defined and evaluated using a tiles emulator and real mobile robots (Kheperas III), which validated the interest of the approach. See CAR’2010 publication [68].

In 2011 we designed with the help of INRIA Grenoble SED a prototype of 9 physical tiles embedding a WSN node able of computation and communication with other tiles. From this experimental device we explored several questions:

- How to follow a person walking on such a discrete and sensitive floor? We proposed a set of distributed algorithms allowing tiles to track a person or a robot (cooperation between neighboring tiles).
- How to make communications between a robot and the tile(s) it occupies? We developed a set of functions using the wifi communication of the tiles and the robot.
- From the work mentioned in the previous items we propose the definition of a new tile network based on the SensLab technology (wire connections between tiles and wireless communications between tiles and robots/humans). This prototype has been ordered and will be installed in the beginning of 2012.

5.2.3.2. Bayesian 3D Human Motion Capture Using Factored Particle Filtering

Participants: Abdallah Dib, Cédric Rose, Amandine Dubois, François Charpillet.

The gait deterioration of elderly people is an important factor in loss of autonomy and it increases the risk of falls. In order to evaluate this risk the MAIA team has been developing since 2003 a markerless human motion capture system that estimates the 3D positions of the body joints over time. The system uses a dynamic Bayesian network and a factored particle filtering algorithm. This year, we have evaluated the impact of using different observation functions for the Bayesian state estimation: chamfer distance, a pixel intersection and finally a pseudo-observation of the subject direction calculated from the previous output of the system. We also compared two methods for the factored generation of the particles. The first one uses a deterministic interval exploration strategy whereas the second one is based on an adaptive diffusion. The capacity of the system to recover after occlusion by obstacles was tested on simulated movements in a virtual scene [57].
An other achievement of the year has been the assessment of the accuracy and precision of this system, especially for measuring the step length of a walking human. This has been realized by Amandine Dubois during her research master [58]. An experiment with young subjects has been designed and realized. Measures of the markerless motion capture system were then compared with real values. These values were obtained through the footprints left by the subjects. Ink swabs placed at the front and rear of the shoes of each subject make it possible to mark a paper strip positioned on the ground. A statistical analysis of the results has been done by Amandine. Thus we were able to determine if the real and measured lengths were significantly different or not.

5.2.3.3. Automatic Evaluation of Vascular Access in Hemodialysis Patients

Participants: Cédric Rose, François Charpillet.

The vascular access that allows to perform the extra-corporeal circulation, is usually a vein of the arm that has been enlarged by a surgical creation of a fistula. The prevention of complications such as stenosis or thrombosis of the vascular access is a key issue in hemodialysis treatment. Many dialysis machines measure ionic dialysance by conductivity measures on the dialysate fluid. Ionic dialysance is an indicator of small molecules transfers through the dialysis membrane. Previous works have shown that the follow-up of the dialysance and the pressures along the extra corporeal circuit can help to detect at an early stage a potential complication on the vascular access. The difficulty of automating the follow-up is the large variability of the measures and the need to detect tendencies. Dynamic Bayesian networks (DBN) allow to formalize expert knowledge as a graphical stochastic model adapted to reasoning under uncertainty. In a DBN the state of the patient and the measurements are represented by interconnected temporal random variables. The relations between those variables are described using probability distributions. The proposed approach [64] is based on a supervised learning of a DBN for classifying the dialysis sessions according to a risk score describing the medical situation (0: no risk, 1: mild risk, 2: severe risk). The training of the system was performed using a dataset labeled by a medical expert. The evaluation of the results was done by performing a double-blind analysis of real data. The result was an 85% agreement rate between the human expert and the automated analysis. The purpose of the system is to assist the human expert by reporting abnormalities. The results show that a score 2 reported by the human is rarely missed by the automated analysis (only 1 case) whereas the opposite is more frequent (8 cases). The final decision to further investigate a case is taken by the human expert.
6. New Results

6.1. Audio and speech content processing

6.1.1. Audio motif discovery

**Participants:** Frédéric Bimbot, Laurence Catanese, Armando Muscariello.

This work was performed in close collaboration with Guillaume Gravier from the Texmex project-team.

As an alternative to supervised approaches for multimedia content analysis, where predefined concepts are searched for in the data, we investigate content discovery approaches where knowledge emerge from the data. Following this general philosophy, we pursued work on motif discovery in audio contents.

Audio motif discovery is the task of finding out, without any prior knowledge, all pieces of signals that repeat, eventually allowing variability. In 2011, we extended our recent work on seeded discovery to near duplicate detection and spoken document retrieval from examples. First, we proposed algorithmic speed ups for the discovery of near duplicate motifs (low variability) in large (several days long) audio streams, exploiting subsampling strategies [muscariello-cbmi-11]. Second, we investigated the use of previously proposed efficient pattern matching techniques to deal with motif variability in speech data [muscariello-icassp-11] in a different setting, that of spoken document retrieval from an audio example. We demonstrated the potential of model-free approaches for efficient spoken document retrieval on a variety of data sets, in particular in the framework of the Spoken Web Search task of the MediaEval 2011 international evaluation [muscariello-is-11, muscariello-mediaeval-11].

This work is carried out in the context of the Quaero Project.

6.1.2. Landmark-driven speech recognition

**Participant:** Stefan Ziegler.

This work is supervised by Guillaume Gravier and Bogdan Ludusan from the Texmex project-team.

Speech recognition is a key issue to access multimedia spoken contents. In this context, speech recognition faces several challenges among which robustness to acoustic and linguistic variability.

In 2011, we initiated research on landmark-driven speech recognition to increase robustness. The idea of this approach consists in accurately detecting in the signal landmarks corresponding to broad phonetic classes (vowels, nasals, etc.). These landmarks, which represent almost certain knowledge about the phonetic content of the signal, are then used to bias the search space in Viterbi decoding towards solutions consistent with the landmarks. We proposed a landmark detection system, which employs numerous attributes extracted from a segment based representation of speech. We use a decision tree for BPC classification, since this allows the evaluation of each BPC on its most informative attributes, selected from a large variety of attributes. Then, each segment is converted into a landmark and a probability estimate for each BPC is provided. Second, we extend a previously proposed landmark-driven decoding strategy by a more flexible implementation, which reinforces paths at the detected landmarks according to the obtained BPC probabilities. Results obtained on French broadcast news data show a relative improvement in word error rate of about 2 % with respect to the baseline.

6.2. Recent results on sparse representations

The team has had a substantial activity ranging from theoretical results to algorithmic design and software contributions in the field of sparse representations, which is at the core of the FET-Open European project (FP7) SMALL (Sparse Models, Algorithms and Learning for Large-Scale Data, see Section 7.2.1 ) and the ANR project ECHANGE (ECHantillonnage Acoustique Nouvelle GEnération, see, Section 6.3.1 ).
6.2.1. A new framework for sparse representations: analysis sparse models
Participants: Rémi Gribonval, Sangnam Nam.
Main collaboration: Mike Davies (Univ. Edinburgh), Michael Elad (The Technion), Hadi Zayyani (Sharif University)

In the past decade there has been a great interest in a synthesis-based model for signals, based on sparse and redundant representations. Such a model assumes that the signal of interest can be composed as a linear combination of few columns from a given matrix (the dictionary). An alternative analysis-based model can be envisioned, where an analysis operator multiplies the signal, leading to a cosparse outcome. Within the SMALL project, we initiated a research programme dedicated to this analysis model, in the context of a generic missing data problem (e.g., compressed sensing, inpainting, source separation, etc.). We obtained a uniqueness result for the solution of this problem, based on properties of the analysis operator and the measurement matrix. We also considered a number of pursuit algorithms for solving the missing data problem, including an L1-based and a new greedy method called GAP (Greedy Analysis Pursuit). Our simulations demonstrated the appeal of the analysis model, and the success of the pursuit techniques presented. These results have been published in international conferences [64] [63], and a journal paper is in preparation.

6.2.2. Theoretical results on sparse representations and dictionary learning
Participants: Rémi Gribonval, Sangnam Nam, Nancy Bertin.
Main collaboration: Karin Schnass (EPFL), Mike Davies (University of Edinburgh), Volkan Cevher (EPFL), Simon Foucart (Université Paris 5, Laboratoire Jacques-Louis Lions), Charles Soussen (Centre de recherche en automatique de Nancy (CRAN)) Jérôme Idier (Institut de Recherche en Communications et en Cybernétique de Nantes (IRCCyN)), Cédric Herzet (Equipe-projet FLUMINANCE (INRIA - CEMAGREF, Rennes)) Morten Nielsen (Department of Mathematical Sciences [Aalborg]), Gilles Puy, Pierre Vandergheynst, Yves Wiaux (EPFL) Mehrdad Yaghoobi, Rodolphe Jenatton, Francis Bach (Equipe-projet SIERRA (INRIA, Paris)) Boaz Ophir, Michael Elad (Technion) Mark D. Plumbley (Queen Mary, University of London)

Sparse recovery conditions for Orthogonal Least Squares : We pursued our investigation of conditions on an overcomplete dictionary which guarantee that certain ideal sparse decompositions can be recovered by some specific optimization principles / algorithms. This year, we extended Tropp’s analysis of Orthogonal Matching Pursuit (OMP) using the Exact Recovery Condition (ERC) to a first exact recovery analysis of Orthogonal Least Squares (OLS). We showed that when ERC is met, OLS is guaranteed to exactly recover the unknown support. Moreover, we provided a closer look at the analysis of both OMP and OLS when ERC is not fulfilled. We showed that there exist dictionaries for which some subsets are never recovered with OMP. This phenomenon, which also appears with ℓ1 minimization, does not occur for OLS. Finally, numerical experiments based on our theoretical analysis showed that none of the considered algorithms is uniformly better than the other. This work has been submitted for publication in a journal [108].

New links between the Restricted Isometry Property and nonlinear approximations : It is now well known that sparse or compressible vectors can be stably recovered from their low-dimensional projection, provided the projection matrix satisfies a Restricted Isometry Property (RIP). We establish new implications of the RIP with respect to nonlinear approximation in a Hilbert space with a redundant frame. The main ingredients of our approach are: a) Jackson and Bernstein inequalities, associated to the characterization of certain approximation spaces with interpolation spaces; b) a new proof that for overcomplete frames which satisfy a Bernstein inequality, these interpolation spaces are nothing but the collection of vectors admitting a representation in the dictionary with compressible coefficients; c) the proof that the RIP implies Bernstein
inequalities. As a result, we obtain that in most overcomplete random Gaussian dictionaries with fixed aspect ratio, just as in any orthonormal basis, the error of best \(m\)-term approximation of a vector decays at a certain rate if, and only if, the vector admits a compressible expansion in the dictionary. Yet, for mildly overcomplete dictionaries with a one-dimensional kernel, we give examples where the Bernstein inequality holds, but the same inequality fails for even the smallest perturbation of the dictionary. This work has been submitted for publication in a journal [102]

**Performance guarantees for compressed sensing with spread spectrum techniques**: We advocate a compressed sensing strategy that consists of multiplying the signal of interest by a wide bandwidth modulation before projection onto randomly selected vectors of an orthonormal basis. Firstly, in a digital setting with random modulation, considering a whole class of sensing bases including the Fourier basis, we prove that the technique is universal in the sense that the required number of measurements for accurate recovery is optimal and independent of the sparsity basis. This universality stems from a drastic decrease of coherence between the sparsity and the sensing bases, which for a Fourier sensing basis relates to a spread of the original signal spectrum by the modulation (hence the name “spread spectrum”). The approach is also efficient as sensing matrices with fast matrix multiplication algorithms can be used, in particular in the case of Fourier measurements. Secondly, these results are confirmed by a numerical analysis of the phase transition of the \(l_1\)-minimization problem. Finally, we show that the spread spectrum technique remains effective in an analog setting with chirp modulation for application to realistic Fourier imaging. We illustrate these findings in the context of radio interferometry and magnetic resonance imaging. This work has been presented at a conference [93] and accepted for publication in a journal [105]

**Dictionary learning**: An important practical problem in sparse modeling is to choose the adequate dictionary to model a class of signals or images of interest. While diverse heuristic techniques have been proposed in the literature to learn a dictionary from a collection of training samples, there are little existing results which provide an adequate mathematical understanding of the behaviour of these techniques and their ability to recover an ideal dictionary from which the training samples may have been generated.

In 2008, we initiated a pioneering work on this topic, concentrating in particular on the fundamental theoretical question of the identifiability of the learned dictionary. Within the framework of the Ph.D. of Karin Schnass, we developed an analytic approach which was published at the conference ISCCSP 2008 [13] and allowed us to describe "geometric" conditions which guarantee that a (non overcomplete) dictionary is "locally identifiable" by \(l_1\) minimization.

In a second step, we focused on estimating the number of sparse training samples which is typically sufficient to guarantee the identifiability (by \(l_1\) minimization), and obtained the following result, which is somewhat surprising considering that previous studies seemed to require a combinatorial number of training samples to guarantee the identifiability: the local identifiability condition is typically satisfied as soon as the number of training samples is roughly proportional to the ambient signal dimension. The outline of the second result was published in conferences [12], [25]. These results have been published in the journal paper [15].

This year we have worked on extending the results to noisy training samples with outliers. A journal paper is in preparation, and the results will be presented at a workshop at NIPS 2011.

**Analysis Operator Learning for Overcomplete Cosparse Representations**: Besides standard dictionary learning, we also considered learning in the context of the cosparse model. We consider the problem of learning a low-dimensional signal model from a collection of training samples. The mainstream approach would be to learn an overcomplete dictionary to provide good approximations of the training samples using sparse synthesis coefficients. This famous sparse model has a less well known counterpart, in analysis form, called the cosparse analysis model. In this new model, signals are characterized by their parsimony in a transformed domain using an overcomplete analysis operator. We proposed two approaches to learn an analysis operator from a training corpus, both published in the conference EUSIPCO 2011 [79], [67].

The first one uses a constrained optimization program based on \(L_1\) optimization. We derive a practical learning algorithm, based on projected subgradients, and demonstrate its ability to robustly recover a ground truth analysis operator, provided the training set is of sufficient size. A local optimality condition is derived,
providing preliminary theoretical support for the well-posedness of the learning problem under appropriate conditions. Extensions to deal with noisy training samples are currently investigated, and a journal paper is in preparation.

In the second approach, analysis "atoms" are learned sequentially by identifying directions that are orthogonal to a subset of the training data. We demonstrate the effectiveness of the algorithm in three experiments, treating synthetic data and real images, showing a successful and meaningful recovery of the analysis operator.

Connections between sparse approximation and Bayesian estimation: Penalized least squares regression is often used for signal denoising and inverse problems, and is commonly interpreted in a Bayesian framework as a Maximum A Posteriori (MAP) estimator, the penalty function being the negative logarithm of the prior. For example, the widely used quadratic program (with an $\ell^1$ penalty) associated to the LASSO / Basis Pursuit Denoising is very often considered as MAP estimation under a Laplacian prior in the context of additive white Gaussian noise (AWGN) reduction.

A first result, which has been published in IEEE Transactions on Signal Processing [35], highlights the fact that, while this is one possible Bayesian interpretation, there can be other equally acceptable Bayesian interpretations. Therefore, solving a penalized least squares regression problem with penalty $\phi(x)$ need not be interpreted as assuming a prior $C \cdot \exp(-\phi(x))$ and using the MAP estimator. In particular, we showed that for any prior $P_X$, the minimum mean square error (MMSE) estimator is the solution of a penalized least square problem with some penalty $\phi(x)$, which can be interpreted as the MAP estimator with the prior $C \cdot \exp(-\phi(x))$. Vice-versa, for certain penalties $\phi(x)$, the solution of the penalized least squares problem is indeed the MMSE estimator, with a certain prior $P_X$. In general $dP_X(x) \neq C \cdot \exp(-\phi(x))dx$.

A second result, obtained in collaboration with Prof. Mike Davies and Prof. Volkan Cevher (a paper is under revision) characterizes the "compressibility" of various probability distributions with applications to underdetermined linear regression (ULR) problems and sparse modeling. We identified simple characteristics of probability distributions whose independent and identically distributed (iid) realizations are (resp. are not) compressible, i.e., that can be approximated as sparse. We prove that many priors which MAP Bayesian interpretation is sparsity inducing (such as the Laplacian distribution or Generalized Gaussian distributions with exponent $p=1$), are in a way inconsistent and do not generate compressible realizations. To show this, we identify non-trivial undersampling regions in ULR settings where the simple least squares solution outperform oracle sparse estimation in data error with high probability when the data is generated from a sparsity inducing prior, such as the Laplacian distribution.

6.2.3. Wavelets on graphs

Participant: Rémi Gribonval.

Main collaboration: Pierre Vandergheynst, David Hammond (EPFL)

Within the framework of the SMALL project 7.2.1, we investigated the possibility of developing sparse representations of functions defined on graphs, by defining an extension to the traditional wavelet transform which is valid for data defined on a graph.

There are many problems where data is collected through a graph structure: scattered or non-uniform sampling, sensor networks, data on sampled manifolds or even social networks or databases. Motivated by the wealth of new potential applications of sparse representations to these problems, the partners set out a program to generalize wavelets on graphs. More precisely, we have introduced a new notion of wavelet transform for data defined on the vertices of an undirected graph. Our construction uses the spectral theory of the graph laplacian as a generalization of the classical Fourier transform. The basic ingredient of wavelets, multi-resolution, is defined in the spectral domain via operator-valued functions that can be naturally dilated. These in turn define wavelets by acting on impulses localized at any vertex. We have analyzed the localization of these wavelets in the vertex domain and showed that our multi-resolution produces functions that are indeed concentrated at will around a specified vertex. Our theory allowed us to construct an equivalent of the continuous wavelet transform but also discrete wavelet frames.
Computing the spectral decomposition can however be numerically expensive for large graphs. We have shown that, by approximating the spectrum of the wavelet generating operator with polynomial expansions, applying the forward wavelet transform and its transpose can be approximated through iterated applications of the graph Laplacian. Since in many cases the graph Laplacian is sparse, this results in a very fast algorithm. Our implementation also uses recurrence relations for computing polynomial expansions, which results in even faster algorithms. Finally, we have proved how numerical errors are precisely controlled by the properties of the desired spectral graph wavelets. Our algorithms have been implemented in a Matlab toolbox that has been released in parallel to the main theoretical article [16]. We also plan to include this toolbox in the SMALL project numerical platform.

We now foresee many applications. On one hand we will use non-local graph wavelets constructed from the set of patches in an image (or even an audio signal) to perform de-noising or in general restoration. An interesting aspect in this case, would be to understand how wavelets estimated from corrupted signals deviate from clean wavelets. In a totally different direction, we will also explore the applications of spectral graph wavelets constructed from brain connectivity graphs obtained from whole brain tractography. Our preliminary results show that graph wavelets yield a representation that is very well adapted to how the information flows in the brain along neuronal structures.

6.2.4. Algorithmic breakthrough in sparse approximation: LoCOMP

Participants: Rémi Gribonval, Frédéric Bimbot, Ronan Le Boulch.

Main collaborations: Pierre Vandergheynst (EPFL), Boris Mailhé (former team member, now with Queen Mary University, London)

Our team had already made a substantial breakthrough in 2005 when first releasing the Matching Pursuit ToolKit (MPTK, see Section 5.3) which allowed for the first time the application of the Matching Pursuit algorithm to large scale data such as hours of CD-quality audio signals. In 2008, we designed a variant of Matching Pursuit called LoCOMP (ubiquitously for LOw Complexity Orthogonal Matching Pursuit or Local Orthogonal Matching Pursuit) specifically designed for shift-invariant dictionaries. LoCOMP has been shown to achieve an approximation quality very close to that of a full Orthonormal Matching Pursuit while retaining a much lower computational complexity of the order of that of Matching Pursuit. The complexity reduction is substantial, from one day of computation to 15 minutes for a typical audio signal [20], [19]. The main effort this year has been to integrate this algorithm into MPTK to ensure its dissemination and exploitation, and a journal paper has been published [22].

6.3. Emerging activities on compressive sensing and inverse problems

6.3.1. Nearfield acoustic holography (ECHANGE ANR project)

Participants: Rémi Gribonval, Nancy Bertin.

Main collaborations: Albert Cohen (Laboratoire Jacques-Louis Lions, Université Paris 6), Laurent Daudet, Gilles Chardon, François Ollivier, Antoine Peillot (Institut Jean Le Rond d’Alembert, Université Paris 6)

Compressed sensing is a rapidly emerging field which proposes a new approach to sample data far below the Nyquist rate when the sampled data admits a sparse approximation in some appropriate dictionary. The approach is supported by many theoretical results on the identification of sparse representations in overcomplete dictionaries, but many challenges remain open to determine its range of effective applicability. METISS has chosen to focus more specifically on the exploration of Compressed Sensing of Acoustic Wavefields, and we have set up the ANR collaborative project ECHANGE (ECHantillonnage Acoustique Nouvelle Génération) which began in January 2009. Rémi Gribonval is the coordinator of the project.

In 2010, the activity on ECHANGE has concentrated on Nearfield acoustic holography (NAH), a technique aiming at reconstructing the operational deflection shapes of a vibrating structure, from the near sound field it generates. In this application scenario, the objective is either to improve the quality of the reconstruction (for a given number of sensors), or reduce the number of sensors, or both, by exploiting a sparsity hypothesis which helps regularizing the inverse problem involved.
Contributions of the team in this task spans: notations and model definitions, experimental setting design and implementation, choice of an adapted dictionary in which the sparsity hypothesis holds, improved acquisition strategies through pseudo-random sensor arrays and/or spatial multiplexing of the inputs, experimental study of robustness issues, and theoretical study of potential success guarantees based on the restricted isometry property (which revealed being not verified in our case, despite improved experimental performance).

A paper about robustness issues and spatial multiplexing (an alternative to building antennas with random sensor position) was published in GRETSI [88]. A journal paper is under revision.

6.3.2. Sparse reconstruction for underwater acoustics (ECHANGE ANR project)

Participants: Rémi Gribonval, Valentin Emiya, Nikos Stefanakis, Nancy Bertin.

Main collaborations: Jacques Marchal, Pierre Cervenka (UPMC Univ Paris 06)

Underwater acoustic imaging is traditionally performed with beamforming: beams are formed at emission to insonify limited angular regions; beams are (synthetically) formed at reception to form the image. We proposed to exploit a natural sparsity prior to perform 3D underwater imaging using a newly built flexible-configuration sonar device. The computational challenges raised by the high-dimensionality of the problem were highlighted, and we described a strategy to overcome them. As a proof of concept, the proposed approach was used on real data acquired with the new sonar to obtain an image of an underwater target. We discussed the merits of the obtained image in comparison with standard beamforming, as well as the main challenges lying ahead, and the bottlenecks that will need to be solved before sparse methods can be fully exploited in the context of underwater compressed 3D sonar imaging. This work has been submitted to ICASSP 2012 and a journal paper is in preparation.

6.3.3. Audio inpainting (SMALL FET-Open project)

Participants: Rémi Gribonval, Valentin Emiya.

Main collaborations: Amir Adler, Michael Elad (Computer Science Department, The Technion, Israel); Maria G. Jafari, Mark D. Plumbley (Centre for Digital Music, Department of Electronic Engineering, Queen Mary University of London, U.K.).

Inpainting is a particular kind of inverse problems that has been extensively addressed in the recent years in the field of image processing. It consists in reconstructing a set of missing pixels in an image based on the observation of the remaining pixels. Sparse representations have proved to be particularly appropriate to address this problem. However, inpainting audio data has never been defined as such so far.

METISS has initiated a series of works about audio inpainting, from its definition to methods to address it. This research has begun in the framework of the EU Framework 7 FET-Open project FP7-ICT-225913-SMALL (Sparse Models, Algorithms and Learning for Large-Scale data) which began in January 2009. Rémi Gribonval is the coordinator of the project. The research on audio inpainting has been conducted by Valentin Emiya in 2010 and 2011.

The contributions consist of:

- defining audio inpainting as a general scheme where missing audio data must be estimated: it covers a number of existing audio processing tasks that have been addressed separately so far – click removal, declipping, packet loss concealment, unmasking in time-frequency;
- proposing algorithms based on sparse representations for audio inpainting (based on Matching Pursuit and on \(l_1\) minimization);
- addressing the case of audio declipping (i.e. desaturation): thanks to the flexibility of our inpainting algorithms, they can be constrained so as to include the structure of signals due to clipping in the objective to optimize. The resulting performance are significantly improved. This work has been reported in [47] and it will appear as a journal paper [96].
Current and future works deal with developing advanced sparse decomposition for audio inpainting, including several forms of structured sparsity (e.g., temporal and multichannel joint-sparsity) and several applicative scenarios (declipping, time-frequency inpainting).

6.4. Music Content Processing and Music Information Retrieval

6.4.1. Acoustic music modeling

Participants: Nancy Bertin, Emmanuel Vincent.

Main collaborations: R. Badeau (Télécom ParisTech), J. Wu (University of Tokyo)

Music involves several levels of information, from the acoustic signal up to cognitive quantities such as composer style or key, through mid-level quantities such as a musical score or a sequence of chords. The dependencies between mid-level and lower- or higher-level information can be represented through acoustic models and language models, respectively.

Our acoustic models are based on nonnegative matrix factorization (NMF) and variants thereof. NMF models an input short-term magnitude spectrum as a linear combination of magnitude spectra, which are adapted to the input under suitable constraints such as harmonicity and temporal smoothness. While our previous work considered harmonic spectra only, we proposed the use of wideband spectra to represent attack transients and showed that this resulted in improved pitch transcription accuracy [77]. Our past work on the convergence properties of NMF was also disseminated [50].

We used the resulting model parameters to identify the musical instrument associated with each note, by means of a Support Vector Machine (SVM) classifier trained on solo data, and obtained improved instrument classification accuracy compared to state-of-the-art Mel-Frequency Cepstral Coefficient (MFCC) features [42], [78].

6.4.2. Music language modeling

Participants: Frédéric Bimbot, Emmanuel Vincent.

Main collaboration: S.A. Raczynski (University of Tokyo, JP)

We pursued our pioneering work on music language modeling, with a particular focus on the joint modeling of "horizontal" (sequential) and "vertical" (simultaneous) dependencies between notes by log-linear interpolation of the corresponding conditional distributions. We identified the normalization of the resulting distribution as a crucial problem for the performance of the model and proposed an exact solution to this problem.

We also applied the log-linear interpolation paradigm to the joint modeling of melody, key, chords and meter, which evolve according to different timelines. In order to synchronize these feature sequences, we explored the use of beat-long templates consisting of several notes as opposed to short time frames containing a fragment of a single note.

Both of these studies are ongoing.

6.4.3. Music structuring

Participants: Frédéric Bimbot, Gabriel Sargent, Emmanuel Vincent.

External collaboration: Emmanuel Deruty (as an independent consultant)

The structure of a music piece is a concept which is often referred to in various areas of music sciences and technologies, but for which there is no commonly agreed definition. This raises a methodological issue in MIR, when designing and evaluating automatic structure inference algorithms. It also strongly limits the possibility to produce consistent large-scale annotation datasets in a cooperative manner.

We have pursued our investigations on autonomous and comparable blocks, based on principles inspired from structuralism and generativism for producing music structure annotation. This work has allowed consolidating the methodology and producing additional annotations (over 400 pieces) [53].
We have also developed an algorithm aiming at the automatic inference of autonomous and comparable blocks using the timbral and harmonic content of music pieces, in combination with a regularity constraint [72]. Tested within the QUAEOR project and during the MIREX 2011 campaign [94], the algorithm ranked among state-of-the-art methods.

6.5. Source separation

6.5.1. A general framework for audio source separation

Participants: Alexis Benichoux, Frédéric Bimbot, Charles Blandin, Ngoc Duong, Rémi Gribonval, Nobutaka Ito, Alexey Ozerov, Emmanuel Vincent.

Main collaborations: H. Tachibana (University of Tokyo, JP), N. Ono (National Institute of Informatics, JP)

Source separation is the task of retrieving the source signals underlying a multichannel mixture signal. The state-of-the-art approach, which we presented in a survey chapter [95], consists of representing the signals in the time-frequency domain and estimating the source coefficients by sparse decomposition in that basis. This approach relies on spatial cues, which are often not sufficient to discriminate the sources unambiguously. Last year, we proposed a general probabilistic framework for the joint exploitation of spatial and spectral cues [39] that was disseminated in several invited talks [43], [44]. This framework relies in particular on the thesis of Ngoc Duong, which was defended this year [30]. It makes it possible to quickly design a new model adapted to the data at hand and estimate its parameters via the EM algorithm. As such, it is expected to become the basis for a number of works in the field, including our own.

Since the EM algorithm is sensitive to initialization, we devoted a major part of our work to reducing this sensitivity. One approach is to set probabilistic priors over the model parameters, including spatial position priors [56] or temporal continuity priors [55]. A complementary approach is to initialize the parameters in a suitable way using source localization techniques specifically designed for environments involving multiple sources and possibly background noise [33], [54], [83]. In a longer-term perspective, we also investigated the design and exploitation of sparsity priors over time-domain acoustic transfer functions [52], [82].

6.5.2. Exploiting filter sparsity for source localization and/or separation

Participants: Alexis Benichoux, Prasad Sudhakar, Emmanuel Vincent, Rémi Gribonval, Frédéric Bimbot.

Main collaboration: Simon Arberet (EPFL)

Estimating the filters associated to room impulse responses between a source and a microphone is a recurrent problem with applications such as source separation, localization and remixing.

We considered the estimation of multiple room impulse responses from the simultaneous recording of several known sources. Existing techniques were restricted to the case where the number of sources is at most equal to the number of sensors. We relaxed this assumption in the case where the sources are known. To this aim, we proposed statistical models of the filters associated with convex log-likelihoods, and we proposed a convex optimization algorithm to solve the inverse problem with the resulting penalties. We provided a comparison between penalties via a set of experiments which shows that our method allows to speed up the recording process with a controlled quality tradeoff. This work was presented at two conferences [52], [82] and a journal paper including extensive experiments with real data is in preparation.

We also investigated the filter estimation problem in a blind setting, where the source signals are unknown. We proposed an approach for the estimation of sparse filters from a convolutive mixture of sources, exploiting the time-domain sparsity of the mixing filters and the sparsity of the sources in the time-frequency (TF) domain. The proposed approach is based on a wideband formulation of the cross-relation (CR) in the TF domain and on a framework including two steps: (a) a clustering step, to determine the TF points where the CR is valid; (b) a filter estimation step, to recover the set of filters associated with each source. We proposed for the first time a method to blindly perform the clustering step (a) and we showed that the proposed approach based on the wideband CR outperforms the narrowband approach and the GCC-PHAT approach by between 5 dB and 20 dB. This work has been published at ICASSP 2011 [49] and submitted for publication as a journal paper.
On a more theoretical side, we studied the frequency permutation ambiguity traditionnally incurred by blind convolutive source separation methods. We focussed on the filter permutation problem in the absence of scaling, investigating the possible use of the temporal sparsity of the filters as a property enabling permutation correction. The obtained theoretical and experimental results highlight the potential as well as the limits of sparsity as an hypothesis to obtain a well-posed permutation problem. This work has been submitted for publication as a journal paper[99].

6.5.3. Towards real-world separation and remixing applications

Participants: Valentin Emiya, Alexey Ozerov, Laurent Simon, Emmanuel Vincent.

Shoko Araki (NTT Communication Science Laboratories, JP), Cédric Févotte (Télécom ParisTech, FR), Antoine Liutkus (Télécom ParisTech, FR), Volker Hohmann (University of Oldenburg, DE)

Following our founding role in the organization of a regular source separation evaluation campaign (SiSEC), we wrote an invited paper summarizing the outcomes of the three latest campaigns[41]. While some challenges remain, this paper highlighted that progress has been made and that audio source separation is closer than ever to successful industrial applications. This is also exemplified by the i3DMusic project and the contract recently signed with MAIA Studio.

In order to exploit our know-how for these real-world applications, we investigated issues such as how to implement our algorithms in real time and how best to exploit human input or metadata[68],[70]. In addition, while the state-of-the-art quality metrics previously developed by METISS remain widely used in the community, we proposed a new set of perceptually motivated metrics which greatly increase correlation with subjective assessments[34].

6.5.4. Source separation for multisource content indexing

Participants: Kamil Adiloglu, Alexey Ozerov, Emmanuel Vincent.

Main collaborations: J. Barker (University of Sheffield, UK), M. Lagrange (IRCAM, FR)

Another promising real-world application of source separation concerns information retrieval from multi-source data. Source separation may then be used as a pre-processing stage, such that the characteristics of each source can be separately estimated. The main difficulty is not to amplify errors from the source separation stage through subsequent feature extraction and classification stages. To this aim, we proposed a principled Bayesian approach to the estimation of the uncertainty about the separated source signals[45] and propagated this uncertainty to the features. We then exploited it in the training of the classifier itself, thereby greatly increasing classification accuracy[69].

While our work in this direction was initially motivated by music applications (e.g. artist recognition by separating the vocals from the accompaniment), we eventually applied it to noise-robust speech recognition, which is a better defined task[71]. In order to motivate further work by the community, we created a new international evaluation campaign on that topic (CHiME)[86].
MIMETIC Team

6. New Results

6.1. Motion Sensing and analysis

6.1.1. Sensing human activity for detecting falling motions

Participants: Franck Multon [contact], Richard Kulpa, Anthony Sorel, Edouard Auvinet.

Sensing human activity is a very active field of research, with a wide range of applications ranging from entertainment and serious games to personal ambient living assistance. MimeTIC aims at proposing original methods to process raw motion capture data in order to compute relevant information according to the application.

In personal ambient living monitoring, we have collaborated with University of Montreal, Department of Computer Science and Operations Research (DIRO) which main activity is biomedical engineering. A co-supervised student is addressing two complementary problems: detecting people falling in everyday environment and providing easy-to-use clinical gait analysis systems for early detection of potential risks of falling. These two problems have been addressed by reconstructing the visual hull of a subject according to synchronized classical of depth cameras. As visual hull is based on videos it’s subject to occlusions which generally occur in natural environment, such as a room with furniture. We have adapted the classical visual hull algorithm in order to be less sensible to occlusions. We also have proposed an index based on 3D silhouette vertical distribution which enhance this property to tackle occlusion problems [1]. This index is based on a ratio: the volume above a given threshold divided by the total body volume. It has been successfully applied to dozens of falling scenarios involving natural occlusions with furniture. The second problem consists in extracting relevant indexes in gait that could enable clinicians to identify elderly people who have a risk of falling. Classical indexes are based on gait regularity and asymmetry in dual tasks protocols (such as walking while counting downward). 3D silhouettes intrinsically contain all the required information in a unique representation contrary to multi-point motion capture systems. However extracting the relevant information from 3D volumes is complex. We have proposed an original approach based on statistical analysis of the volumes in order to compute indexes for gait asymmetry while simply using 3 depth-cameras (Microsoft Kinects) [1] (see figure 4).

Figure 4. 3D silhouettes reconstructed with three depth-cameras in order to analyze gait asymmetry.
In entertainment and serious games, the problem is different as we need to accurately now the action performed by the user in order to react in a convenient manner. Collaboration with Artefacto Company enabled us to develop such motion recognition methods in serious games scenarios. Given motion capture data provided by an optical motion capture system lead to large state vectors in which the relevant information is hidden. Mixture of Gaussians is generally used as an input of Hidden Markov Models to recognize a motion according to this raw data. To simplify, features are generally introduced in order to capture the relevant geometrical property of the motion with either general information (such as joint angles or Cartesian positions) or application-specific information. The former type of information has the advantage to be generic but leads to recognizers that are very sensitive to style and morphology variations. We have proposed a new generic feature based on morphology-independent representation that enables to tackle this problem (submitted to Eurographics2012). The recognition rate is above 75% for very similar upper-limb motions (see figure 5) while classical methods fail to recognize the same motions (recognition rate below 50%).

![Figure 5. The motion of a user equipped with reflective markers is automatically recognized with a HMM system based on morphology-independent features.](image)

### 6.1.2. The Joyman: a novel immersive locomotion device for virtual environments

**Participant:** Julien Pettré [contact].

We proposed a novel interface called Joyman, designed for immersive locomotion in virtual environments [18]. Whereas many previous interfaces preserve or stimulate the users proprioception, the Joyman aims at preserving equilibrioception in order to improve the feeling of immersion during virtual locomotion tasks. The proposed interface is based on the metaphor of a human-scale joystick. The device has a simple mechanical design that allows a user to indicate his virtual navigation intentions by leaning accordingly. We also propose a control law inspired by the biomechanics of the human locomotion to transform the measured leaning angle into a walking direction and speed - i.e., a virtual velocity vector. We aim at using this interface to enable natural interaction with virtual humans with low-cost devices. The Joyman is patented and was presented at the Emerging Technologies, Siggraph Asia, Hong-Kong [20].

These are joint results with the VR4i team (Anatole Lécuyer and Maud Marchal).

### 6.2. VR and Sports

**Participants:** Richard Kulpa, Benoit Bideau [contact], Brault Sébastien, Burns Anne-Marie.
In the past, we have worked on the interaction between two opponents in handball. We have designed a framework to animate virtual throwers in a reality center and to analyze the gestures of real goalkeepers whose objective was to intercept the corresponding virtual balls. This VR framework was then validated by showing that behaviors in real and virtual environments were similar. These works have been extended by using perception-action coupling and perception-only studies to evaluate the anticipation of opponents. In order to evaluate the importance of perceived parameters, the ball and/or the character animation was successively hidden to determine their importance and the same kind of study was done on the graphical level of details.

These works have been extended to the study of deceptive movements and gaps evaluation in rugby. Combining perceptual analysis based on the use of cutoffs with biomechanical analysis, we have extracted important kinematic information that could explain differences between experts and novices. Indeed, thanks to the cutoffs, it is possible to determine how early each of these two levels of practice can perceive the correct final direction of the opponent. Then this information is correlated to kinematical parameters of this player. Finally, we have embedding these knowledge on the evaluation of novices and experts to create models of rugby defenders. We are currently working on integrating these models in a VR experiment in which the real user is this time the attacker and our model the virtual defender.

Concurrently, studies are experimented to determine if VR can be used for training in sports [9]. The first step was to compare if trainees learned the same way in real situation, facing a video of the lesson or facing a virtual teacher that is animated from the motions of the real course. Based on evaluation of an expert, the results showed that the three groups evolved the same way and reached the same level of practice. The second step is then to have more experts to complete the evaluation but also to combine these results with objective analyses based on kinematics data.

This work is partially funded by the Biofeedback project (DGCIS “Serious Gaming” project) of the M2S laboratory, University Rennes 2. Its goal was to create a training tool that can be used and configured by coaches in order to train athletes to repetitive motions such as katas in karate. The evaluation is made by an automatic module that compares the kinematics of the trainee to a database of expert movements.

6.3. Crowds

Participants: Julien Pettré [contact], Richard Kulpa, Anne-Hélène Olivier, Samuel Lemercier, Yijiang Zhang, Jonathan Perrinet.

6.3.1. Perception of collision in crowds

We designed a level-of-detail (LOD) selection function to determine where collision avoidance constraints in crowd simulation can be relaxed without being perceived by spectators [4]. Collision avoidance is probably the most time-consuming part of crowd simulator. However, when only believable results are required, we argue that visually similar results can be obtained at low computational costs based on macroscopic crowd simulation. Based on a perception study, we determined the conditions for collision to be or not to be detected. We discovered that the camera tilt angle was having a great effect on perception, whereas distance to camera (usually used in previous works) was only the third most important factor to be considered.

6.3.2. Mixed Reality Crowds

In the task of making virtual crowds and real people interact together, we explore a mixed reality solution [22]. The seamless integration of virtual characters into dynamic scenes captured by video is a challenging problem. In order to achieve consistent composite results, both the virtual and real characters must share the same geometrical space and their interactions must adhere to the physical coherence criteria. One essential question is how to detect the motion of real objects - such as real characters moving in the video - and how to steer virtual characters accordingly to avoid unrealistic collisions. We propose an online solution. First, by analysis of the input video, the motion states of the real objects are recovered into a common world 3D coordinate system. Meanwhile, a simplified accuracy measurement is defined to represent the confidence of the motion estimate. Then, under the constraints imposed by the real dynamic objects, the motion of virtual characters are accommodated by a uniform steering model. The final step is to merge virtual objects back the
real video scene by taking into account visibility and occlusion constraints between real foreground objects and virtual ones.

6.3.3. Experiments on crowds

Evaluating crowd simulation models is a difficult task. In the frame of the ANR PEDIGREE project, we put in a lot of effort to perform experiments on groups of walking people in order to dispose of a reference database on groups motion. In order to obtain high-quality data, we measure people locomotion by using optoelectronic motion capture systems. In 2011, we starting obtaining detailed analysis on such motion after large efforts on motion analysis and processing. We had to develop dedicated reconstruction techniques because of the challenging conditions in which we performed our motion capture [ 12 ]. We submitted two papers on following modeling and simulation stages (submitted to Eurographics 2012 and Physical Review E).

6.4. Interactive Virtual Cinematography

Participants: Marc Christie [contact], Christophe Lino.

The domain of Virtual Cinematography explores the operationalization of rules and conventions pertaining to camera placement, light placement and staging in virtual environments. In 2011, two major challenges were tackled (i) the proposition of intelligent interactive assistants integrating users in the process of selecting viewpoints and editing a virtual movie, with the capability of adapting to the user choices, and (ii) the design and implementation of evaluation functions for precisely ranking the quality of viewpoints of a virtual 3D environment.

Our intelligent assistant is designed around (i) an intelligent cinematography engine that can compute, at the request of the filmmaker, a set of suitable camera placements (called suggestions) for starting a shot, representing semantically and cinematically distinct choices for visualizing the current narrative, considering established cinema conventions of continuity and composition along with the filmmaker’s previous selected suggestions, and also his or her manually crafted camera compositions, by a machine learning component that adapts shot editing preferences from user-created camera edits, (ii) a user interface, where the suggestions are presented as small movie frames, arranged in a grid whose rows and columns correspond to visual composition properties of the suggested cameras, and (iii) a motion-tracked camera system that makes the user able of modifying the low-level parameters of the camera in shots in the same way a real operator would. The result of this work [ 16 ] is a novel workflow based on interactive collaboration of human creativity with automated intelligence that enables efficient exploration of a wide range of cinematographic possibilities, and rapid production of computer-generated animated movies. A full prototype has been built and demonstrated at ACM Multimedia conference [ 15 ] as well as ParisFX conference. A patent protecting this technology is currently under evaluation [ 25 ].

The second challenge is related to the design of efficient and precise metrics for measuring the quality of viewpoints. For efficiency, we have proposed parallel GPU-based evaluation techniques for the estimation of multiple viewpoints [ 8 ] coupled within a Genetic Algorithm (Swarm Particle Optimization) to rapidly explore the space of possible viewpoints. For preciseness, we have designed a large range of quality functions relative to screen composition and transition between shots, and employed these functions to either automatically generate movies from actions occuring in the virtual environment [ 13 ] or interactively generating movies by letter the users select best shots and best transitions between shots [ 14 ].

Finally we have been exploring the use of tactile devices to the interactive construction of narratives following Prop’s computational model of stories [ 10 ].

6.5. Biomechanics and Motion Analysis

6.5.1. Balance in highly dynamic situations

Participants: Franck Multon [contact], Ludovic Hoyet.
Balance is a key problem in humans as people stand on two feet which leads to a small base of support area compared to the overall body volume. This unstable state has been widely analyzed in static situation but is still difficult to understand when velocity and acceleration reach ineligible values. We thus have proposed an experimental protocol in order to evaluate if criteria published in the literature for specific motions could be generalized to any dynamic motions (see figure 7). To this end, each studied criterion was tested on various dynamic motions and the number of false falling alarms was reported in each case: the number of frames where the criterion detects loss of balance while the subject is actually balanced. The tested criteria where: the projection of the center of mass on the ground which should remain in the base of support, the Zero Moment Point widely used in robotics, the Zero Rate of Angular Momentum, the Foot Rotation Index and the extrapolated center of mass which was introduced in biomechanics recently. The results demonstrate that none of the criteria succeeded in correctly predicting loss of balance in highly dynamic motions [7]. It thus demonstrate the need to continue some fundamental work on this topic which is a key problem in many applications, including robotics, detection and prevention of falls in the Elderly, understanding performance in sports, improving realism in virtual human simulation...

6.5.2. Interaction strategies between two walkers to avoid collision
Participants: Armel Crétual, Julien Pettré [contact], Anne-Hélène Olivier, Jan Ondrej, Antoine Marin.

In the everyday life situation where two humans walk in the same nearness, each can be considered as a moving obstacle for the other one. They adapt their locomotion with respect to this external disturbance to avoid any collision. Collision avoidance between two humans has been largely neglected in the literature despite it lets us expect specific interactions. The main question we raised was to identify the conditions that induce avoidance manoeuvres in locomotor trajectories: what are the relations between the respective positions and velocities which induce motion adaptations? To answer this question, we proposed an original experiment: thirty participants were asked to walk two-by-two in a motion captured area. We assigned them locomotion tasks in order to provoke varied situations of potential future collisions (see figure 8 a). Following the hypothesis of a reciprocal interaction, we suggested a variable which is common to both of the walkers, the Minimum Predicted Distance (MPD), to predict the actual presence of physical interactions as well as to describe their properties. At each instant t, MPD was computed as the distance the walkers would meet if they did not perform motion adaptation after this instant t. Results showed that walkers adapted their motions only when required, i.e., when initial MPD was too low (<1m). We concluded that human walkers are able to accurately estimate future crossing distance and to mutually adapt it. The evolution of MPD enabled decomposing collision avoidance into 3 successive phases: observation, reaction, and regulation (see figure 8 b). Respectively, these phases corresponded to periods of time when, first, MPD was constant, second, increased to acceptable values by motion adaptation and, third, reached a plateau and even slightly decreased. This final phase demonstrates that collision avoidance is actually performed with anticipation. Future work is needed to inspect individual motion adaptations and to relate them with variation of MPD.

Figure 8. a) Experimental set-up to study collision avoidance between 2 walkers. b) Mean MPD evolution for trials for which initial MPD is below 1m. Interaction follows 3 successive phases: observation, reaction and regulation phases. Normalized time was computed between the time when participants were able to see each other to the time they crossed.

6.5.3. Quantification of pathological gait in adults

Participants: Armel Crétual [contact], Kristell Bervet.

Quantifying gait deviation is still a challenge in adults patient follow-up within a rehab process. This quantification can be done on several levels. Among them, the most useful for practitioners are surely kinematics and muscular activation. On the first one, Gillette Gait Index (GGI) has now become a common tool in rehabilitation centers to assess gait abnormalities. However, one limitation of this index is that it is based on some peak values and is thus sensitive to measurement noise. A new index, the Gait Deviation Index (GDI) which is based on joints angles patterns has been developed by the same team to avoid this problem. Nevertheless, both of them have only been validated in children with cerebral palsy. On the second level, no
satisfying global index has yet been developed. The first part of our study was to validate the GDI in adults. From a database of 74 healthy subjects and 48 patients we did demonstrate that GDI is a relevant index to quantify kinematic gait pattern in adults. Then, we developed a new index called KeR-EGI (for Kerpace-Rennes EMG-based Gait Index) which accounts for the muscular activation patterns of patients. There also, results were conclusive relying on a good correlation between GDI and KeR-EGI. Finally, our recommendation to practitioners was to use both of these two index as they account for complementary aspects of pathology. This allow to better understand if gait disorder is more due to neurological injury or on the opposite to mechanical constraints such as joint stiffness.

6.5.4. Modeling gesture in sports: tennis serve

Participants: Nicolas Bideau [contact], Guillaume Nicolas, Benoit Bideau, Caroline Martin, Richard Kulpa.

Most experimental studies on tennis focus on the segmental coordination in connection with the ball speed, but do not consider the resulting traumatisms. To this end, we currently develop an inverse dynamics modeling approach, based on muskuloskeletal parameters. As a fist step to this work, we calculated the joint constraints on the upper limb in the tennis serve, for professional and regional players. Eleven high level, professional players were compared to seven regional players during this specific motion. Each player was equipped with 42 reflectiver markers and tennis serve was analyzed using an optoelectronical system composed of 12 infrared cameras cadenced at 300Hz. For each player, values of force, power and internal work (in absolute value) were calculated for the three joints (shoulder, elbow, wrist) using a musculoskeletal model. The results showed that professional players produced higher power and internal work for each joint in comparison with the regional players. Results also showed a decrease in the values of internal work from the distal (wrist joint) to the proximal joint (shoulder joint). These results may explain shoulder pain in tennis, which is commonly depicted for high-level players in epidemiological studies. A first perspective to this work is to better take individual parameters (inertia, muscle parameters, pre-constraint, etc.) into account. Another perspective is to test various objective functions in order to predict which parameter is optimized during tennis serve.

6.5.5. Modeling gesture in sports: fin swimming

Participants: Nicolas Bideau [contact], Guillaume Nicolas, Benoit Bideau, Richard Kulpa.

In swimming, experimental approaches are commonly used to analyze performance. However, due to obvious limitations in experimental approaches (impossibility to standardize any situations etc.), it is difficult to characterize surrounding fluid. To overcome this limitation, we currently develop analysis, modeling and simulation of aquatic locomotion, using CFD computer simulation and new methods based on animation of virtual characters.

A first application of this topic enables to evaluate the influence of swim fin flexibility on efficiency during swimming based on a CFD structure interaction model. Finite elements simulations are carried out for various material properties and various prescribed kinematics. Besides the significant effect of flexibility on propulsive forces, the results indicate that the propulsive efficiency is greatly influenced by the stroke frequency and the initial angle of attack. For the selected material properties, the results show that efficiency increases from 3.6 percents to 11.9 percents when the stroke frequency is increased from 0 to 1.7 Hz. Moreover efficiency is clearly increased from 5.0 percents to 24.2 percents when increasing the angle of attack from 0 to 45 degrees. Therefore, an interesting prospect of the present work could be an enhancement of the design of better performing swim fins.

A second application of this topic related to aquatic propulsion deals with a new method to evaluate cross-sectional area based on computer animation of swimming. Indeed, reducing cross sectional area (CSA) during starts and turns is a key part of performance optimisation. Different methods have been used to obtain this parameter without any standard: total human body volume to the power 2/3, wetted area or frontal area based on planimetry technique (PT). These different methods can lead to discrepancies in drag values. Recently, we used two synchronized camcorders to evaluate drag parameters during the different phases of an undulatory stroke cycle.
However, such a technique needs accurate synchronization and calibration of the different camcorders views. The aim of this study is to provide a new method based on animation of virtual characters to obtain instantaneous cross-sectional area in an undulatory stroke cycle. Its main advantage is to obtain cross-sectional area as well as biomechanical analysis with a single camcorder in a sagittal plan and without space calibration. From this, we intend to better understand swimming hydrodynamics and the way CSA influences active drag. More generally, this approach has been designed to provide new practical insights into swimming analysis protocols.

6.6. Path planning and environment analysis

6.6.1. Space-Time planning in dynamic environments

**Participants:** Fabrice Lamarche [contact], Thomas Lopez.

When automatically populating 3D geometric databases with virtual humanoids, modeling the navigation behavior is essential since navigation is used in most exhibited behaviors. In many application fields, the need to manage navigation in dynamic environments arises (virtual worlds taking physics laws into account, numerical plants in which step stools can be moved,...). This study focuses on the following issue: how to manage the navigation of virtual entities in such dynamic environments where topology may change at any time i.e. where unpredictable accessibility changes can arise at runtime. In opposition to current algorithms, movable items are not only considered as obstacles in the environment but can also help virtual entities in their navigation.

The proposed algorithm [17] splits that problem into two complementary processes: a topology tracking algorithm and a path planning algorithm. The aim of the topology tracking algorithm is to continuously detect and update topological relations between moving objects i.e. accessibility or obstruction, while storing temporal information when recurring relations are observed. The path planning algorithm uses this information to plan a path inside the dynamic environment. The coupling of those algorithms endows a virtual character with the ability to immediately use inserted / moved object to reach previously unreachable locations. Moreover, this algorithm is able to find a path through moving platforms to reach a target located on a surface that is never directly accessible.

6.6.2. Automated environment analysis

**Participants:** Fabrice Lamarche [contact], Carl-Johan Jorgensen.

To populate a virtual environment, modeling the navigation behavior is crucial. This behavior relies on the ability of planning a path inside a complex environment, which itself relies on an adequate representation of the environment structure. Most often, virtual environments are represented has 3D databases that are analyzed to produce data structures that are suitable for path planning and navigation. However, without any user intervention, those data structures lack of information about the nature of identified navigable zones that are crucial for navigation credibility.

We proposed an environment analysis algorithm [11] that automatically extracts a meaningful spatial representation of 3D virtual environments, suitable for spatial reasoning. This algorithm automatically differentiates indoor, outdoor and covered parts of the environment. It separates buildings into floors linked by stairs and represent floors as rooms linked by doorsteps. On this basis, we propose a natural hierarchical representation of the environment. This representation is used for spatial reasoning including zone selection and multi-criterion path planning that enhances path credibility.
6. New Results

6.1. Improvement of the force-feedback in a 1-ddl device

Participants: Michel Amberg, Frédéric Giraud, Betty Lemaire-Semail.

Traveling Wave Ultrasonic Motor have many advantages compared to the classical electromagnetic motors: they are lightweight, they don’t need any speed reducer and they make no noise. In a 1-ddl force feedback device, they can help to reduce the bulk size of the mechanism by simplifying the kinematic chain. However, their control has to be very precise because the torque produced is not a straightforward function of the electrical parameters. Previously, we proposed several control algorithms and we obtained good results. But at low speed, problems still remains, like a stick-slip phenomena which makes the motor producing a cogging torque.

To cope with this problem, we first proposed an accurate modeling of the motor and its torque production [24]. We introduced a friction torque $T_f$ which holds the non linearity of the torque production. The evolution of this friction torque has been identified through an experimental study. Then we obtained by inversion a control scheme [13]. The basic idea is to compensate the virtual friction torque. In order to achieve a more accurate control of the torque, we proposed to identify on-line the parameters of the equation of $T_f$.

A one-degree-of-freedom force feedback lever was built to verify the control laws. The experiment involves use of the lever of the digitracker which is free to rotate about the horizontal axis, and is presented in figure 1. In the same figure, we plotted the output torque of the motor, compared to its reference. Both are consistent, showing a good accuracy of the torque controller. To achieve that, the estimator’s parameters are time-varying.

(a) (b)

Figure 1. (a) The 1-ddl haptic device; (b) experimental run of the torque controller with its parameter estimator; in H, results are plotted in the Torque-position plane, reference is in green and measurements is in blue while C, D and G show the estimator’s parameters.

6.2. Haptic Perception of Curvature through active touch

Participants: Michel Amberg, Frédéric Giraud, Betty Lemaire-Semail.
Haptic perception of curvature can be achieved by passive or active finger touch. In this study we proposed a new haptic device that could independently orient, elevate and translate a flat plate. User is free to move his finger on the plate; by controlling plate’s orientation and position accordingly to the position of the finger, we can render a curved shape. The device is composed of two 6-dof haptic devices (Novint Falcon) on which we attached the plate (Figure 2). A force sensor is used to compute the position of the finger on the plate.

Several modelings have been proposed to calculate the orientation and position of the plate. We then measured how accurate simulations of curved shapes are. To achieve that work, we simulated several curved surfaces with different curvature. As the perception performance of curvature is dependent on local surface orientation, the plate was always kept tangent to a virtual shape at the contact point. We then asked people to compare simulated curvature to the real ones. We found that users are able to find the real shape (among five) corresponding to the simulated one [21].

6.3. Tactile input with programmable friction

Participants: Michel Amberg, Géry Casiez, Frédéric Giraud, Betty Lemaire-Semail, Paolo Olivo, Nicolas Roussel.

Our work on programmable friction relies on a particular technology we have been developing for several years. The STIMTAC is a touchpad device that supports friction reduction by means of a squeeze film effect [29]. It uses a controlled vibration at an ultrasonic frequency with a few micrometers amplitude to create an air bearing between a user’s finger and the device’s surface. As the frequency is outside skin mechanoreceptors’ bandwidth, one does not feel this vibration but its effect on tribological contact mechanisms: the touchpad feels more slippery as the amplitude is raised.

We have used this touchpad to create Surfpad, a pointing facilitation technique that operates in the tactile domain. Experiments comparing this technique to semantic pointing [30] and constant control-display gain with and without distractor targets clearly show the limits of traditional target-aware control-display gain
adaptation in the latter case, and the benefits of the tactile approach in both cases [17]. Surfpad leads to a performance improvement close to 9% compared to unassisted pointing at small targets with no distractor. It is also robust to high distractor densities, keeping an average performance improvement of nearly 10% while semantic pointing can degrade up to 100%. Our results also suggest the performance improvement is caused by tactile information feedback rather than mechanical causes, and that the feedback is more effective when friction is increased on targets using a simple step function.

This year’s work on the hardware aspects of the STIMTAC resulted in a compact and quiet prototype powered by the USB cable used for data communication and supporting precise and reliable finger tracking based on multiple force sensors (Figure 3, left). Within the context of the 3DTOUCH project, efforts have also been targeted at the adaptation of the STIMTAC operating principles to off-the-shelf transparent touch sensors. Our latest prototypes demonstrate the compatibility of our approach with resistive (Figure 3, right) and capacitive technologies. In order to facilitate the design and evaluation of novel interaction techniques taking advantage of these prototypes, we have started developing a specific library, tIO, that supports all of them in a unified way.

![Figure 3. Compact opaque STIMTAC and transparent resistive prototype](image)

6.4. Methods and tools to characterize, replicate and compare pointing transfer functions

**Participants:** Géry Casiez, Damien Marchal, Nicolas Roussel.

Transfer functions are the only pointing facilitation technique actually used in modern graphical interfaces involving the indirect control of an on-screen cursor. But despite their general use, very little is known about them. We developed EchoMouse, a device we created to characterize the transfer functions of any system, and libpointing, a toolkit that we developed to replicate and compare the ones used by Windows, OS X and Xorg [16]. We described these functions and reported on an experiment that compared the default one of the three systems. Our results show that these default functions improve performance up to 24% compared to a unitless constant CD gain. We also found significant differences between them, with the one from OS X improving performance for small target widths but reducing its performance up to 9% for larger ones compared to Windows and Xorg. These results notably suggest replacing the constant CD gain function commonly used by HCI researchers by the default function of the considered systems.

6.5. Multimodal pen input for interactive multitouch surfaces

**Participant:** Géry Casiez.
Touch interaction is arguably more immediate and natural in many situations, but fingers are imprecise and difficult to write with. Alternatively, using a pen (or stylus) makes writing more natural and pointing more precise. Luckily, this does not need to be a unilateral choice; pen and touch can be used simultaneously. However, without non-dominant hand coordination or graphical buttons, the pen itself supports few modes. This makes single-handed mobile usage difficult and reduces the number of combined touch and pen modes. When frequently switching between pen-oriented modes, such as drawing, handwriting, gestures, and lasso selection, this can hurt performance. Inferring modes is difficult, and most users prefer explicit control. Schemes for squeezing multiple explicit modes from a pen include adding barrel buttons and classifying pressure, tilt, barrel rotation, or grip. But these can be error-prone and ambiguous. A simple way to add a second mode is by adding an “eraser,” a second contact point. The pencil analogy lends intuition and users have explicit control.

Conté is a small input device inspired by the way artists manipulate a real Conté crayon. By changing which corner, edge, end, or side is contacting the display, the operator can switch interaction modes using a single hand. Conté’s rectangular prism shape enables both precise pen-like input and tangible handle interaction. Conté also has a natural compatibility with multi-touch input: it can be tucked in the palm to interleave same-hand touch input, or used to expand the vocabulary of bimanual touch. Inspired by informal interviews with artists, we catalogue Conté’s characteristics, and use these to outline a design space. We describe a prototype device using common materials and simple electronics. With this device, we demonstrate interaction techniques in a test-bed drawing application.

6.6. Perceived difficulty of pen gestures

Participants: Géry Casiez, Laurent Grisoni.

There are three primary factors which contribute to a successful gesture-based interface: the acquisition technology, the recognizer, and the design of the gesture set. Technologies to acquire gestures, and gesture recognition algorithms, are now quite robust and widely available. However, developing techniques and criteria to help designers create an intuitive and easy-to-perform gesture set remain an active area of research. The challenge is that in order to successfully integrate into an application, a gesture has to satisfy multiple criteria: it must be unambiguously recognized; fit well with its associated function; be easy to learn and recall; and be efficient to perform.
Our empirical results show that users perceive the execution difficulty of single stroke gestures consistently, and execution difficulty is highly correlated with gesture production time. We use these results to design two simple rules for estimating execution difficulty: establishing the relative ranking of difficulty among multiple gestures; and classifying a single gesture into five levels of difficulty. We confirm that the CLC model does not provide an accurate prediction of production time magnitude, and instead show that a reasonably accurate estimate can be calculated using only a few gesture execution samples from a few people. Using this estimated production time, our rules, on average, rank gesture difficulty with 90% accuracy and rate gesture difficulty with 75% accuracy. Designers can use our results to choose application gestures, and researchers can build on our analysis in other gesture domains and for modeling gesture performance [18].
6. New Results

6.1. Mesh repair with topology control

**Participant:** Franck Hétroy.

This work is done in collaboration with Carlos Andújar, Pere Brunet and Álvar Vinacua from Universitat Politecnica de Barcelona, Spain, and has been published in the CAD journal [7]. The purpose is to propose an efficient method to create 2-manifold meshes from real data, obtained as soups of polygons with combinatorial, geometrical and topological noise (see Figure 3). We propose to use a voxel structure called a discrete membrane and morphological operators to compute possible topologies, between which the user chooses.

![Figure 3. Two topologically different 2-manifold mesh repairs, from the same polygon soup.](image)

6.2. Topology computation on simplicial shapes

**Participants:** Dobrina Boltcheva, Franck Hétroy.

This work is a part of the BQR project IDEAL (see Section 8.1.1) which is performed in collaboration with Leila de Floriani from the University of Genova in Italy. The main goal of this project is to study non-manifold geometrical models and to find out features allowing to classify these models and criteria for determining their shape. We are interested in non-manifold models such as idealized industrial CAD models, since they are still ill-understood even if they are frequently used in computer graphics and many engineering applications.

We have developed an efficient method to compute the homology of a large (non-manifold) simplicial complex, from the homologies of its sub-complexes. Computed topological invariants play a crucial role in the field of shape description and analysis. This work has been published in the CAD journal [5] and presented at the SIAM conference on geometric and physical modeling (GD/SPM’11).

6.3. Scale Space Representations on Manifolds

**Participant:** Edmond Boyer.
In collaboration with Radu Horaud and Andrei Zaharescu, we developed a novel approach for the scale-space representations of scalar functions defined over Riemannian manifolds. One of the main interest in such representations stems from the task of 3D modelling where 2D surfaces, endowed with various physical properties, are recovered from images. Multi-scale analysis allows to structure the information with respect to its intrinsic scale, hence enabling a wide range of low-level computations, similar to what is usually used for representing images. In contrast to the Euclidean image domain, where scale spaces can be easily obtained through convolutions with Gaussian kernels, surfaces require a more general approach that must handle non-Euclidean spaces. Such a generalized scale-space framework is the main contribution of this work, which builds on the spectral decomposition available with the heat-diffusion framework to derive a computational approach for representing scalar functions on 2D Riemannian manifolds using an intrinsic scale parameter. In addition, we proposed a feature detector and a region descriptor, based on these representations, extending the widely used DOG detector and HOG descriptor to manifolds. Experiments on real datasets with various physical properties, i.e., scalar functions, demonstrated the validity and the interest of this approach[16].

![Figure 4. Fine-to-coarse representations of scalar functions (color) defined over 2D Riemannian manifolds.](image_url)

**6.4. Topologically-Robust 3D Shape Matching**

*Participant:* Edmond Boyer.

3D Shape matching is an important problem in computer vision. One of the major difficulties in finding dense correspondences between 3D shapes is related to the topological discrepancies that often arise due to complex kinematic motions. In this work done in collaboration with Jan Cech, Radu Horaud and Avinash Sharma a shape matching method is proposed that is robust to such changes in topology. The algorithm starts from a sparse set of seed matches and outputs dense matching. We use a shape descriptor based on properties of the heat-kernel and which provides an intrinsic scale-space representation. This descriptor incorporates (i) heat-flow from already matched points and (ii) self diffusion. At small scales the descriptor behaves locally and hence it is robust to global changes in topology. Therefore, it can be used to build a vertex-to-vertex matching score conditioned by an initial correspondence set. This score is then used to iteratively add new correspondences based on a novel seed-growing method that iteratively propagates the seed correspondences to nearby vertices. The matching is farther densified via an EM-like method that explores the congruency between the two shape embeddings. The method is compared with two recently proposed algorithms and we show that we can deal with substantial topological differences between the two shapes[15].

**6.5. Motion-based segmentation of mesh sequences**

*Participants:* Romain Arcila, Franck Hétroy.
Mesh animations, or sequences of meshes, represent a huge amount of data, especially when acquired from scans or videos. In collaboration with the university of Lyon (LIRIS lab), we address the problem of partitioning these sequences, in order to both recover motion information and be able to compress them. Following last year’s method, we proposed this year a second and third motion-based segmentation algorithm, which clusters mesh vertices into static or rigidly moving components (see Figure 5). These methods are based on spectral clustering of the vertex transformations and are more robust and general than the previous one. This work has been submitted for publication to a journal, and is part of the PhD thesis of Romain Arcila [1].

Figure 5. Temporally varying segmentation of a mesh sequence, into rigidly moving components. Right hand and arm are merged at some time since they start to move accordingly. They are later split, when they start to follow different motions.

6.6. Surface Flow

Participants: Antoine Letouzey, Benjamin Petit, Jean-Sébastien Franco, Edmond Boyer.

Recovering dense motion information is a fundamental intermediate step in the image processing chain upon which higher level applications can be built, such as tracking or segmentation. For that purpose, pixel observations in the image provide useful motion cues through temporal variations of the intensity function. We have studied the estimation of dense, instantaneous 3D motion fields over non-rigidly moving surface observed by multi-camera systems. The motivation arises from multi-camera applications that require motion information for arbitrary subjects, in order to perform tasks such as surface tracking or segmentation. To this aim, we have proposed a novel framework that allows to efficiently compute dense 3D displacement fields using low level visual cues and geometric constraints. The main contribution is a unified framework that combines flow constraints for small displacements with temporal feature constraints for large displacements and fuses them over the surface using local rigidity constraints. The resulting linear optimization problem allows for variational solutions and fast implementations. Experiments conducted on synthetic and real data demonstrated the respective interests of flow and feature constraints as well as their efficiency to provide robust surface motion cues when combined[14], [18].

As an extension of this work, we also studied the situation where a depth camera and one or more color cameras are available, a common situation with recent composite sensors such as the Kinect. In this case, geometric information from depth maps can be combined with intensity variations in color images in order to estimate smooth and dense 3D motion fields. We propose a unified framework for this purpose, that can handle both arbitrary large motions and sub-pixel displacements. The novelty with respect to existing scene flow approaches is that it takes advantage of the geometric information provided by the depth camera to define a surface domain over which photometric constraints can be consistently integrated in 3D. Experiments on real and synthetic data provide both qualitative and quantitative results that demonstrated the interest of the approach[13].
6.7. Learning Temporally Consistent Rigidities

Participants: Jean-Sébastien Franco, Edmond Boyer.

We present a novel probabilistic framework for rigid tracking and segmentation of shapes observed from multiple cameras. Most existing methods have focused on solving each of these problems individually, segmenting the shape assuming surface registration is solved, or conversely performing surface registration assuming shape segmentation or kinematic structure is known. We assume no prior kinematic or registration knowledge except for an over-estimate \( k \) of the number of rigidities in the scene, instead proposing to simultaneously discover, adapt, and track its rigid structure on the fly. We simultaneously segment and infer poses of rigid subcomponents of a single chosen reference mesh acquired in the sequence. We show that this problem can be rigorously cast as a likelihood maximization over rigid component parameters. We solve this problem using an Expectation Maximization algorithm, with latent observation assignments to reference vertices and rigid parts. Our experiments on synthetic and real data show the validity of the method, robustness to noise, and its promising applicability to complex sequences. This work was presented at the CVPR 2011 conference [11].
6.8. Ontology-based mesh segmentation

Participants: Sahar Hassan, Franck Hétroy.

Patient-specific 3D virtual models of anatomical organs are becoming more and more useful in medicine, for instance for diagnosis or follow-up care purposes. These models are usually created from 2D scan or MRI images. However, small or thin geometrical features, such as ligaments, are sometimes not visible on these images. We propose to use an anatomical ontology, called MyCorporisFabrica [http://www.mycorporisfabrica.org/](http://www.mycorporisfabrica.org/), to add missing parts to reconstructed virtual organs. This ontology describes definitions of and relationships between organs: e.g., femur is part of the leg. The first step towards the full achievement of this process is to segment virtual models, often represented by 2D meshes, into meaningful parts. In our case, “meaningful” means “related to the ontology”: each part should refer to an organ defined in the ontology. An algorithm to decompose a given organ into sub-organs according to the ontology has been proposed in the PhD thesis of Sahar Hassan [2]: first, we approximate organ shapes by geometric primitives, then we segment a given organ mesh by optimizing objective functions which are related to these primitives.

6.9. Detection and quantification of brain aneurysms

Participants: Sahar Hassan, Franck Hétroy.

Aneurysms are excrescences on blood vessels. They can break, letting the blood propagate outside the vessel, which often leads to death. In some cases, the blood clots sufficiently fast so that people survive. However, a neurosurgeon or a neuroradiologist should intervene very quickly in order to reparer the vessel before the aneurysm breaks once more.

The purpose of our research is to help neurosurgeons and neuroradiologists to plan surgery, by giving them quantitative information about the size, shape and geometry position of aneurysms. This work was part of the PhD of Sahar Hassan [2], and has presented at the International Conference on Computer Analysis of Images and Patterns (CAIP) [12]. The method we propose first extracts a centered skeleton from the input voxel set of the vascular tree, then detects aneurysms by studying variations of vessel diameters along the skeleton. The name of an aneurysm-carrying vessel is also given thanks to a partial graph matching technique, and accurate measures to decide the treatment are provided.

6.10. Dimensionality reduction for character animation

Participants: Maxime Tournier, Lionel Reveret.

This work investigates and proposes a mathematical framework to perform statistical analysis and dimensionality reduction on rotational trajectories derived from motion capture data. Motion capture data consists in a set of trajectories in the space of 3D rotations (SO(3)) and as such do not present properties of an Euclidian space. Consequently there is no easy to way to apply standard dimensionality reduction techniques on these data. Using the formalism of exponential maps and Principle Geodesics Analysis (PGA), it has been shown that it is possible to rigorously derive a dimensionality reduction analysis on such data. This reduction can be typically applied for compression of motion capture data and probabilistic implementation of the Inverse Kinematics problem. This approach has shown good properties in the context of physically-based animation with a Lagrangian formulation of rigid body dynamics coupled with geometric integrators. These integrators allow a good preservation of momentum using only first order equations, achieving both real-time and high level of realism. These works were developed through the PhD thesis of Maxime Tournier [4]. Early development of PGA on motion capture data had been published at Eurographics in 2009. Its integration into a GPLVM framework has been published this year in the IEEE CG&A journal [9]. Its extension into the context of physically-based animation is currently under preparation for publication.

6.11. Animation of quadrupeds locomotion

Participant: Lionel Reveret.
Following a study on locomotion of quadrupeds by a team in the National Museum of Natural History (MNHN), a new theory on motion planning has been proposed. This theory, the Antero-Posterior Sequences (APS), allows a characterization of the sequence of foot placement for quadrupeds for all regular gaits with very few parameters, as well as transition between gaits, starting from stop to full gallop. In collaboration with the MNHN and the robotics department of the University of Versailles-Saint Quentin en Yvelines (UVSQ), a rigorous software implementation has been specified and developed. This software allows automatically generating foot planning of quadrupeds locomotion according to a desired speed transition. Co-workers for this project were Ludovic Maes and Anick Abourachid at the MNHN and Vincent Hugel at the UVSQ. A patent has been written and finalized for this project.

In parallel, collaboration on physical simulation of quadrupeds locomotion has been carried on with Stelian Coros (previously at University of British Columbia (UBC), now at Disney Research) and Michiel van de Panne (UBC). Automatic video analysis of dog walking, trotting and running has been used to optimize parameters of physical controllers. This work has been published at SIGGRAPH 2011 [6].
6. New Results

6.1. Modeling XML document transformations

Participants: Joachim Niehren, Sophie Tison, Slawek Staworko, Aurélien Lemay, Anne-Cécile Caron, Yves Roos, Shunichi Amano, Camille Vacher, Benoît Groz, Antoine Ndione, Tom Sebastian.

Query answering on XML streams. In [16], Gauwin and Niehren introduce the notion of finite streamability for query languages, and classify fragments of XPath that are finitely streamable or not. They show that if a query language is finitely streamable, then its satisfiability problem can be solved in polynomial time, which in turn is known to fail for mostly all fragments of XPath. They also show that FXP, the fragment of ForwardXPath with child and descendant axis, conjunction, and negation becomes finitely streamable if bounding the number of conjunctions. Since 3 conjunctions are enough in many practical applications, FXP is most relevant in practice. Without any bound, FXP is not finitely streamable, since its satisfiability problem is DEXPTIME hard. The positive result for FXP with a bounded number of conjunctions is obtained by compilation of FXP to deterministic nested word automata. The compiler is in exponential time in the number of conjunctions, and thus polynomial if this parameter is bounded.

Answer enumeration for n-ary queries. Bagan, Filiot, Gauwin, and Niehren investigated answer enumeration algorithms for dialects of XPath with variables. The problem with n-ary queries is that answer sets may grow exponentially in |t|^n, so that algorithms depending polynomially on the size of the answer set might still be unfeasible. In such case, it might still be possible to enumerate elements of answer sets on need. The questions is then whether enumeration can be done efficiently without duplicates and failures, that is with constant delay between subsequent answers and polynomial time preprocessing in the size of the query and the tree. We obtained positive results on answer enumeration with constant delay enumeration for acyclic conjunctive queries over so called X-doublebar structures that we introduce [24]. These subsume tree structures with child, next-sibling and next-sibling* axis, but not the descendant axis. Our result can be lifted to a dialect of ConditionalXPath with variables, that is FO-complete on trees of bounded depth, so that the descendant axis is not needed.

Tree automata global constraints. TAGEds are a new class of tree automata with constraints that currently receive much interest from top conferences on theoretical computer science. During its postdoc in Lille, Vacher improved complexity bounds for some fragments of decidability results [12].

Sequential tree-to-word transducers. Laurence, Lemay, Niehren, Staworko, Tommasi considered deterministic sequential top-down tree-to-word transducers (STWa), that capture the class of deterministic top-down nested-word to word transducers. While reordering and copying are not allowed, STWa are nevertheless very expressive because they allow concatenation of outputs, deletion of inner nodes and they can produce context free languages as output. Their expressiveness is incomparable with DTOPs (plus concatenation, but minus copying). While objecting for learning algorithms, they study normalization of STWa in a first step and then develop unique minimalization algorithms for normalized STWa in a second step in [19]. The idea of normalization is to produce the output in an earliest manner, when reading the input in document order. This works only on binary trees, but can be lifted to unranked trees modulo the binary top-down encoding. The normalization algorithm is by far nontrivial. The natural continuation of this approach will be toward learning algorithms for earliest STWa.

Access control for XML views. The PhD project of Groz, supervised by Staworko and Tison, is centered on access control for XML databases, and in particular on security of user views over XML documents. He obtained results on query rewriting for read-only queries, and translation for update queries. More precisely, given an XML view definition and a user defined query (resp. update program) q, the problem is to find a source query (resp. update program) that is equivalent to q on the view. Caron, Groz, Roos, Staworko and Tison study update programs and views represented by recognizable tree languages in [15], and devise algorithms for update translation in different settings, namely without or with constraints on the authorised source updates.

**Participants:** Jérôme Champavère, Jean Decoster, Jean-Baptiste Faddoul, Antonino Freno, Gemma Garriga, Rémi Gilleron, Mikaela Keller, Grégoire Laurence, Aurélien Lemay, Joachim Niehren, Sławek Staworko, Marc Tommasi, Fabien Torre.

**Induction of tree automata.** Champavère, Gilleron, Lemay and Niehren proposed to use schemas for improving induction algorithms for monadic queries represented by tree automata [26]. The idea is to use pruning strategies to eliminate useless parts of trees when learning from partially annotated trees such that only the structure of relevant fragments is learned. This allows to avoid generalization errors and to learn from fewer annotations. They define schema-guided pruning strategies. They define stable queries w.r.t. a pruning strategy and show that stable queries are learnable.

**Further Results.** In [21], Staworko proposed learning twig and path queries. Prioritized repairing and consistent query answering in relational databases was tackled in [13] and Bounded repairability for regular tree languages in [20]. Torre and Terlutte explored the combination of automata and words balls for sequences classification in [14]. Tommasi participated in the writing of a chapter of a book on conditional Markov fields for information extraction [23].

Garriga and collaborators from the Fraunhofer Institute in Bonn, studied fixed parameter tractable algorithms for the discovery of maximal order preserving submatrices in bioinformatic applications in [17].

We begun a new activity on learning for social network and information network supported by the arrivals of Gemma GARRIGA, Mikaela KELLER and Antonino FRENO. Freno, Garriga and Keller [22] proposed a model for predicting new links in a network which exploit both the current structure of the network and the content of its node.
6. New Results

6.1. The Mining of Complex Data


Formal concept analysis, together with itemset search and association rule extraction, are suitable symbolic methods for KDDK, that may be used for real-sized applications. Global improvements may be carried on the scope of applicability, the ease of use, the efficiency of the methods, and on the ability to fit evolving situations. Accordingly, the team is working on extensions of such symbolic data mining methods to be applied on complex data such as biological or chemical data or textual documents, involving objects with multi-valued attributes (e.g., domains or intervals), n-ary relations, sequences, trees and graphs.

6.1.1. FCA, RCA, and Pattern Structures

Recent advances in data and knowledge engineering have emphasized the need for Formal Concept Analysis (FCA) tools taking into account structured data. There are a few extensions of FCA for handling contexts involving complex data formats, e.g., graphs or relational data. Among them, Relational Concept Analysis (RCA) is a process for analyzing objects described both by binary and relational attributes [116]. The RCA process takes as input a collection of contexts and of inter-context relations, and yields a set of lattices, one per context, whose concepts are linked by relations. RCA has an important role in KDDK, especially in text mining [85], [84].

Another extension of FCA is based on Pattern Structures (PS) [90], which allows to build a concept lattice from complex data, e.g., nominal, numerical, and interval data. In (major [5]), pattern structures are used for building a concept lattice from intervals, in full compliance with FCA, thus benefiting of the efficiency of FCA algorithms. Actually, the notion of similarity between objects is closely related to these extensions of FCA: two objects are similar as soon as they share the same attributes (binary case) or attributes with similar values or the same description (at least in part). Various results were obtained in the study of the relations existing between FCA with an embedded explicit similarity measure and FCA with pattern structures [48]. Moreover, similarity is not a transitive relation and this lead us to the study of tolerance relations. In addition, a new research perspective is aimed at using frequent itemset search methods for mining interval-based data being guided by pattern structures and biclustering as well [50], [49].

Pattern structures in association with a similarity measure were applied in the field of decision support in agronomy. In this domain, a set of agro-ecological indicators is aimed at helping farmers to improve their agricultural practices by estimating the impact of cultivation practices on the “agrosystem”. The modeling and the assessment of environmental risk require a large number of parameters whose measure is imprecise. The propagation of the imprecision and the different types of imprecision have to be taken into account in the computation of the value of indicators for decision support. Actually, based on pattern structures with a associated similarity measure, this problem has been approached as an information fusion problems with substantial results [34], [35].

6.1.2. Miscellaneous in FCA and Pattern Mining

In the field of medicine, an approach based on a combination of FCA with sequential pattern mining was developed to explore patients care trajectories (PCT) [46]. When PCT are modeled as multidimensional and multilevel sequences [108], the results of a frequent sequential itemsets search feed an FCA step in order to compute interests measures such as concept stability. These measures help the experts to find the most interesting sequential patterns.
In the context of environmental sciences, research work is in concern with the mining of complex hydroecological data with concept lattices. FCA was compared and combined with statistical approaches to deal with multi-valued contexts in hydroecology [31], [27], [39]. Regarding the preparation of agronomical data, we have developed an episode-based analysis about the design of information systems (actually, this work was carried out during the ANR-ADD COPT project between 2005 and 2008). We focused on the experience of persons in charge of building observatoires, i.e. information systems, for the monitoring and the management of rural territories [32]. Moreover, Florence Le Ber—as a member of UMR 7517 Lhyges, Strasbourg— is the scientific head of an ANR project named “FRESQUEAU” (2011–2014) dealing with FCA and data mining and hydroecological data (see http://fresqueau.engees.eu/).

For completing the work on itemset search, there is still on-going work on frequent and rare itemset search, for being able to build lattices from very large data and completing the algorithm collection of the Coron platform. This year, results were obtained on the design of an integrated and modular algorithm for searching for closed and generators itemsets, and equivalence classes of itemsets, thus enabling the construction of the associated lattice [56]. This research aspect is also linked to the research carried on within a the PICS CaDoE research project (see Section 8.1.3).

6.1.3. Skylines, sequences and privacy

Pattern discovery is at the core of numerous data mining tasks. Although many methods focus on efficiency in pattern mining, they still suffer from the problem of choosing a threshold that influences the final extraction result. The goal of a study done this current year (2011) is to make the results of pattern mining useful from a user-preference point of view. That is, take into account some domain knowledge to guide the pattern mining process. To this end, we integrate into the pattern discovery process the idea of skyline queries in order to mine skyline patterns in a threshold-free manner. This forms the basis for a novel approach to mining skyline patterns. The efficiency of our approach was illustrated over a use case from chemoinformatics and we showed that small sets of dominant patterns are produced under various measures that are interesting for chemical engineers and researchers [55].

Sequence data is widely used in many applications. Consequently, mining sequential patterns and other types of knowledge from sequence data has become an important data mining task. The main emphasis has been on developing efficient mining algorithms and effective pattern representation.

However, important fundamental problems still remained open: (i) given a sequence database, can we have an upper bound on the number of sequential patterns in the database? (ii) Is the efficiency of the sequence classifier only based on accuracy? (iii) Do the classifiers need the entire set of extracted patterns or a smaller set with the same expressiveness power?

In three different works on sequences, we study the problem of bounding sequential patterns with the combinatorial complexity of sequences and the problem of sequence classifiers with the constraints of optimizing both accuracy and earliness [53], [46].

Orpailleur is one of the few project-teams working on privacy challenges which are becoming a core issue with different scientific problems in computer science. Privacy-preserving data publication has been studied intensely in the past years. In our recent works, we introduce two different data anonymization methodologies based on different usability scenarios [57], [58].

6.1.4. KDDK in Text Mining

Ontologies help software and human agents to communicate by providing shared and common domain knowledge, and by supporting various tasks, e.g. problem-solving and information retrieval. In practice, building an ontology depends on a number of “ontological resources” having different types: thesaurus, dictionaries, texts, databases, and ontologies themselves. We are currently working on the design of a methodology and the implementation of a system for ontology engineering from heterogeneous ontological resources. This methodology is based on both FCA and RCA, and was previously successfully applied in contexts such as astronomy and biology. At present, an engineer is in charge of implementing a robust system
being guided by the previous research results and preparing the way for some new research directions involving trees and graphs.

In another work in text mining [19], we propose a method based on a syntactic parsing for extracting rich semantic relationships between pairs of entities co-occurring in a single sentence. The method was applied in pharmacogenomics (study of the impact of individual genomic variation on drug responses) and we obtained a resource encoded in RDF that summarizes pharmacogenomics relationships mentioned in roughly 17 million Medline abstracts. This resource appears to be of major interest since it is used to guide human curation of biomedical databases, and to derive new knowledge about drug-drug interactions [92].

6.2. KDDK in Life Sciences


One of the major challenges in the post genomic era consists in analyzing terabytes of biological data stored in hundreds of heterogeneous databases (DBs). The extraction of knowledge units from these large volumes of data would give sense to the present data production effort with respect to domains such as disease understanding, drug discovery, and pharmacogenomics or systems biology. Research reported here addresses these important issues and shows the spreading of KDDK over such domains.

6.2.1. Ontology-based Functional Classification of Genes

Functional classification involves grouping genes according to their molecular functions or the biological processes they participate in. This unsupervised classification task is essential for interpreting gene datasets produced by postgenomic experiments. As the functional annotation of genes is mostly based on the Gene Ontology (GO), many similarity measures using the GO have been described, but few of them have been used for clustering [107]. We have evaluated a functional classification of genes using our previously described IntelliGO semantic similarity measure with the help of reference sets [38]. The IntelliGO measure computes semantic similarity between genes for discovering biological functions shared by genes and takes into account domain knowledge represented in Gene Ontology [82]. The reference sets consist of genes taken from human and yeast KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways and Pfam clans. Hierarchical clustering and heatmap visualization were used to illustrate the advantages of IntelliGO over several other measures. Because genes often belong to more than one reference set, the fuzzy C-means clustering algorithm was then applied to the datasets using IntelliGO. The F-score method was used to estimate the quality of clustering and the optimal number of clusters. The results were compared with those obtained from the state of the art DAVID (Database for Annotation Visualization and Integrated Discovery) functional classification method. Overlap analysis allows to study the matching between clusters and reference sets, and leads us to propose a set-difference method for discovering missing information [38]. The IntelliGO similarity measure, the clustering tool and the reference sets used for the evaluation are available at http://plateforme-mbi.loria.fr/intelligo.

6.2.2. Use of Domain Knowledge for Dimension Reduction

Data complexity is a major challenge for knowledge discovery approaches. High dimensionality of datasets can impair the execution of most data mining programs and/or lead to the production of numerous and complex patterns, improper for interpretation by the supervising expert. Thus, an important research orientation is dimension reduction as part of the data preparation step [93]. Domain knowledge is essential for achieving such dataset modification with minimal loss of information. The Life Sciences constitute a suitable domain for testing knowledge-guided approaches for dimension reduction because of the continuous increase in the number of both complex datasets and bio-ontologies. Most of these bio-ontologies are used for annotating biological objects leading to high-dimensional datasets. We propose a new approach for reducing dimensions
in a dataset by exploiting semantic relationships between terms of an ontology structured as a rooted directed acyclic graph [40]. Term clustering is performed thanks to the IntelliGO similarity measure and the term clusters are further used as descriptors for data representation. The technique was applied to a set of drugs associated with their side effects collected from the SIDER database. Terms describing side effects belong to the MedDRA terminology. The hierarchical clustering of about 1,200 MedDRA terms into an optimal collection of 112 term clusters led to a reduced data representation. Two data mining experiments were conducted to illustrate the advantage of using such reduced data representation.

Results obtained in the frame of the collaborative Grand Challenge project (see previous report 2009 and 2010) have been published this year. We have designed the HIV-PDI (Protein-Drug Interactions) resource as a decision making tool to propose alternative antiretroviral drugs (ARVs) for personalized antiretroviral treatment [22]. The HIV-PDI is an integrated database in which sequence mutations of viral proteins can be mapped onto three-dimensional structural interactions between these proteins and ARVs. Thus, critical loss of interactions leading to resistance can be detected and serve as indicators for proposing appropriate ARVs escaping the resistance. As a first step, the HIV-PDI was populated with data relating to HIV protease: clinical information on patients, resistance to ARVs treatments, HIV protease structures and mutations, ARV drugs and their 3D interactions with HIV protease models. Possible queries include protein, drug and treatment conditions, coupled with dedicated tools for visualization/analysis of 3D Protein-Drug interactions. Case-studies demonstrate the capabilities of the HIV-PDI resource for retrieving information associated with patients and for analyzing structural data relating proteins and ligands [23].

6.2.3. Mining Agronomical Data with stochastic models

In the framework of agricultural landscape data mining, we have developed an original approach combining two methods used separately so far: the identification of explicit farmer decision rules through on-farm surveys methods and the identification of landscape stochastic regularities through data-mining of the mosaic of agricultural parcels, following preceding work [96]. This approach was assessed in a study on the Niort plain (West of France) database. In this database, provided by the CEBC (UPR CNRS), the land use occupations of the fields covering a 400 km² area are recorded during 12 years. It results a segmentation of the landscape, based on both its spatial and temporal organization and partly explained by generic farmer decision rules. This consistency between results points out that the two modelling methods interact and may be combined for land-use modelling at landscape scale and for understanding the driving forces of spatial organization. Based on farm surveys, we were able to retrieve and measure changes in land use occupation and link some farmer decision and spatiotemporal regularities that were observed in the landscapes.

6.3. Structural Systems Biology and Docking

Participants: Thomas Bourquard, Marie-Dominique Devignes, Anisah Ghoorah, Bernard Maigret, Lazaros Mavridis, Violeta Pérez-Nueno, Dave Ritchie, Malika Smail-Tabbone, Vishwesh Venkatraman.

Structural systems biology aims to describe and analyze the many components and interactions within living cells in terms of their three-dimensional (3D) molecular structures. Much of our work in this area has been funded by the ANR Chaîres d’Excellence project entitled “High Performance Algorithms for Structural Systems Biology” (HPASSB) which was awarded to Dave Ritchie (January 2009 – September 2011). A related follow-on ANR project entitled “Polynomial Expansions of Protein Structures and Interactions” (PEPSI) has recently started (November 2011). The HPASSB project complements existing competencies in the Orpailleur team represented by Marie-Dominique Devignes (CR CNRS) who is coordinating the MBI project (Modelling Biomolecules and their Interactions, http://bioinfo.loria.fr ), Malika Smail-Tabbone (MCU Nancy University) who is working on data integration and relational data-mining approaches, and Bernard Maigret (DR CNRS) who has an extensive experience of molecular dynamics and virtual screening. We are currently developing advanced computing techniques for molecular shape representation, protein-protein docking, protein-ligand docking, high-throughput virtual drug screening, and knowledge discovery in databases dedicated to protein-protein interactions. The PEPSI project is a collaboration with Sergei Grudinin at INRIA Grenoble (project Nano-D) and Valentin Gordeliy at the Institut de Biologie Structurale in Grenoble. This new project will
Involving developing further the above techniques and using them to help solve the structures of large molecular systems experimentally.

6.3.1. Accelerating protein docking calculations using graphics processors

We have recently adapted the Hex protein docking software to use modern graphics processors (GPUs) to carry out the expensive FFT part of a docking calculation [115]. Compared to using a single conventional central processor (CPU), a high-end GPU gives a speed-up of 45 or more. This software is publicly available at [hex.loria.fr](http://hex.loria.fr). A public GPU-powered server has also been created (hexserver.loria.fr) [99]. These advances have facilitated further work on modeling the assembly of multi-component molecular structures using a particle swarm optimization technique [69].

6.3.2. Eigen-Hex: Modeling protein flexibility during docking

Although the Hex protein docking software can often make reasonably good predictions about how two proteins might fit together, a major limitation of many current algorithms, including Hex, is that they assume that proteins are rigid objects. In fact, proteins can be highly flexible, and the internal conformations of their atoms often change on going from the unbound forms in the free proteins to the bound conformations in the complex. We have developed a novel approach to model such flexibility using a principal component analysis (PCA) technique to identify and predict the main atomic motions during a docking calculation. Our approach gives better results than rigid body docking, although the flexible docking problem is still by no means solved. A journal article describing this work has been submitted.

6.3.3. 3D-Blast: A new approach for protein structure alignment and clustering

We recently developed a new sequence-independent protein structure alignment approach called 3D-Blast [102], which exploits the spherical polar Fourier (SPF) correlation technique used in the Hex protein docking software [114]. This approach recently performed very well in a blind shape comparison experiment organized by Orpailleur as part of Eurographics Workshop on 3D Object Retrieval [103]. The utility of this approach has been demonstrated by clustering subsets of the CATH protein structure classification database [106] for each of the four main CATH fold types, and by searching the entire CATH database of some 12,000 structures using several protein structures as queries. Overall, the automatic SPF clustering approach agrees very well with the expert-curated CATH classification, and ROC-plot analysis of database searches show that the approach has very high precision and recall. We recently proposed that the 3D-Blast approach could ultimately provide a novel way to enumerate and index protein fold space (major [7]).

6.3.4. KBDOCK: Protein docking using Knowledge-Based approaches

Protein docking is the difficult computational task of predicting how a pair of three-dimensional protein structures come together to form a complex. Historically, there has been considerable interest in developing ab initio docking algorithms such as the Hex docking program developed by Dave Ritchie. However, as structural genomics initiatives continue to populate the space of protein 3D structures, and as several on-line databases of protein interactions have recently become available, using structural database systems to perform docking by homology will become an increasingly powerful approach to predicting protein interactions. In order to explore such possibilities, Anisah Ghoorah has recently developed the KBDOCK system as part of her doctoral thesis project. KBDOCK combines residue contact information from the 3DID database [117] with the Pfam protein domain family classification [89] together with coordinate data from the Protein Data Bank [86] in order to describe and analyze all known protein-protein interactions for which the 3D structures are available. In a recent publication [24] we demonstrated the utility of this approach for template-based docking using 73 complexes from the Protein Docking Benchmark [94]. KBDOCK is available at [kbdock.loria.fr](http://kbdock.loria.fr).

6.3.5. V-Dock: scoring protein-protein interactions using Voronoi fingerprints
There is growing interest in using machine learning techniques to analyze and populate protein-protein interaction (PPI) networks [104]. The aim of this project is to investigate the use of Voronoi fingerprints [16] as a way to distinguish cognate and non-cognate pairs of protein-protein interfaces. In collaboration with colleagues in the INRIA AMIB and INRA Bios teams, we recently applied our Voronoi fingerprint representation (V-Dock) to re-score rigid body docking predictions from Hex [60], and we demonstrated that it could be used to improve the ranking of 7 out of 9 docking targets from the CAPRI protein docking experiment [60]. This approach was also used to predict the stability of engineered protein structures for another recent CAPRI target [21].

6.3.6. DOVSA: Developing new algorithms for virtual screening

In 2010, Violeta Pérez-Nueno joined the Orpailleur team thanks to a Marie Curie Intra-European Fellowship (IEF) award to develop new virtual screening algorithms (DOVSA). The aim of this project is to advance the state of the art in computational virtual drug screening by developing a novel consensus shape clustering approach based on spherical harmonic (SH) shape representations [110]. The main disease target in this project is the acquired immune deficiency syndrome (AIDS), caused by the human immuno-deficiency virus (HIV) [109]. However, the approach will be quite generic and will be broadly applicable to many other diseases. So far, good progress has been made on calculating and clustering spherical harmonic “consensus shapes” which represent rather well the essential features of groups of active molecules [30]. Recent progress on this project has been presented orally at the 5th Journée Nationale de Chémoinformatique in Cabourg, the 9th International Conference on Chemical Structures in Noordwijkerhout, and at 3rd International Conference on Drug Discovery and Therapy in Dubai. A review of the state of the art in drug promiscuity was also recently published [29].

6.4. Around the Kasimir research project

Participants: Nicolas Jay, Jean Lieber, Bart Lamiroy, Amedeo Napoli, Thomas Meilender.

This special research project involves researchers working around the Kasimir project and Bart Lamiroy who was attached to the Orpailleur Team during his “INRIA délégation” (2010-2011) and at the same time was a visiting scientist at Lehigh University, USA. The background of Bart Lamiroy is in document and image analysis. Recently he was interested in investigating the application of KDDK to numerical and structural data including document images. The objective is to extend mining tools towards complex and semi-structured multi-media data on the one hand, and to associate image analysis with KDDK techniques on the other hand.

The main research direction which is followed at the moment is in concern with the Kasimir project. Actually, oncology protocols are mainly documented and represented in diagram formats. The classification and CBR techniques used in the Kasimir project require that the ontologies and decision protocols have to be represented in OWL. Based on previous work, we started modeling the mapping of visual features in diagram charts with semantics of the medical domain ontology. The mapping between the visual ontology and the domain ontology should guide a more complete extraction of the protocols from the diagrams for completing the domain ontology of the Kasimir system.

Moreover, during his stay at Lehigh University, Bart Lamiroy developed a new approach for recovering useful information within image data. By recording a wide range of “provenance information” related to complex image analysis processes, the DAE platform (http://dae.cse.lehigh.edu) provides a large set of metadata that can be used by KDDK methods. For example, this allows the correlation and combination of numerical and symbolic aspects, e.g. relating image aspects and domain symbolic representations (within domain ontologies). This work bridges the gap between formal knowledge representation and signal-based pattern recognition and offers a robust experimental environment for further application of KDDK on image data.

6.5. Around the Taaable research project

The Taaable project (http://taaable.fr) has been originally created as a challenger of the Computer Cooking Contest (CCC, organized during the ICCBR Conference). A candidate to this contest is a system whose goal is to solve cooking problems on the basis of a recipe book (common to all candidates), where each recipe is a shallow XML document with an important plain text part. The size of the recipe book (about 1500 recipes) prevents from a manual indexing of recipes: this indexing is performed using semi-automatic techniques.

After being ranked twice second, in the 2008 and 2009 CCCs organized during the ICCBR conference, Taaable won the first price and the adaptation challenge, in 2010 (note that no contest was organized in 2011). Beyond its participation to the CCCs, the Taaable project aims at federating various research themes: case-based reasoning, information retrieval, knowledge acquisition and extraction, knowledge representation, minimal change theory, ontology engineering, semantic wikis, text-mining, etc.

The most important original features of this version are:

A module for refining the domain ontology for improving the case retrieval. In Taaable, the retrieval of similar cases is based on a query generalization using an ontology of the cooking domain. In order to make the case retrieval more progressive and more precise, a enrichment of the domain ontology, and especially the ingredient hierarchy, has been studied and implemented [42]. The refinement process consists in inserting intermediate classes into the initial hierarchy of the system for better distinguishing classes that were initially not distinguishable. In order to introduce new classes into the initial hierarchy, the initial classes of ingredients have been characterized with additional properties. These additional properties are cooking actions that can be applied to ingredients and that have been extracted from the texts of recipes. FCA has been used on these new properties for restructuring the initial hierarchy.

A module for computing adaptation knowledge. Adaptation knowledge discovery has been performed for better adapting cooking recipes to user constraints. This paper extends the approach proposed in 2009 [80] for extracting this kind of adaptation knowledge. The adaptation knowledge comes from the interpretation of closed itemsets whose items correspond to the ingredients that have to be removed, kept, or added. An original approach focusing on a restrictive binary context building and on a specific ranking based on the form of the closed itemsets has been proposed [47].

Several theoretical studies have been carried out that should be applied to some future versions of Taaable:

- The representation of preparations in temporal representation formalisms [63].
- An algorithm for adapting cases defined in the expressive description logic $\mathcal{ALC}$ [43], [11].
- The study of the relations between adaptation based on belief revision and other approaches to adaptation [61], [11].
- The study of the extension of the domain ontology to make the retrieval step of a case-based reasoning system more accurate [42].
- The study of adaptation knowledge discovery based on variation of ingredients between pairs of recipes [42].
6. New Results

6.1. Speech Analysis and Synthesis

Participants: Anne Bonneau, Vincent Colotte, Dominique Fohr, Yves Laprie, Joseph di Martino, Slim Ouni, Asterios Toutios, Sébastien Demange, Fadoua Bahja, Agnès Piquard-Kipffer, Utpala Musti.

6.1.1. Acoustic-to-articulatory inversion

6.1.1.1. Building new articulatory models

The possibility of generating the same sounds as those uttered by the speaker (or at least vocal tract transfer functions not too far from those observed) via the articulatory model and the acoustic simulation constitutes the underlying hypothesis of an analysis by synthesis method of acoustic-to-articulatory inversion. The articulatory model, and consequently its construction, thus plays a crucial role in inversion. An geometrical adaptation procedure has been developed in order to account for new speakers [28], [29]. It uses two scaling factors, one for the mouth cavity and the second for the pharyngeal cavity. In addition the model can be rotated globally and a second rotation controls the relative position of the pharynx with respect to the mouth cavity. In order to ensure a smooth transition from the mouth cavity to the pharynx cavity the angle of the rotation is a function of the distance with respect to the mouth axis.

The adaptation and model have been tested by using the X-ray data used by Maeda to construct his model. It should be noted that there are very few X-ray data with articulatory contour information available. These data correspond to a female speaker. The RMS reconstruction error reached by the adapted articulatory model is 0.550 mm what is very good for this particular speaker. Other data will be used in the future to validate the model and the adaptation procedure as soon as the contours will be delineated. An anatomical adaption procedure will also be developed in the future.

6.1.1.2. Determination of the vocal tract centerline

The connection of the articulatory model with the acoustic simulation requires the area function to be decomposed into elementary uniform tubes. The decomposition should respect the plane wave propagation. For that purpose the central line of the vocal tract has to be determined. The quality of the centerline strongly influences the closeness between natural and artificial formant frequencies.

We designed two complementary algorithms. The first exploits a dynamic programming approach to select points on interior and exterior walls of the vocal tract which minimize a global criterion combining the length of the centerline and the angle between the normal to the segments linking the points selected on both walls and the centerline [29]. It turned out that this first algorithm provides an insufficient smoothness of the centerline. A second algorithm has been designed by using an active curve which maximizes the smoothness of the centerline and the distance from any point of the centerline with exterior and interior walls. This second algorithm provides very good results.

6.1.1.3. Adaptation of cepstral coefficients for inversion

The inversion of speech requires spectra of natural speech to be compared with spectra synthesized via the articulatory synthesizer. This comparison cannot be carried out directly because the source is not taken into account in the synthetic spectra. Last year we thus investigated an affine adaptation of all the cepstral coefficients. This adaptation brings the spectral peaks of natural and synthetic spectra closer but at the same time tends to flatten the spectra. Moreover, it also appears that adaptation of only the very first cepstral coefficients (the first two except \( C_0 \) which represents energy) were sufficient to capture the spectral tilt. Since it is important to keep clear spectral peaks to explore the articulatory space, we used the bilinear transform in order to bring the two spectra closer [15]. The results are now better and the bilinear transform will be used to recover inverse solutions.
6.1.1.4. Acoustic-to-articulatory inversion using a generative episodic memory

We have developed an episodic based inversion method. Episodic modeling is interesting for two reasons. First, it does not rely on any assumption about the mapping relationship between acoustic and articulatory, but rather it relies on real synchronized acoustic and articulatory data streams. Second, the memory structurally embeds the naturalness of the articulatory dynamics as speech segments (called episodes) instead of single observations as for the codebook based methods. Estimating the unknown articulatory trajectories from a particular acoustic signal, with an episodic memory, consists in finding the sequence of episodes, which acoustically best explains the input acoustic signal. We refer to such a memory as a concatenative memory (C-Mem) as the result is always expressed as a concatenation of episodes. Actually a C-Mem lacks from generalization capabilities as it contains only several examples of a given phoneme and fails to invert an acoustic signal, which is not similar to the ones it contains. However, if we look within each episode we can find local similarities between them. We proposed to take advantage of these local similarities to build a generative episodic memory (G-Mem) by creating inter-episodes transitions. The proposed G-Mem allows switching between episodes during the inversion according to their local similarities. Care is taken when building the G-Mem and specifically when defining the inter-episodes transitions in order to preserve the naturalness of the generated trajectories. Thus, contrary to a C-Mem the G-Mem is able to produce totally unseen trajectories according to the input acoustic signal and thus offers generalization capabilities. The method was implemented and evaluated on the MOCHA corpus, and on a corpus that we recorded using an AG500 articulograph. The results showed the effectiveness of the proposed G-Mem which significantly outperformed standard codebook and C-Mem based approaches. Moreover similar performances to those reported in the literature with recently proposed methods (mainly parametric) were reached. [18]

The paradigm of episodic memories was also used for speech recognition. We do not extend the acoustic feature with any explicit articulatory measurements but instead we used the articulatory-acoustic generative episodic memories (G-mem). The proposed recognizer is made of different memories each specialized for a particular articulator. As all the articulators do not contribute equally to the realization of a particular phoneme, the specialized memories do not perform equally regarding each phoneme. We showed, through phone string recognition experiments that combining the recognition hypotheses resulting from the different articulatory specialized memories leads to significant recognition improvements. [19].

6.1.2. Using Articulography for Speech production

Since we have an articulograph (AG500, Carstens Medizinelektronik) available, we can easily acquire articulatory data required to study speech production. The articulograph is used to record the movement of the tongue (this technique is called electromagnetography - EMA). The AG500 has a very good time resolution (200Hz), which allows capturing all articulatory dynamics. The articulograph was used in a study about inversion (see the previous section) and to investigate pharyngealization.

Pharyngealized phonemes are commonly described as having the same place of articulation (dental) as their non-pharyngealized counterparts, but differ by the presence of a secondary articulation involving mainly the back of the tongue.

To study pharyngealized phonemes in Arabic from an articulatory point of view, our articulograph was used to record the movement of the tongue. Although EMA is not known as an optimal technique to cover the back of the tongue, good placement of the sensors and good interpretation of their positions can help to define pharyngealization relevantly. In fact, it is important to set one sensor as far as possible on the tongue (in our case, at 7cm from the tongue tip).

A corpus of several CVCVCVs was recorded using this articulograph, then phonetically labeled, and analyzed. The main finding of this work is that the coarticulation effect of the pharyngealized phonemes extends the immediate surrounding phonemes to influence the phonemes up to four-phoneme distance from the pharyngealized phoneme. The pharyngealization affects indifferently the previous and the following vowels and consonants.
We also investigated the effect of pharyngealization in Modern Standard Arabic (MSA) and Dialectal Arabic (DA). The acoustic material was more important than EMA. Although, we studied one speaker for EMA, the obtained results are encouraging to record more arabic speakers. [42]

6.1.3. Labial coarticulation

Results show that protrusion is a fragile cue to the rounding feature. Although we observe for each speaker a clear (but not large) separation between vowels /i/ and /y/ produced in isolation, many realizations of /i/ and /y/ come very close together and even overlap in few cases for vowels in contexts. The efficiency of the parameter depends on speakers and contexts. The distance between the corners is probably the most fragile cue to vowel roundedness. Many overlapping areas are observed for vowels in context. This is not good news for speech specialists since this parameter is easy to measure (with cameras and markers painted on the speaker’s face) and its evaluation can be fully automatic. Each of the three lip opening parameters constitutes a very efficient cue to the rounding feature. For vertical opening, the opposition between /i/ and /y/ in initial position appears to be endangered in bilabial context, due to the anticipation of lip closing during /i/. Nevertheless, the temporal variations of lip opening during the initial /i/ are very important, and more analyses, taking into account these variations, will be necessary to analyse /i/ vs. /y/ phonetic distinction more thoroughly.

6.1.4. Speech synthesis

Visual data acquisition was performed simultaneously with acoustic data recording, using an improved version of a low-cost 3D facial data acquisition infrastructure. The system uses two fast monochrome cameras, a PC, and painted markers, and provides a sufficiently fast acquisition rate to enable an efficient temporal tracking of 3D points. The recorded corpus consisted of the 3D positions of 252 markers covering the whole face. The lower part of the face was covered by 70% of all the markers (178 markers), where 52 markers were covering only the lips so as to enable a fine lip modeling. The corpus was made of 319 medium-sized French sentences uttered by a native male speaker and corresponding to about 25 minutes of speech.

We designed a first version of the text to acoustic-visual speech synthesis based on this corpus. The system uses bimodal diphones (an acoustic component and a visual one) and unit selection techniques (see 3.2.4). We have introduced visual features in the selection step of the TTS process. The result of the selection is the path in the lattice of candidates found in the Viterbi algorithm, which minimizes a weighted linear combination of three costs: the target cost, the acoustic joined cost, and the visual joined cost.

Finding the best set of weights is a difficult problem by itself mainly because of their highly different nature (linguistic, acoustic, and visual considerations). This year, we added the first derivative of the visual trajectories in the visual joint cost and we developed a method to determine automatically the weights applied to each cost, using a series of metrics that assess quantitatively the performance of synthesis [37].

This year, more progress have been made regarding the definition of the target cost. Now, The target cost includes both acoustic target cost and visual target cost.

The visual target cost includes visual and articulatory information. We implemented and evaluated two techniques [32]: (1) Phonetic category modification, where the purpose was to change the current characteristics of some phonemes which were based on phonetic knowledge. The changes modified the target and candidate description for the target cost to better take into account their main characteristics as observed in the audio-visual corpus. The expectation was that their synthesized visual speech component would be more similar to the real visual speech after the changes. (2) Continuous visual target cost, where the visual target cost component is now considered as real value, and thus continuous, based on the articulatory feature statistics.

6.1.5. Phonemic discrimination evaluation in language acquisition and in dyslexia and dysphasia

6.1.5.1. Phonemic segmentation in reading and reading-related skills acquisition in dyslexic children and adolescents

Our computerized tool EVALEC was published [67] after the study of reading level and reading related skills of 400 hundred children from grade 1 to grade 4 (from age 6 to age 10) [69]. This research was supported by a grant from the French Ministry of Health (Contrat 17-02-001, 2002-2005). This first computerized battery
of tests in French language assessing reading and related skills (phonemic segmentation, phonological short term memory) comparing results both to chronological age controls and reading level age control in order to diagnostic Dyslexia. Both processing speed and accuracy scores are taken into account. This battery of tests is used by speech and language therapists. We keep on examining the reliability (group study) and the prevalence (multiple case study) of 15 dyslexics’ phonological deficits in reading and reading related skills in comparison with a hundred reading level children [68], and by the mean of longitudinal studies of children from age 5 to age 17 [66]. This year, we started the development of a project which examined multimodal speech both with SLI, dyslexics and control children (30 children). Our goal is to examine visual contribution to speech perception across different experiments with a natural face (syllables with several conditions). Our goal is to search what can improve intelligibility in children who have severe language acquisition difficulties.

6.1.5.2. Language acquisition and language disabilities (deaf children, dysphasic children)

Providing help for improving French language acquisition for hard of hearing (HOH) children or for children with language disabilities was one of our goal: ADT (Action of Technological Developpement) Handicom [piquardkipffer:2010:inria-00545856:2]. The originality of this project was to combine psycholinguistical and speech analyses researches. New ways to learn to speak/read were developed. A collection of three digital books has been written by Agnès Piquard-Kipffer for both 2-6, 5-9, 8-12 year old children (kindergarten, 1-4th grade) to train speaking and reading acquisition regarding their relationship with speech perception and audio-visual speech perception. A web interface has been created (using Symfony and AJAX technologies) in order to create others books for language impaired children. A workflow which transforms a text and an audio source in a video of digital head has been developed. This workflow includes an automatic speech alignment, a phonetic transcription, a speech synthesizer, a French cued speech coding and speaking digital head. A series of studies (simple cases studies, 5 deaf children and 5 SLI children and group studies with 2 kindergarten classes) were proposed to investigate the linguistical, audio-visual processing. . . . presumed to contribute to language acquisition in deaf children. Publication are submitted.

6.1.6. Enhancement of esophageal voice

6.1.6.1. Detection of F0 in real-time for audio: application to pathological voices

The work first rested on the CATE algorithm developed by Joseph Di Martino and Yves Laprie, in Nancy, 1999. The CATE (Circular Autocorrelation of the Temporal Excitation) algorithm is based on the computation of the autocorrelation of the temporal excitation signal which is extracted from the speech log-spectrum. We tested the performance of the parameters using the Bagshaw database, which is constituted of fifty sentences, pronounced by a male and a female speaker. The reference signal is recorded simultaneously with a microphone and a laryngograph in an acoustically isolated room. These data are used for the calculation of the contour of the pitch reference. When the new optimal parameters from the CATE algorithm were calculated, we carried out statistical tests with the C functions provided by Paul BAGSHAW. The results obtained were very satisfactory and a first publication relative to this work was accepted and presented at the ISIVC 2010 conference [46]. At the same time, we improved the voiced / unvoiced decision by using a clever majority vote algorithm electing the actual F0 index candidate. A second publication describing this new result was published at the ISIC 2010 conference [45].

6.1.6.2. Voice conversion techniques applied to pathological voice repair

Voice conversion is a technique that modifies a source speaker’s speech to be perceived as if a target speaker had spoken it. One of the most commonly used techniques is the conversion by GMM (Gaussian Mixture Model). This model, proposed by Stylianou, allows for efficient statistical modeling of the acoustic space of a speaker. Let “x” be a sequence of vectors characterizing a spectral sentence pronounced by the source speaker and “y” be a sequence of vectors describing the same sentence pronounced by the target speaker. The goal is to estimate a function F that can transform each source vector as nearest as possible of the corresponding target vector. In the literature, two methods using GMM models have been developed: In the first method (Stylianou), the GMM parameters are determined by minimizing a mean squared distance between the transformed vectors and target vectors. In the second method (Kain), source and target vectors are combined in a single vector “z”. Then, the joint distribution parameters of source and target speakers is estimated using
the EM optimization technique. Contrary to these two well known techniques, the transform function F, in our laboratory, is statistically computed directly from the data: no needs of EM or LSM techniques are necessary. On the other hand, F is refined by an iterative process. The consequence of this strategy is that the estimation of F is robust and is obtained in a reasonable lapse of time. This interesting result was published and presented at the ISIVC 2010 conference [70].

6.1.7. Perception and production of prosodic contours in L1 and L2

6.1.7.1. Language learning (feedback on prosody)

Feedback on L2 prosody based upon visual displays, speech modifications and automatic diagnosis has been elaborated and a pilot experiment undertaken to test its immediate impact on listeners. Results show that the various kinds of feedback provided by the system enable French learners with a low production level to improve their realisations of English lexical accents more than (simple) auditory feedback. These results should be confirmed with a large number of speakers but based upon the important differences between results obtained for speakers in test and control conditions, we are confident in the interest of the system presented here [41]. In particular, the system analyses learners’ realisations and provide indications on what they should correct, a guidance which is considered as necessary by specialists in the oral aspects of language learning.

6.1.7.2. Production of prosody contour

We report here relevant observations for the study continuation in French. These observations were obtained in an ongoing project about non-conclusive prosodic patterns in French and English (“Intonale” project 7.2.2). We specifically discuss slope variations, estimated in semitones, concerning two kinds of non-conclusive configurations, which are inside a clause, or at the end of a clause, respectively: (i) the final segment of a subject NP in an assertive sentence, followed or not by another syntagm ended by a continuation contour (ii) the final segment of a A clause, in a two clause utterance AB, where A and B are assertive clauses connected by an discourse relation, marked or not with a conjunction.

Intonation slopes are computed as regression slopes using F0 values in semitones estimated every 10 ms. Slopes are calculated on the two last syllables of the target segments of every sentence. Results show that slopes for segments which are not at the end of a clause, and segments at the end of a clause followed by a conjunction are typically rising, and not significantly different the ones from the others. On the contrary, slopes for ends of clauses not followed by a conjunction are significantly different from the previous ones. More than 50

We are presently studying English sentences, in particular continuation contours, produced by French speakers, in order to determine the impact of their native language (French) on their English pronunciations.

6.1.8. Pitch detection

Over the last two years, we have proposed two new real time pitch detection algorithms (PDAs) based on the circular autocorrelation of the glottal excitation, weighted by temporal functions, derived from the CATE [64] original algorithm (Circular Autocorrelation of the Temporal Excitation), proposed initially by J. Di Martino and Y. Laprie. In fact, this latter algorithm is not constructively real time because it uses a post-processing technique for the Voiced/Unvoiced (V/UV) decision. The first algorithm we developed is the eCATE algorithm (enhanced CATE) that uses a simple V/UV decision less robust than the one proposed later in the eCATE+ algorithm.

We propose a recent modified version called the eCATE++ algorithm which focuses especially on the detection of the F0, the tracking of the pitch and the voicing decision in real time. The objective of the eCATE++ algorithm consists in providing low classification errors in order to obtain a perfect alignment with the pitch contours extracted from the Bagshaw database by using robust voicing decision methods. The main improvement obtained in this study concerns the voicing decision, and we show that we reach good results for the two corpora of the Bagshaw database.

6.2. Automatic Speech Recognition

Participants: Christophe Cerisara, Sébastien Demange, Dominique Fohr, Christian Gillot, Jean-Paul Haton, Irina Illina, Denis Jouvet, Odile Mella, Luiza Orosanu, Othman Lachhab, Larbi Mesbah.
6.2.1. Core recognition

6.2.1.1. Broadcast News Transcription

In the framework of the Technolangue project ESTER, we have developed a complete system, named ANTS, for French broadcast news transcription (see section 5.4).

Extensions of the ANTS system have been studied, including the possibility to use the Sphinx recognizers. Training scripts for building acoustic models for the Sphinx recognizers are now available and take benefit of the computer cluster for a rapid optimization of the model parameters. The Sphinx models are also used for speech/text alignment on both French and English speech data. A new speech decoding program has been developed for efficient decoding on the computer cluster, and easy modification of the decoding steps (speaker segmentation and clustering, data classification, speech decoding in one or several passes, ...). It handles both the Julius and Sphinx (versions 3 and 4) decoders.

This year, we have proposed an approach to grapheme-to-phoneme conversion based on a probabilistic method: Conditional Random Fields (CRF). CRF gives a long term prediction, and assume a relaxed state independence condition. Moreover, we proposed an algorithm to the one-to-one letter to phoneme alignment needed for CRF training. This alignment is based on discrete HMMs. The proposed system was validated on two pronunciation dictionaries. Different set of input features were studied: POS-tag, context size, unigram versus bigram. Our approach compared favorably with the performance of the state-of-the-art Joint-Multigram Models (JMM) for the quality of the pronunciations, but provided better recall and precision measures for multiple pronunciation variants generation [22][21].

As the pronunciation lexicon is one the key-points of a speech recognition system, we have investigated to which extent wiktionary data can be used to build such a lexicon. Collecting the pronunciations available for many entries of the wiktionary make possible the creation of an initial pronunciation lexicon. Such initial lexicon is then used for training grapheme-to-phoneme conversion systems (either CRF-based of JMM-based), in order to obtain pronunciation variants for words that are not in the initial pronunciation lexicon extracted from the web wiktionary data. Combining the pronunciation variants generated by the 2 grapheme-to-phoneme systems provides the best results. Although the achieved results are not as good as those obtained with a hand-made pronunciation lexicon, this automatic approach makes possible an easy creation of a pronunciation lexicon for a new language [26].

Confidence measures aim at estimating the confidence of a hypothesis result provided by the speech recognition engine. Two word confidence measures were proposed, which can be computed without waiting for the end of the audio stream; one frame-synchronous and one local. Our local measures achieved performance very close to a state-of-the-art measure which requires the recognition of the whole sentence. A preliminary experiment to assess the contribution of our confidence measure in improving the comprehension of automatic transcription results by hearing impaired was also conducted [10].

6.2.1.2. Speech recognition for interaction in virtual worlds

Automatic speech recognition is investigated for vocal interaction in virtual worlds, in the context of serious games in the EMOSPEECH project. This year, a wizard-of-oz experiment was carried out to collect speech data corresponding to the dialogs from 5 players interacting with a serious game. The players were invited to speak freely to any character of the game with whom it is possible to interact, while the wizard of Oz (a game expert localized in the same room) answered them. Hence, the recorded interactions between the player and the characters of the game are natural dialogs. The audio sessions have been manually transcribed. Each session comprises roughly 30 speech turns (one player’s sentence plus one wizard’s sentence).

For training the language models, the text dialogs recorded by the TALARIS team (Midiki corpus) on the same serious game (but in a text-based interaction), have been used on addition of available broadcast news corpus. For this purpose we have also manually corrected the Midiki sentences, in order to handle the numerous typos and misspellings as well as chat specific "words" such as smileys ("mdr" or "lol"), emphasized punctuations ("!!!!!") or over-segmentations such as "é-lec-tro-nique". This normalization step is a strong requirement for speech recognition models. Different language models have then been created using different vocabulary sizes.
The acoustic models are adapted from the radio broadcast news models, using state-of-the-art Maximum A Posteriori adaptation algorithm. This reduces the mismatch in recording conditions between the game devices and the original models trained on radio streams. We are currently investigating solutions to integrate this adaptation within the speech recognition component and perform it online. At runtime, the targeted strategy is to ask the player to utter some few predefined sentences and to use these sentences to adapt the generic acoustic models to the player’s voice.

6.2.2. Speech recognition modeling

Robustness of speech recognition to multiple sources of speech variability is one of the most difficult challenges that limits the development of speech recognition technologies. We are actively contributing to this area via the development of the following advanced modeling approaches.

6.2.2.1. Detection of Out-Of-Vocabulary words

One of the key problems for large vocabulary continuous speech recognition is the occurrence of speech segments that are not modeled by the knowledge sources of the system. An important type of such segments are so-called Out-Of-Vocabulary (OOV) words (words that are not included in the lexicon of the recognizer). Mostly OOV words yield more than one error in the transcription result because the error can propagate due to the language model.

We have investigated, with Frederik Stouten (postdoctoral), to what extent OOV words can be detected. For this we used a classifier that makes a decision about each speech frame whether it belongs to an OOV word or not. Acoustic features for this classifier are derived from three recognition systems. On top of the acoustic features we also used four language model features: the ngram probability, the order of the gram that was used to calculate the language model probability, the unigram probability for the current word and a binary indicator that takes the value one if the word is preceded by a first name.

We propose to exploit the fact that 38% of the OOV word observations in the broadcast news data are pronounced more than one time in a time period of less than 1 minute. To improve the detection of repeated OOV words, we design a clustering module working on the detected OOV word segments. This algorithm is based on the estimation of the entropy. The proposed incremental clustering algorithm has been evaluated on the broadcast news corpus ESTER and gave better performance than a classical baseline incremental clustering algorithm based on a distance threshold [36].

6.2.2.2. Detailed modeling exploiting uncertainty

Modeling pronunciation variation is an important topic for automatic speech recognition. It has been widely observed that speech recognition performance degrades notably on spontaneous speech, and more precisely, that the word error rate increases when the degree of spontaneity increases. The rate of speech is also an important variability source which impacts notably on the acoustic realization of the sounds as well as on the pronunciation of the words, and consequently affects recognition performance. Large increases in word error rates are observed when speaking rate increases. And, it should be noted that rate of speech and spontaneous speech are not completely independent as the rate of speech is an important cue for detecting spontaneous speech.

This year, we have investigated further the detailed modeling of the probabilities of pronunciation variants for large vocabulary continuous speech recognition, and evaluated it on broadcast news transcriptions. In particular we have refined the modeling of the probabilities of the pronunciation variants dependent on the speaking rate. This was achieved by taking into account the uncertainty in the estimation of the speaking rate that results from the word and phoneme boundary uncertainty (speech signal - phoneme alignment errors). Such uncertainty was handled both in the training process and in the decoding step, leading to speech recognition performance improvements [25].

Detailed acoustic modeling was also investigated using automatic classification of speaker data. With such an approach it is possible to go beyond the traditional four class models (male vs female, studio quality vs telephone quality). However, as the amount of training data for each class gets smaller when the number of classes increases, this limits the amount of classes that can efficiently be trained. Hence, this year we have
investigated introducing a classification marging in the classification process. With such a marging, which handle boundary classification uncertainty, speech data at the class-boundary may belong to several classes. This increases the amount of training data in each class, which makes the class acoustic model parameters more reliable, and finally improved the overall recognition performance.

6.2.2.3. Speech recognition using distant recording

Speech recognition of distant recording of speech commands was investigated. A set of domotic commands were recorded from a few speakers using a far talking microphone. Acoustic models were adapted to this context using some training data played with a loud speaker, and recorded using a distant microphone. Among other results, preliminary experiments showed the benefit of adapting the models, as well as using a noise robust acoustic analysis when dealing with noisy data.

6.2.2.4. Training HMM acoustic models

At the beginning of his second internship at INRIA Nancy research laboratory, Othman Lachhab focused on the finalization of a speech recognition system based on context-independent HMMs models, using bigram probabilities for the phonotactic constraints and a model of duration following a normal distribution $\mathcal{N}(\mu, \sigma^2)$ incorporated directly in the Viterbi search process. Currently, he built a reference system for speaker-independent continuous phone recognition using Context-Independent Continuous Density HMM (CI-CDHMM) modeled by Gaussian Mixture Models (GMMs). In this system he developed his own training technique, based on a statistical algorithm estimating the classical optimal parameters. This new training process compares favorably with already published HMM technology on the same test corpus (TIMIT).

6.2.3. Speech/text alignment

6.2.3.1. Alignment with native speech

Speech to text alignement is a research objective that is derived from speech recognition. While it seems easier to solve at first sight, expectations are also higher and new problems appear, such as how to handle very large audio documents, or how to handle out-of-vocabulary words. Another important challenge that motivated our work in this area concerns how to improve our results and meet the user expectation by exploiting as much as possible the interactions and feedback loop between the end-user and the system. This year, we kept on improving the open-source JTrans software platform for this task as described in Section (see section 5.6 ). We further submitted an ANR Corpus proposal in collaboration with University Paris 3. We also sent a new version of the software to the "Timecode" company to help them investigating the usefulness of this approach in the application context of foreign film dubbing (see section 7.4.1 ).

6.2.3.2. Alignment with non-native speech

Non-native speech alignment with text is one critical step in computer assisted foreign language learning. The alignment is necessary to analyze the learner’s utterance, in view of providing some prosody feedback (as for example bad duration of some syllables - too short or too long -). However, non-native speech alignment with text is much more complicated than native speech alignment. This is due to the pronunciation deviations observed on non-native speech, as for example the replacement of some target language phonemes by phonemes of the mother tongue, as well as errors in the pronunciations. Moreover, these pronunciation deviations are strongly speaker dependent (i.e. they depend on the mother tongue of the speaker, and on its fluency in the target foreign lanaguage) which makes their prediction difficult.

In this application context, the precision of phoneme boundaries is critical. Hence, speech-text alignment was investigated on non-native speech. A large non-native speech corpus has been manually segmented for building a reference corpus. Then automatic phonetic segmentation (resulting from the speech-text alignment) has been analyzed. The results shows that rather reliable boundaries are obtained for some phonetic classes [31] and that better results are obtained when only frequent pronunciation deviations are kept as variants in the pronunciation lexicon [27]. Further work is on-going to determine automatically a confidence value on the proposed alignments.
6.2.4. Computing and merging linguistic information on speech transcripts

The raw output of speech recognition is difficult to read for humans, and difficult to exploit for further automatic processing. We thus investigated solutions to enrich speech recognition outputs with non-lexical information, such as dialog acts, punctuation marks and syntactic dependencies. Computing such a linguistic information requires a corpus to train stochastic models, and we also worked out new semi-supervised training algorithms for building a French corpus dedicated to syntactic parsing of oral speech. The creation of this corpus is realized in collaboration with the TALARIS team. Finally, we designed a new solution to improve our core language models by integrating into them lexical semantic distances.

An important information for post-processing speech transcripts concerns dialog acts and punctuation marks. We initiated some work in this area several years ago with the Ph.D. thesis of Pavel Kral. Since then, we continued our collaboration in this domain by successively investigating specific challenges, such as finding the most relevant features, models and testing the adaptation of our approaches in two languages, Czech and French [59]. We further proposed this year an approach to improve commas generation with the help of syntactic features [17].

Inferring syntactic dependencies is an extremely important step towards structuring the text and an absolute prerequisite for working with relations between words and next interpreting the utterance. Yet, no state-of-the-art solutions designed for parsing written texts can be reliably adapted to parsing speech, and even less transcribed speech. The lack of such methods and resources is especially blatant in French. We started, in collaboration with the TALARIS team, to address this issue by building a new French treebank dedicated to speech parsing [52], as well as a software platform dedicated to working with this corpus (see section 5.5). We exploited this year this corpus to study specific syntactic structures, such as negations (Master internship in 2011) and left dislocations in French [13].

While a large part of our work is dedicated to enriching the output of our speech recognition system, we also tried integrating within the speech decoding process itself new information coming from the higher-levels. We thus extended the new approach proposed in 2010 about language model smoothing with a new probabilistic smoothing that takes into account much longer words history thanks to a Levenshtein-based clustering of the training sentences [20].

6.3. Speech-to-Speech Translation and Language Modeling

Participants: Kamel Smaïli, David Langlois, Sylvain Raybaud.

Our work on Confidence Measures is now published in Machine Translation [9]. Now we are working on Speech-to-Text translation. We have proposed a method to segment audio input stream for machine translation. First, audio stream is split into overlapping segments useful for speech recognition; then, these segments are transcribed and regrouped; last, the obtained text flow is segmented into machine translation-friendly segments and translated. We incorporated this work into our speech-to-text machine translation system, and we evaluated our system for French-English broadcast news translation [34]. The following step consists in integrating Confidence Measures into the system in order to improve the integration between the both recognition-translation processes.

Moreover, we pursued our collaboration with Chiraz Latiri from the URPAH Team, University of Tunis. Running on our previous works (based on word-based machine translation system) we compared our respective methods in the scope of phrase-based machine translation [30].
6. New Results

6.1. Calibration of a mixed camera system
An approximately Euclidean representation of the visible scene can be obtained directly from a range, or
time-of-flight, camera. An uncalibrated binocular system, in contrast, gives only a projective reconstruction of
the scene. This paper analyzes the geometric mapping between the two representations, without requiring an
intermediate calibration of the binocular system. The mapping can be found by either of two new methods, one
of which requires point correspondences between the range and colour cameras, and one of which does not. It
is shown that these methods can be used to reproject the range data into the binocular images, which makes it
possible to associate high-resolution colour and texture with each point in the Euclidean representation.

6.2. Computation of scene flow
A simple seed growing algorithm for estimating scene flow in a stereo setup is presented. Two calibrated and
synchronized cameras observe a scene and output a sequence of image pairs. The algorithm simultaneously
computes a disparity map between the image pairs and optical flow maps between consecutive images. This,
together with calibration data, is an equivalent representation of the 3D scene flow, i.e. a 3D velocity vector
is associated with each reconstructed point. The proposed method starts from correspondence seeds and
propagates these correspondences to their neighborhood. It is accurate for complex scenes with large motions
and produces temporally-coherent stereo disparity and optical flow results. The algorithm is fast due to inherent
search space reduction. An explicit comparison with recent methods of spatiotemporal stereo and variational
optical and scene flow is provided.

6.3. 3D shape analysis and registration
We address the problem of 3D shape registration and we propose a novel technique based on spectral graph
theory and probabilistic matching. Recent advancement in shape acquisition technology has led to the capture
of large amounts of 3D data. Existing real-time multi-camera 3D acquisition methods provide a frame-
wise reliable visual-hull or mesh representations for real 3D animation sequences. The task of 3D shape
analysis involves tracking, recognition, registration, etc. Analyzing 3D data in a single framework is still a
challenging task considering the large variability of the data gathered with different acquisition devices. 3D
shape registration is one such challenging shape analysis task. The main contribution of this chapter is to
extend the spectral graph matching methods to very large graphs by combining spectral graph matching with
Laplacian embedding. Since the embedded representation of a graph is obtained by dimensionality reduction
we claim that the existing spectral-based methods are not easily applicable. We discuss solutions for the
exact and inexact graph isomorphism problems and recall the main spectral properties of the combinatorial
graph Laplacian; We provide a novel analysis of the commute-time embedding that allows us to interpret the
latter in terms of the PCA of a graph, and to select the appropriate dimension of the associated embedded
metric space; We derive a unit hyper-sphere normalization for the commute-time embedding that allows us to
register two shapes with different samplings; We propose a novel method to find the eigenvalue-eigenvector
ordering and the eigenvector sign using the eigensignature (histogram) which is invariant to the isometric
shape deformations and fits well in the spectral graph matching framework, and we present a probabilistic
shape matching formulation using an expectation maximization point registration algorithm which alternates
between aligning the eigenbases and finding a vertex-to-vertex assignment.
6.4. A differential model for the complex cell

The receptive fields of simple cells in the visual cortex can be understood as linear filters. These filters can be modelled by Gabor functions, or by Gaussian derivatives. Gabor functions can also be combined in an energy model of the complex cell response. This work proposes an alternative model of the complex cell, based on Gaussian derivatives. It is most important to account for the insensitivity of the complex response to small shifts of the image. The new model uses a linear combination of the first few derivative filters, at a single position, to approximate the first derivative filter, at a series of adjacent positions. The maximum response, over all positions, gives a signal that is insensitive to small shifts of the image. This model, unlike previous approaches, is based on the scale-space theory of visual processing. In particular, the complex cell is built from filters that respond to the 2-D differential structure of the image. The computational aspects of the new model are studied in one and two dimensions, using the steerability of the Gaussian derivatives. The response of the model to basic images, such as edges and gratings, is derived formally. The response to natural images is also evaluated, using statistical measures of shift insensitivity. The relevance of the new model to the cortical image-representation is discussed.

6.5. Audiovisual fusion based on a mixture model

The problem of multimodal clustering arises whenever the data are gathered with several physically different sensors. Observations from different modalities are not necessarily aligned in the sense there there is no obvious way to associate or to compare them in some common space. A solution may consist in considering multiple clustering tasks independently for each modality. The main difficulty with such an approach is to guarantee that the unimodal clusterings are mutually consistent. In this paper we show that multimodal clustering can be addressed within a novel framework, namely conjugate mixture models. These models exploit the explicit transformations that are often available between an unobserved parameter space (objects) and each one of the observation spaces (sensors). We formulate the problem as a likelihood maximization task and we derive the associated conjugate expectation-maximization algorithm. The convergence properties of the proposed algorithm are thoroughly investigated. Several local/global optimization techniques are proposed in order to increase its convergence speed. Two initialization strategies are proposed and compared. A consistent model-selection criterion is proposed. The algorithm and its variants are tested and evaluated within the task of 3D localization of several speakers using both auditory and visual data.
6. New Results

6.1. A Lightweight Augmented Virtuality System for Providing a Faithful and Spatially Manipulable Visual Hand Representation

**Participants:** Sabine Coquillart, Olivier Martin, Andreas Pusch.

We introduced the technical foundations of a system designed to embed a lightweight, faithful and spatially manipulable representation of the user’s hand into an otherwise virtual world - Augmented Virtuality (AV). A highly intuitive control during pointing-like near space interaction can be provided to the user, as well as a very flexible means to experimenters, in a variety of contexts. Our approach essentially relies on stereoscopic video see-through Augmented Reality (AR) technology and a generic, extendible framework for managing 3-D visual hand displacements. Research from human-computer interaction, perception and motor control has contributed to the elaboration of our proposal which combines a) acting in co-location, b) avoiding occlusion violations by assuring a correct scene depth ordering and c) providing a convincing visual feedback of the user’s hand. This system has already successfully been used in one case and further promising applications are studied [17], [18].

6.2. Effects of Hand Feedback Fidelity on Near Space Pointing Performance and User Acceptance

**Participants:** Sabine Coquillart, Olivier Martin, Andreas Pusch.

We conducted an experiment to test the effects of different hand representations on near space pointing performance and user preference. Subjects were presented with varying levels of hand realism, including real hand video, a high and a low level 3D hand model and an ordinary 3D pointer arrow. Behavioural data revealed that an abstract hand substitute like a 3D pointer arrow leads to significantly larger position estimation errors in terms of lateral target overshooting when touching virtual surfaces with only visual hand movement constraints. Further, questionnaire results show that a higher fidelity hand is preferred over lower fidelity representations for different aspects of the task [18].
6. New Results

6.1. Introduction

This year Pulsar has tackled several issues related to its two main research axes: scene understanding for activity recognition and software engineering for activity recognition.

6.1.1. Scene Understanding for Activity Recognition


This year Pulsar has proposed new algorithms in computer vision (people head and face detection and people re-identification), in reasoning (activity recognition and uncertainty handling). More precisely, the new results for this research axis concern:

- People detection in monocular video sequences (6.2)
- Online Parameter Tuning for Object Tracking Algorithms (6.3)
- Fiber Based Video Segmentation (6.4)
- Multiple Birth and Cut Algorithm for Multiple Object Detection (6.5)
- Exhaustive Family of Energies Minimizable Exactly by a Graph Cut (6.6)
- Steepest Descent in Banach Spaces with Application to Piecewise-Rigid Evolution of Curves (6.7)
- Object Tracking Using a Particle Filter based on SIFT Features (6.8)
- Human Re-identification using Riemannian Manifolds (6.9)
- Global Tracking of Multiples Actors (6.10)
- Crowd Data Collection from Video Recordings (6.11)
- Events Recognition and Performance Evaluation (6.12)
- Group interaction and group tracking for video-surveillance in underground railway stations (6.13)
- Action Recognition in Videos (6.14)
- Activity Recognition Applied on Health Care Application (6.15)
- A Cognitive Vision System for Nuclear Fusion Device Monitoring (6.16)
- Scenario Recognition with depth camera (6.17)
- Trajectory Clustering for Activity Learning (6.18)

6.1.2. Software Engineering for Activity Recognition

This year Pulsar has improved the SUP platform. This latter is the backbone of the team experiments to implement the new algorithms proposed by the team in perception, understanding and learning. We improve our meta-modeling approach to support the development of video surveillance applications based on SUP. We continue the development of a scenario recognition module relying on formal methods to support activity recognition in SUP platform. We also continue to study the definition of multiple services for device adaptive platform for scenario recognition. Finally, we are implementing the new theoretical results obtained last year to improve the Clem toolkit.

The new results related to this research axis concern:

- SUP Software Platform (6.19)
- Model-Driven Engineering and Video-surveillance (6.20)
- Scenario Analysis Module (6.22)
- Multiple Services for Device Adaptive Platform for Scenario Recognition (6.23)
- The Clem Toolkit (6.24)

### 6.2. People detection in monocular video sequences

**Participants:** Etienne Corvee, François Brémond, Silviu-Tudor Serban, Vasanth Bathrinaryanan.

A video understanding system analyzes human activity by detecting people in video sequences and tracking their displacement and movement throughout the sequences. The better the detection quality, the higher the semantic level of the information is. People activity can differ greatly from one application to another e.g. the presence of a person in one zone can simply be detected from a moving pixel region in a manually specified zone whereas detecting people fighting in a subway requires more complex information. For people activity to be recognized, one needs to detect people accurately in videos and at real time frame rate. Current state of the art algorithms provide generic people detection solutions but with limited accuracy. In the people monitoring domain, although cameras remain mostly fixed, many issues occur in images. For example, outdoor scenes display strong varying lighting conditions (e.g. sunny/cloudy illumination, important shadows), public spaces can be often crowded (e.g. subways, malls) and images can be obtained with a low resolution and can be highly compressed. Hence, detecting and tracking objects in such complex environment remains a delicate task to perform. In addition, detecting people has to face one major difficulty which is caused by occlusion where important information is hidden. When people overlap onto the image plane, their foreground pixels cannot be separated using a standard thresholding operation from a background reference frame. Therefore vision algorithms need to use information held by the underlying pixels and located at specific locations such as body parts.

We have extended our work by implementing and testing a novel people, head and face detection algorithm using Local Binary Pattern based features and Haar like features. The traditional and efficient Adaboost training scheme is adopted to train object features from publicly available databases. This work has been published in the ICVS Workshop [36].

The work has been tested for group tracking in Vanaheim videos (see section 8.2.1.2) and for people tracking in VideoId videos. The VideoId project aims to re-identify people across a network of non overlapping cameras using iris, face and human appearance recognition. An example of tracked people, head and faces in a testing database is shown in figure 7. An example of re-identified face is shown in figure 8 by the VideoId interface in a Paris underground video.

We have evaluated our people detection algorithm on the test human dataset provided by INRIA against state of the art algorithms which we refer as HOG [59] and LBP-HOG [77]. The INRIA human dataset is composed of 1132 human images and 453 images of background scenes containing no human. The results are displayed in figure 9 which shows that we obtain slightly better performances than the HOG-LBP technique in terms of missed detection rate vs. FPPI i.e. False Positive Per Image. In this figure, two extreme functioning modes could be chosen: approximately 2 noisy detections are obtained every 1000 background images for 50% true positive detections or 1 noisy detection every 2 frames for a detection rate of approximately 88%.
Figure 7. Example of tracked people, head and face

Figure 8. Face recognition in Paris underground
The same evaluation scheme of people detection above is used for face detection evaluation. The FPPI rates are obtained on 997 NICTA [66] background images of 720x576 pixels. 180 faces provided by a CMU test face image database are used to evaluate true positive rates. We have compared our results with the 2 versions of Haar feature provided by the OpenCv library i.e. the standard ’default’ and alternative ’alt’ training parameters. The results in table 1 show that the Haar ’alt’ technique performs better than the traditional Haar one. And our haar based technique called CCR provides similar face detection rates while giving a less false alarm rate. The proposed approach is approximately 1% less successful in detecting faces than the Haar technique while this latter is 32% more noisier than our CCR technique.

<table>
<thead>
<tr>
<th>technique</th>
<th>TP(%)</th>
<th>FPPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haar (default)</td>
<td>91.57</td>
<td>4.132</td>
</tr>
<tr>
<td>Haar (alt)</td>
<td>92.13</td>
<td>1.685</td>
</tr>
<tr>
<td>CCR</td>
<td>91.01</td>
<td>1.274</td>
</tr>
</tbody>
</table>

6.3. Online Parameter Tuning for Object Tracking Algorithms

Participants: Duc Phu Chau, Monique Thonnat, François Brémond.

Many approaches have been proposed to track mobile objects in a scene. However the quality of tracking algorithms always depends on scene properties such as: mobile object density, contrast intensity, scene depth and object size. The selection of a tracking algorithm for an unknown scene becomes a hard task. Even when the tracker has appropriately selected, it is difficult to tune online its parameters to get the best performance.

Therefore we propose a new control approach for mobile object tracking. More precisely in order to cope with the tracking context variations, this approach learns how to tune the parameters of object appearance-based tracking algorithms. The tracking context of a video sequence is defined as a set of features: density of mobile objects, their occlusion level, their contrasts with regard to the background and their 2D areas. In an offline supervised learning phase, satisfactory tracking parameters are searched for each training video sequence. Then these video sequences are classified by clustering their contextual features. Each context cluster is associated with the learned tracking parameters. In the online control phase, two approaches are...
proposed. In the first one, once a context change is detected, the tracking parameters are tuned for the new context using the learned values. In the second approach, the parameter tuning is performed when the context changes and the tracking quality (computed by an online performance evaluation algorithm \[ 56 \]) is not good enough. An online learning process enables to update the context/parameter relations.

We have also proposed two new tracking algorithms to experiment the proposed control method. The first tracker relies on a Kalman filter and a global tracking which aims at fusing trajectories belonging to the same mobile object. This work has been published in \[ 35 \]. The second tracker relies on the similarities of eight object descriptors (2D and 3D positions, area, shape ratio, HOG, color histogram, color covariance and dominant color) to build object trajectories. This work has been published in \[ 34 \].

Figure 10. (a) CARETEKER: Illustration of the Caretaker video; (b) CAVIAR: Illustration of the Caviar video

The proposed controller has been experimented on a long, complex video belonging to the Caretaker European project \(^1\) (see figure 10 (a)) and 26 videos of Caviar dataset \(^2\) (see figure 10 (b)). For the Caretaker video, when the controller is used, the tracking quality increases from 52% to 78%. For the Caviar dataset, the experimental results show that the tracking performance increases from 78.3% to 84.4% when using the controller. The tracking results on Caviar videos with the proposed controller are as good as the ones obtained with manual parameter tuning.

6.4. Fiber Based Video Segmentation

Participants: Ratnesh Kumar, Guillaume Charpiat, Monique Thonnat.

The aim of this work is to segment objects in videos by considering videos as 3D volumetric data (space×time). Figure 11 shows 2D slices of a video volume. Bottom right corner of each figure shows the current temporal depth in the volume, while top right shows the X-time slice and bottom left shows Y-time slice. In this 3D representation of videos, points of static background form straight lines of homogeneous intensity over time, while points of moving objects form curved lines. Analogous to the fibers in MRI images of human brains, we name fibers, these straight and curved lines of homogeneous intensity. So, in our case, to segment the whole video volume data, we are interested in a dense estimation of fibers involving all pixels.

For the detection of these fibers, we use motion flow vectors and intensity correlation of 2D patches over time. As these techniques are not reliable everywhere in the image domain, we sort the fibers based on the reliability of the detections from these techniques. The subsequent goal is then to pick high ranked fibers to propagate motion information and boundary fronts to other regions of the 3D volume.

\(^1\) http://cordis.europa.eu/ist/kct/caretaker_synopsis.htm
\(^2\) http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/
To reliably propagate information from a fiber, we express the reliability of detection of a fiber and the cost of propagation of information from it. The later can be based on a distance measure of a pixel from a fiber, while reliability of a fiber involves motion coherency, color homogeneity, duration along time axis etc. Our work closely relates to [72]. A video is represented by a set of particles (trajectory of an image point sample). The algorithm then extends and truncates particle trajectories to model motion near occlusion boundaries.

Figure 12 shows some straight fibers found in a video volume. The reliability of these fibers is based on temporal length. Fibers which have temporal span same as that of the video are colored in green, while fibers which have temporal span of less than 10% of the video are colored blue. Red colored fibers have temporal length in between green and blue colored fibers.

6.5. Multiple Birth and Cut Algorithm for Multiple Object Detection

Participant: Guillaume Charpiat.

In collaboration with the Ariana team (Ahmed Gamal-Eldin, Xavier Descombes and Josiane Zerubia), we developed a new optimization method which we call Multiple Birth and Cut (MBC). It combines the recently proposed Multiple Birth and Death (MBD) algorithm and the Graph-Cut algorithm. MBD and MBC optimization methods are applied to energy minimization of an object based model, the marked point process. The most important advantage of the MBC over MBD is the reduction of number of parameters. By proposing good candidates throughout the selection phase in the birth step, the speed of convergence is increased. In
this selection phase, the best candidates are chosen from object sets by a belief propagation algorithm. The algorithm is applied on the flamingo counting problem in a colony [37], [26].

6.6. Exhaustive Family of Energies Minimizable Exactly by a Graph Cut

Participant: Guillaume Charpiat.

Graph cuts are widely used in many fields of computer vision in order to minimize in small polynomial time complexity certain classes of energies. These specific classes depend on the way chosen to build the graphs representing the problems to solve. We study here all possible ways of building graphs and the associated energies minimized, leading to the exhaustive family of energies minimizable exactly by a graph cut. To do this, we consider the issue of coding pixel labels as states of the graph, i.e. the choice of state interpretations. The family obtained comprises many new classes, in particular energies that do not satisfy the submodularity condition, including energies that are even not permuted-submodular.

We studied in details a generating subfamily, in particular we proposed a canonical form to represent Markov random fields, which proves useful to recognize energies in this subfamily in linear complexity almost surely, and then to build the associated graph in quasilinear time. We performed a few experiments to illustrate the new possibilities offered [33]. We have also started to use this technique to minimize exactly approximations of Markov random field energies instead of minimizing approximately the exact energies, by projecting energies on the family we know to solve globally efficiently.

6.7. Steepest Descent in Banach Spaces with Application to Piecewise-Rigid Evolution of Curves

Participant: Guillaume Charpiat.

This is joint work with Gabriel Peyré (CNRS, Ceremade, Université Paris-Dauphine). We intend to favor piecewise-rigid motions, i.e. articulated movements, during shape evolutions, especially when computing morphings or image segmentation with shape prior. To do this, we first need a dissimilarity measure between shapes, whose gradient is meaningful. We formulate one using kernels and bistochastization.

The parameters of these kernels are automatically estimated in a fixed-point scheme that guarantees physical relevance, and the notion of bistochastization is extended to continuous distributions. Finally, piecewise rigidity is ensured during gradient descents by a change of the norm from which the gradient is derived. This norm is formulated so as to favor sparse second derivatives, which produces articulated movements without knowing by advance the location of the articulations.

The formula of the norm is actually elegantly simple, involving simple geometric quantities, derivatives, and the $L_1$ norm. Note that this norm does not derive from an inner product but defines a gradient in the sense of [5] as the minimizer of an energy. It turns out that in our case the energy defining the gradient is actually convex, and efficient minimization follows.

6.8. Object Tracking Using a Particle Filter based on SIFT Features

Participants: Malik Souded, François Brémont.

The approach consists in detecting SIFT points of interest on the objects to track, calculating their SIFT descriptors, tracking these points with a particle filter, and finally achieving tracking process by linking them along the time with links which are weighted by measures on SIFT descriptors and reliability.

The main contributions in this work are on three points.

The first point consists in techniques of detection and selection of SIFT points, allowing a better distribution of points of interest on the target and allowing better management of partial occlusions, and secondly an optimized computing time thanks to the parallelization of these SIFT computation on modern processors (see figure 13).
The second point concerns the weighting of the particles during tracking. This is done with a combination of two kinds of information: the similarity measure of the SIFT descriptor and the state of motion of pixels corresponding to the particles. This allows more robust tracking of SIFT points regardless of the quality of the background subtraction providing the detected objects.

The last point concerns the selection of temporal links between tracked objects and detected ones. These links are selected according to their weight. The weight of each link is based on the proportion of common SIFT points to both objects (two successive images) potentially linked, and the reliability of each of these SIFT point. This reliability is calculated for each point by measuring the variation of the SIFT descriptor during the tracking time.

The occlusion management is performed using three types of information: SIFT descriptors used for tracking (matching after reappearance) the dominant colors of the object of interest and finally the width, height, and speed (in real world) of the object, which are learned in Gaussian models during the tracking (tracking being used with video cameras which have been calibrated), see figure 14.

The approach was tested on 121 sequences of four different datasets: 80 sequences from CAVIAR, 34 sequences from ETISEO, 3 sequences from PETS2001 and 2 sequences from VS_PETS2003. The obtained results are satisfying. The comparison of these results with the state of the art shows improvements for the benchmarking dataset (ETISEO) The following table compares the proposed approach and state of the art results on ETISEO data base:

<table>
<thead>
<tr>
<th></th>
<th>Metrics</th>
<th>ETI-VS1-BE-18-C4</th>
<th>ETI-VS1-BE-16-C4</th>
<th>ETI-VS1-MO-7-C1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed tracker</td>
<td>M1</td>
<td>0.68</td>
<td>0.54</td>
<td>0.90</td>
</tr>
<tr>
<td>Chau et al. (VISAPP 2011)</td>
<td>M1</td>
<td>0.64</td>
<td>0.36</td>
<td>0.87</td>
</tr>
<tr>
<td>Best team in ETISEO Project (2005)</td>
<td>M1</td>
<td>0.48</td>
<td>0.44</td>
<td>0.77</td>
</tr>
</tbody>
</table>

This work was published in [49]

### 6.9. Human Re-identification using Riemannian Manifolds

**Participants:** Sławomir Bąk, Etienne Corvée, François Brémond, Monique Thonnat.

This work addresses the human re-identification problem, which is defined as a requirement to determine whether a given individual has already appeared over a network of cameras. This problem is particularly hard by significant appearance changes across different camera views. In order to re-identify people, a human signature should handle difference in illumination, pose and camera parameters.
We propose new appearance models based on the mean riemannian covariance (MRC) matrices combining the appearance information from multiple images. These mean covariance matrices not only keep information on feature distribution but also carry out essential cues about temporal changes of an appearance. Using MRC-s, we propose two methods for an appearance representation:

- **Learned Covariance Patches (LCP)** [25] - a distinctive representation is extracted by a boosting scheme. The structure of MRC patches (size, position) is learned using boosting algorithm based on confidence-rated predictions (see Fig. 15 (a)). These confidence-rated coefficients are employed to weight appearance characteristics of a specific individual w.r.t. the reference (training) dataset of humans.

- **Mean Riemannian Covariance Grid (MRCG)** [30] - less computationally demanding technique than LCP. We represent a human appearance by a grid of MRC cells (see Fig. 15 (b)). Relevant cells are identified by an efficient discriminant analysis. This analysis takes into account variance of MRC patch in the class of humans (reference dataset). MRC is assumed to be more significant when its variance is larger in the class of humans: (1) the most common patterns belong to the background (the variance is small); (2) the patterns which are far from the rest are at the same time the most discriminative (the variance is large). All operations, such as mean or variance, are performed on covariance manifold specified as Riemannian.

Both methods are evaluated and compared with the state of the art using publicly available datasets. We demonstrate that the proposed approaches outperform state of the art methods. Further, we extract new sets of individuals from i-LIDS data to investigate more carefully advantages of using many images for human re-identification.

The computation complexity is analyzed in the context of distance operator between two signatures. Comparing two human signatures, it is necessary to compute distance between covariance matrices, which requires solving the generalized eigenvalues problem. This operation is computationally heavy. In [31], we propose an implementation for finding generalized eigenvalues and eigenvectors for distance operator, using NVIDIA GPU architecture. We improve significantly the performance, reaching 66 speedup using Tesla S1070.
6.10. Global Tracking of Multiples Actors

Participants: Julien Badie, François Brémond.

We propose a new approach for long term tracking of individuals. Our main objective is to design a tracking algorithm for people reidentification [30] that can track people even if they come back in the scene after leaving it. This algorithm is based on covariance matrix and we have also added some contextual information of the scene (for instance, zones where people can enter the scene) to improve tracking performance. In addition, a basic noise detection system and a tracking correction system are proposed in order to handle short-term tracking errors such as multiplication of IDs corresponding to only one individual. The noise detection system is designed to find and remove objects that are detected in a very small number of consecutive frames (for instance 4) and disappear afterward. The tracking correction system associates IDs recently lost with IDs that have just started to be tracked based on geometrical features and 3D distance criteria.

As a result, the tracking quality is significantly improved on 5 video sequences tested from the ETISEO dataset[^3]. The people reidentification algorithm gives encouraging results for future work. The number of IDs associated to one individual is reduced (on average 50% less) and the tracking quality improves due to the IDs stability. This algorithm can detect not only people re-entering the scene but also trajectory interruptions due to occlusions or misdetections. This approach could enable the detection of new kinds of events on video sequences such as long range people tracking on a camera network.

6.11. Crowd Data Collection from Video Recordings

Participants: Jihed Joobeur, François Brémond.

The aim of this work is to analyze crowd behaviors by detecting specific situations: panic, congestion, fighting etc. We validate our work with subway station videos from VANAHEIM project. We use Mixture of Gaussian based segmentation to extract moving point and then detecting moving objects. Subsequently inside these moving objects we detect FAST feature points and compute HOG descriptors for tracking these points. We compute different features based on these points like speed and orientation. To estimate the crowd density we use features based on Grey-Level Co-occurrence Matrix. As these features depend on the distance of people from the camera, we divide the scene into different zones which have each zone same distance from the

[^3]: http://www-sop.inria.fr/orion/ETISEO/
Figure 16. The green lines are the ground-truth and the blue lines are the tracking results using both reidentification and noise removal (left) and with only noise removal (right). The man is tracked throughout the whole video with a single ID (68).
camera. In each area, compiling all the information on speed, direction and learned over a threshold density of the crowd, we can learn and detect different situations. For example, if the density increases and the average speed decreases in a pre-defined zone, that may correspond to a congestion situation.

On figure 17 the FAST feature points are shown in blue points, while the tracking of these points is shown in yellow.

![Figure 17. Feature points detection and tracking in different zones.](image)

### 6.12. Events Recognition and Performance Evaluation

**Participants:** Ricardo Cezar Bonfim Rodrigues, François Brémond.

The goal of this work is to evaluate the accuracy and performance of events detection, see workflow in Figure 18. The experiments will be performed using the tools developed in Pulsar team, such as Scene Understand Platform (SUP) a plugin for events detection [50] and ViseVal 5.

The experiments were performed using video sequences of a subway station (VANAHEIN dataset) where the goal was to detect events such as people waiting, entering, buying tickets and so on. Preliminary results showed a very low accuracy and demonstrated that the scenario configuration parameters are very sensitive in this problem. It means many of the expected events were missed or misclassified, specially composite events (when more than one activity recognition is required) see some issues on Figure 19.

Based on the issues, a second experiment was configured using 3 different video sequences. In this new experiment, the scenario was adjusted to give more tolerances to people detection issues, the camera calibration was refined and some events were remodeled. After these changes the results were significantly improved. This last experiments showed that the engine proposed by Pulsar team is able to detect events accurately however events modeling can be very sensitive to the scenario configuration, see the results in Table 3.

<table>
<thead>
<tr>
<th>Sequence 1</th>
<th>Sequence 2</th>
<th>Sequence 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision (global)</td>
<td>0.73</td>
<td>1.00</td>
</tr>
<tr>
<td>Sensitivity (global)</td>
<td>0.82</td>
<td>0.90</td>
</tr>
</tbody>
</table>

---

4 http://raweb.inria.fr/rapportsactivite/RA2010/pulsar/uid27.html
5 http://www-sop.inria.fr/teams/pulsar/EvaluationTool/VSEvAI_Description.html
Figure 18. Events detection workflow

Figure 19. Blue bounding boxes correspond to Annotations (Reference data) and red detected objects. It's possible to observe if there is a detection issue on the left image.
6.13. Group interaction and group tracking for video-surveillance in underground railway stations

Participants: Sofia Zaidenberg, Bernard Boulay, Carolina Garate, Duc-Phu Chau, Etienne Corvée, François Brémond.

One goal in the European project VANAHEIM is the tracking of groups of people. Based on frame to frame mobile object tracking, we try to detect which mobiles form a group and to follow the group through its lifetime. We define a group of people as two or more people being close to each other and having similar trajectories (speed and direction). The dynamics of a group can be more or less erratic: people may join or split from the group, one or more can disappear temporarily (occlusion or disappearance from the field of view) but reappear and still be part of the group. The motion detector which detects and labels mobile objects may also fail (misdetections or wrong labels). Analyzing trajectories over a temporal window allows handling this instability more robustly. We use the event-description language described in [50] to define events, described using basic group properties such as size, type of trajectory or number and density of people and perform the recognition of events and behaviors such as violence or vandalism (alarming events) or a queue at the vending machine (non-alarming events). Two approaches to this problem have been implemented. The first approach takes as input the frame-to-frame tracking results of individual mobiles and tries to gather them into groups based on their trajectories through the temporal window. Each group has a coherence coefficient. This coefficient is a weighted sum of three quantities characterizing a group: the group density (average of distances between mobiles), the similarity of mobile’s speed and the similarity of their motion directions. The update of a group consists in re-calculating the group coherence with new mobiles from the current frame. If adding the mobile does not put the coherence under a defined threshold, the mobiles are added to the group. A pre-selection is made by only considering mobiles that are close enough to the center of gravity of the group. After the update step, all mobiles that have not been assigned to a group are analyzed to form new groups if possible.

A first improvement has been done by integrating the use of the LBP-based people detector described in [36]. This makes the algorithm more robust to false detections such as train doors closing. But on the other hand, it also introduces false negatives as, among other things, people are only detected if fully visible in the image. The group tracking algorithm has been tested both with the original, background subtraction-based mobile object detection (noted S hereafter) and the LBP-based people detection (noted LBP hereafter).

For evaluating the detection, we used 3 annotated sequences: Sequence 1 is a short sequence of 128 frames with just one ground truth object (one group), Sequence 2 has 1373 frames and 9 ground truth objects, and Sequence 3 is 17992 frames long and 25 ground truth objects were annotated. Detection and tracking results are shown in table 4.

The whole algorithm chain has been integrated into the common VANAHEIM platform and sent to partners for pre-integration.

We also used videos from the ViCoMo project, recorded in the Eindhoven airport to test our approach. No formal evaluation has been done yet on these sequences due to the lack of ground truth. Nevertheless, these videos contain several acted scenes which could be successfully recognized: groups merging, splitting and entering a forbidden zone.

This work has been published in [50].

In parallel, a new approach is being developed, making use of a long-term tracker described in [35]. This tracker provides more robust individual trajectories to the group tracker, containing less confusions in cases where people cross each other. We apply the Mean Shift clustering algorithm on trajectories of people through a sliding time window (e.g. 10 frames). If the target is lost in one or several frames, we interpolate its positions. The clustering brings together mobiles having similar trajectories, which is our definition of a group. At each frame, clusters are calculated and then a matching is done to associate clusters to existing groups in the previous frame, and thus track groups. Looking backwards (within a window) on the trajectory of a mobile we might find a mobile on that trajectory that belongs to a group. If such a group is found, it is called the probable
### Table 4. Segmentation (S) and Human Detector (HD) Results

<table>
<thead>
<tr>
<th></th>
<th>Sequence 1</th>
<th></th>
<th>Sequence 2</th>
<th></th>
<th>Sequence 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S</td>
<td>HD</td>
<td>S</td>
<td>HD</td>
<td>S</td>
<td>HD</td>
</tr>
<tr>
<td>True Positives (TP)</td>
<td>72</td>
<td>67</td>
<td>1395</td>
<td>1079</td>
<td>5635</td>
<td>3679</td>
</tr>
<tr>
<td>False Positives (FP)</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>111</td>
<td>1213</td>
<td>642</td>
</tr>
<tr>
<td>False Negatives (FN)</td>
<td>6</td>
<td>11</td>
<td>269</td>
<td>585</td>
<td>3686</td>
<td>5642</td>
</tr>
<tr>
<td>Precision (global)</td>
<td>1</td>
<td>1</td>
<td>0.99</td>
<td>0.90</td>
<td>0.82</td>
<td>0.85</td>
</tr>
<tr>
<td>Sensitivity (global)</td>
<td>0.92</td>
<td>0.84</td>
<td>0.83</td>
<td>0.65</td>
<td>0.60</td>
<td>0.40</td>
</tr>
<tr>
<td>Tracking confusion</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.99</td>
<td>0.92</td>
<td>0.96</td>
</tr>
</tbody>
</table>

**group** of the current mobile. Each trajectory cluster is associated to the group that is the probable group of most mobiles in the cluster. Several clusters may be associated to the same group. This cluster association makes the algorithm robust to cases where one or several mobiles temporarily separate from the group. If the separation is longer than the time window, the probable group of these mobiles will be empty and a split will be detected.

Additionally, we work on improving the people detection by combining both methods: background subtraction-based and LBP-based. We compare overlapping mobiles from both methods and choose the best one based on their respective confidence values and their sizes. If a target was detected by only one of the two methods, we keep the target given that the confidence is high enough. If a mobile from the background subtraction method is big enough to cover several LBP-detected people (the LBP-based people detection output targets have the size of a human, whereas the background subtraction can detect a bigger mobile with the size of a **GROUP_OF_PEOPLE**), we attach the LBP-people as sub-mobiles of the group mobile so no information is lost. This method is a work in progress and no evaluation have been done yet.

Figure 20 shows two examples of group and event detection.


**Participants:** Piotr Bilinski, François Brémond.

The aim of this work is to learn and recognize short human actions in videos. We perform an extensive evaluation of local spatio-temporal descriptors, then we propose a new action recognition approach for RGB camera videos. We also propose a new approach for RGB-D cameras. For all our experiments, we develop an evaluation framework based on the bag-of-words model, SVM and cross-validation technique. We use the bag-of-words model to represent actions in videos and we use non-linear multi-class Support Vector Machines together with leave-one-person-out cross-validation technique to perform action classification.

Local spatio-temporal descriptors have shown to obtain very good performance for action recognition in videos. Over the last years, many different descriptors have been proposed. They are usually evaluated using too specific experimental methods and using different datasets. Moreover, existing evaluations make assumptions that do not allow to fully compare descriptors. In order to explore capabilities of descriptors, we perform an extensive evaluation of local spatio-temporal descriptors for action recognition in videos.

Four widely used state-of-the-art descriptors (HOG, HOF, HOG-HOF and HOG3D) and four video datasets (Weizmann, KTH, ADL and KECK) have been selected. In contrast to other evaluations, we test all the computed descriptors, we perform experiments on several codebook sizes and use several datasets, differing in difficulty. Our results show how the recognition rate depends on the codebook size and the dataset. We
observe that usually the HOG descriptor alone performs the worst but outperforms other descriptors when it is combined with the HOF descriptor. Also, we observe that smaller codebook sizes lead to consistently good performance across different datasets. This work has been published in [32].

We also propose a new action recognition method for RGB camera videos based on feature point tracking and a new head estimation algorithm. We track feature points along a video and compute appearance features (HOG-HOF) for each trajectory. Additionally, we estimate a head position for each visible human in the video, using the following chain: segmentation, person, head and face detectors. Finally, we create an action descriptor based on the combination of all these sources of information. Our approach has been evaluated on several datasets, including two benchmarking datasets: KTH and ADL, and our new action recognition dataset. This new dataset has been created in cooperation with the CHU Nice Hospital. It refers to people performing daily living activities like: standing up, sitting down, walking, reading a magazine etc.

We also study the usefulness of low-cost RGB-D camera for action recognition task. We propose a new action recognition method using both RGB and depth information. We track feature points using RGB videos and represent trajectories in a four-dimensional space using additionally depth information. Experiments have been successfully performed on our new RGB-D action recognition dataset, recorded using Microsoft’s Kinect device.

6.15. Activity Recognition Applied on Health Care Application

Participants: Rim Romdhane, Veronique Joumier, François Brémond.

The aim of this work is to propose a constraint-based approach for video event recognition with probabilistic reasoning for handling uncertainty. This work was validated on health care applications.

6.15.1. Event Recognition

We propose an activity recognition framework which is able to recognize composite events with complex temporal relationships. We consider different aspects of the uncertainty of the recognition during the event modeling and the event recognition process to overcome the noise or missing observations which characterize real world applications.
To reach this goal, we manage the uncertainty in the event modeling and event recognition processes by a combination of logical and probabilistic reasoning for handling uncertainty. We improve the event description language developed in Pulsar team and introduce a new probabilistic description based approach to gain in flexibility for event modeling by adding the notion of utility. Utility expresses the importance of sub-events to the recognition of the whole event. We compute the probability of recognition for both primitive (i.e. elementary) events and composite events based on Bayesian theory.

We compute the probability that the event $e$ is recognized given a sequence of observations $O$ as described in [48]. The observations consist of the set of the physical objects $p_{o_i}$ moving in the scene. If the probability of an event is over a predefined threshold, the event is recognized.

6.15.2. Health Care Application

The proposed event recognition approach is validated using the videos from the health care application SWEETHOME (http://cmrr-nice.fr/sweethome/) and CIUSante (https://extranet.chu-nice.fr/ciu-sante). We have worked in close collaboration with clinicians from Nice hospital to evaluate the behaviours of Alzheimer patients. We have first model 69 event models for health care application using our event modeling formalism. With the help of clinicians we have established a scenario protocol. The scenario is composed of three parts: (1) directed activities (10 min), (2) semi directed activities (20 min), (3) free activities (30 min). Experiments have been performed in a room of Nice hospital equipped with 2 video cameras where 45 elderly volunteers have spent between 15 min to 1 hour. Volunteers include Alzheimer patients, MCI (mild cognitive impairment) and healthy elderly.

The study described in [38] and [27] shows the ability of the proposed automatic video activity recognition system to detect activity changes between elderly subjects with and without dementia during a clinical experimentation. A total of 28 volunteers (11 healthy elderly subjects, 17 Alzheimer’s disease patients (AD)) participate to the experimentation. The proposed study shows that we could differentiate the two profiles of participants based on motor activity parameters, such as the duration of the recognized activities, the strike length and the walking speed, computed from the proposed automatic video activity recognition system. These primary results are promising and validating the interest of automatic analysis of video as an objective evaluation tool providing comparative results between participants and over the time.


Participants: Guillaume Charpiat, Vincent Martin, François Brémond, Monique Thonnat.

In collaboration with Victor Moncada, Jean-Marcel Traver and Thierry Loarer (CEA Cadarache), we propose a cognitive vision-based system for the intelligent monitoring of tokamaks during plasma operation, based on multi-sensor data analysis and symbolic reasoning. The practical purpose is to detect and characterize in real time abnormal events such as hot spots measured through infrared images of the in-vessel components in order to take adequate decisions. Our system is made intelligent by the use of a priori knowledge of both contextual and perceptual information for ontology-driven event modeling and task-oriented event recognition. The system is made original by combining both physics-based and perceptual information during the recognition process. Real time reasoning is achieved thanks to task-level software optimizations. The framework is generic and can be easily adapted to different fusion device environments. The developed system and its achievements on real data of the Tore Supra tokamak imaging system can be found in [39].

6.17. Scenario Recognition with depth camera

Participants: Bernard Boulay, Daniel Zullo, Swaminathan Sankaranarayanan, François Brémond.

Thanks to Microsoft and its kinect sensor, RGB-depth camera becomes popular and accessible. The basic idea of depth camera is to combine a visible camera, with an IR camera associated to a laser to determine the depth of each image pixel. This kind of sensor is well adapted for applications which monitor people (e.g. monitoring Alzheimer patient in hospital): because the people are in a predefined area and near the camera.
The depth cameras have two main advantages: first, the output image contains depth information and second, the sensor is independent from the light changes (IR sensor).

In our work, we propose to use the kinect sensor to acquire 3D images, detect the people and recognize interesting activities (see Figure 21).

![Figure 21. People detection and activity recognition in day and night conditions with kinect sensor.](image)

The nestk library is used to manage the kinect sensor. This library is based on OpenNI framework (an open source driver) to acquire the images. Moreover the library is able to compute some treatments (e.g. people detection) and to provide a true 3D map of the scene in the referential of the kinect.

Basic attributes are computed for each detected person: 3D position, 3D height,... These attributes are then used to compute more complex information: speed, global posture to recognize interesting activities thanks to the ScreKs framework (scenario recognition based on expert knowledge) of the SUP platform as walking, stopping, standing, sitting,... (following a protocol delivered by doctors). Then we have a plug and play system able to recognize basic activities associated to a person.

Moreover, if information on the scene, as interesting zones, or equipment are available, complex activities can be recognized as nurse explaining exercise or nurse switching off the light.

The next step, is to use the human skeleton detection to recognize precisely the posture of the patient in order to understand more precise activities and infer a behaviour model.

6.18. Trajectory Clustering for Activity Learning

Participants: Jose Luis Patino, Guido Pusiol, Hervé Falciani, Nedra Nefzi, François Brémond, Monique Thonnat.

The discovery, in an unsupervised manner, of significant activities observed from a video sequence, and its activity model learning, are of central importance to build up on a reliable activity recognition system. We have deepened our studies on activity extraction employing trajectory information. In previous work we have shown that rich descriptors can be derived from trajectories; they help us to analyze the scene occupancy and its topology and also to identify activities [67], [68], [70], [55]. Our new results show how trajectory information can be more precisely employed, alone or in combination with other features for the extraction of activity patterns. Three application domains are currently being explored: 1) Monitoring of elderly people
at home; 2) Monitoring the ground activities at an airport dock-station (COFRIEND project\(^6\)); 3) Monitoring activities in subway/street surveillance systems.

**6.18.1. Monitoring of elderly people at home**

We propose a novel framework to understand daily activities in home-care applications; the framework is capable of discovering, modeling and recognizing long-term activities (e.g. “Cooking”, “Eating”) occurring in unstructured scenes (i.e. “an apartment”).

The framework links visual information (i.e., tracked objects) to the discovery and recognition of activities by constructing an intermediate layer of primitive events automatically.

The primitive events characterize the global spatial movements of a person in the scene (“in the kitchen”), and also the local movements of the person body parts (“opening the oven”). The primitive events are built from interesting regions, which are learned at multiple semantic resolutions (e.g. the “oven” is inside the “kitchen”). An example of the regions and possible activities for a single resolution is displayed in Fig. 22.

![Figure 22. Example of some learned regions and possible activities.](image)

A probabilistic model is learned to characterize each discovered activity. The modeled activities are automatically recognized in new unseen videos where a pop-up with a semantic description appears when an activity is detected. Examples of semantic labels are illustrated in Fig. 23 (a, b, c).

Recently we introduced 3D (MS. Kinect) information to the system. The preliminary results show an improvement superior to the 30% of the recognition quality. Also, the system can recognize activities in challenging situations as the lack of light. - See Fig. 2 (b) and Fig. 2 (c) -. The approach can be used to recognize most of the interesting activities in a home-care application and has been published in [43]. Other examples and applications are available online in [http://www-sop.inria.fr/pulsar/personnel/Guido.Pusiol/Home4/index.php](http://www-sop.inria.fr/pulsar/personnel/Guido.Pusiol/Home4/index.php).

**6.18.2. Monitoring the ground activities at an airport dock-station**

The COFRIEND project aims at creating a system for the recognition and interpretation of human activities and behaviours at an airport dock-station. Our contribution is a novel approach for discovering, in an unsupervised manner, the significant activities from observed videos. Spatial and temporal properties from detected mobile objects are modeled employing soft computing relations, that is, spatio-temporal relations graded with

Figure 23. Examples of the recognized activities in 3D and under different light conditions.

different strengths. Our system works off-line and is composed of three modules: The trajectory speed analysis module, The trajectory clustering module, and the activity analysis module. The first module is aimed at segmenting the trajectory into segments of fairly similar speed (tracklets). The second aims at obtaining behavioural displacement patterns indicating the origin and destination of mobile objects observed in the scene. We achieve this by clustering the mobile tracklets and also by discovering the topology of the scene. The latter module aims at extracting more complex patterns of activity, which include spatial information (coming from the trajectory analysis) and temporal information related to the interactions of mobiles observed in the scene, either between themselves or with contextual elements of the scene. A clustering algorithm based on the transitive closure calculation of the final relation allows finding spatio-temporal patterns of activity. An example of discovery is given in the figure below. This approach has been applied to a database containing near to 25 hours of recording of dock-station monitoring at the Toulouse airport. The discovered activities are: ‘GPU positioning’, ‘Handler deposits chocks’, ‘Frontal unloading operation’, ‘Frontal loading operation’, ‘Rear loading operation’, ‘Push back vehicle positioning’. An example of discovered activity (Frontal loading) is given on the figure 24. When comparing our results with explicit ground-truth given by a domain expert, we were able to identify the events in general with a temporal overlap of at least 50%. The comparison with a supervised method on the same data indicates that our approach is able to extract the interesting activities signalled in the ground-truth with a higher True Positive Rate (74% TPR for the supervised approach against 80% TPR with our unsupervised method). This work has been published in [42].

Figure 24. Example of an activity cluster obtained. The left panel presents the tracklets of the mobiles participating in the Frontal Loading activity. Filled circles indicate the beginning of a tracklet. Empty circles indicate the end of a tracklet. The right panel presents the start frame of the activity.

6.18.3. Monitoring activities in subway/street surveillance systems
In this work we have built a system to extract from video and in an unsupervised manner the main activities that can be observed from the a subway scene. We have setup a processing chain broadly working on three steps: The system starts in a first step by the unsupervised learning of the main activity areas of the scene. In a second step, mobile objects are then characterized in relation to the learned activity areas: either as ‘staying in a given activity zone’ or ‘transferring from an activity zone to another’ or a sequence of the previous two behaviours if the tracking persists long enough. In a third step we employ a high-level relational clustering algorithm to group mobiles according to their behaviours and discover other characteristics from mobile objects which are strongly correlated. We have applied this algorithm to two domains. First, monitoring two hours of activities in the hall entrance of an underground station and showed what are the most active areas of the scene and how rare/abnormal (going to low occupied activity zones) and frequent activities (e.g. buying tickets) are characterized. In the second application, monitoring one hour of a bus street lane, we were able again to learn the topology of the scene and separate normal from abnormal activities. When comparing with the available ground-truth for this application, we obtained a high recall measure (0.93) with an acceptable precision (0.65). This precision value is mostly due to the different levels of abstraction between the discovered activities and the ground-truth. The incremental learning procedure employed in this work is published in [52] while the full activity extraction approach was published in [41].

6.19. SUP Software Platform

Participants: Julien Gueytat, Leonardo Rocha, Daniel Zullo, François Brémond.

SUP is a software platform developed by PULSAR team, written in C and C++ for generating activity recognition systems. These systems should be able to perceive, analyze, interpret and understand a 3D dynamic scene observed through a network of sensors. These activity recognition systems are a combination of algorithms developed by members of Pulsar or state of the art computer vision libraries. The SUP dissemination is targeted for use in real-world applications requiring high-throughput.

SUP is made as a framework allowing several computer vision workflows to be implemented. Currently, the workflow is static for a given application but our goal is to make it dynamic. A given workflow is the composition of several plugins, each of them implementing an algorithmic step in the video processing chain (i.e. the segmentation of images, the classification of objects, etc.). The design of SUP allows to execute at run-time the selected plugins.

During 2011 several tasks have been accomplished:

- A stable packaged release is available
- 3D simulation from a scenario description
- Existing algorithms have been improved in performance and accuracy
- Kinect sensor has been added to the hardware supported

Several plugins are available:

- 2 plugins are wrappers on industrial implementations of video processing algorithms (made available by Keeneo). They allow a quick deployment of a video processing chain encompassing image acquisition, segmentation, blob construction, classification and short-term tracking. These algorithms are robust and efficient algorithms, but with the drawback that some algorithms can lack accuracy.
- Several implementations by the Pulsar team members which cover the following fields:
  1. Image acquisition from different types of image and camera video streaming.
  2. Segmentation removing the shadows.
  3. Two classifiers, one being based on postures and one on people detection.
  4. Four frame-to-frame trackers, using as algorithm:
1. a simple tracking by overlapping,
2. neural networks,
3. tracking of feature points,
4. tracking specialized for the tracking of people in a crowd.
5. Three scenario recognizers, one generic algorithm allowing expression of probabilities
   on the recognized events, the second one focusing on the recognition of events based on
   postures and the third one (see section Extendable Event Recognition algorithm: SED in
   this document) uses the complete ontology of the domain as a parameter (e.g. the definition
   of objects of interest, scenario models, etc.).
6. 3D animation generation, it generates a virtual 3D animation from information provided
   by different plugins of the processing chain together with 3D contextual environment.
7. 3D simulation from description, it generates a virtual 3D animation from information
   provided from a text file with the description of the scenario.

From a software engineering point-of-view, the goal is to obtain a flexible platform being dynamically
reconfigurable for the generated scene understanding systems to be autonomous and adaptable for handling
changing environment.

SUP relies on DTK, a generic platform developed by the DREAM service at INRIA Research Center Sophia-
Antipolis Méditerranée.

The purpose of DTK is to provide a software infrastructure allowing the generation of a new system by the
composition of plugins, each plugin being an algorithmic step of the whole processing chain. SUP is oriented to
help developers building activity recognition systems and describing their own scenarios dedicated to specific
applications. By relying on the DTK software infrastructure, the possibilities are:

- To simplify the exchanges of algorithms between different INRIA teams using the DTK.
- To use the facilities already provided by the DTK allowing to compose quickly existing plugins. Cur-
  rently a python interface is operational, and we plan to take advantage of the graphical composer to
  prototype quickly new work-flows, or reconfigure existing ones, for the experimentation conducted
  by the team.

In order to be confident on the results obtained with the SUP platform, an important effort is done to check:

- The correct behavior of the platform from a software engineering point of view, i.e. that the
  functionality of the SUP software is correctly provided, or is not broken by modifications.
- A qualitative evaluation tool (see ViSEvAI in this document) for the algorithms, which compares
  and assesses the results obtained with the algorithms to ground truth for several reference videos.

Both kinds of test are performed on a daily basis and on several hardware/software architectures.

6.20. Model-Driven Engineering and Video-surveillance

Participants: Sabine Moisan, Jean Paul Rigault, Sagar Sen, François Brémond.

In the framework of our research on model engineering techniques for video-surveillance systems, we have
focused this year on the runtime adaptation of such systems.

Video-surveillance systems are complex and exhibit high degrees of variability along several dimensions.
At the specification level, the number of possible applications and type of scenarios is large. On the
software architecture side, the number of components, their variations due to possible choices among different
algorithms, the number of tunable parameters... make the processing chain configuration rather challenging.
Moreover, the context of an application may change in real time, requiring dynamic reconfiguration of the
chain. This huge variability raises problems at design time (finding the configurations needed by the chain,
foreseeing the different possible contexts), at deployment time (selecting the initial configuration), and at run
time (switching configurations to react to context changes).
The first step was to formalize in an unified way all the necessary concerns —at the specification as well as at the component level— and their relations. To this end, we rely on Feature Models and (semi) automatic model transformations. Feature Models are widely used to represent systems with many possible variation points. Moreover they are liable to formal analysis (using propositional logic and satisfiability techniques) and thus lead to valid configurations, by construction. We have developed two feature models, one for the specification of the application (type of application, context of execution, expected quality of service, etc.) and one for the implementation representation (components and their assembly). Each model has its own internal constraints. Moreover, the two models are not independent: they are connected by cross model transformation rules that formalize the bridge between application requirements and component assemblies that realize them.

Second, we propose a framework to derive valid possible system configurations and to adapt running configurations to context changes. Users can select features describing their application in the specification model, through a simple graphic interface. The outcome is a sub-model of the specification model. Based on the cross model transformation rules, our framework automatically transforms this sub-model into a sub-model of the component model. The latter represents all possible component configurations of the target video-surveillance system that satisfy the specifications. Both sub-models will be kept throughout the system life. They are used while the system is running to adjust its configuration in response to context changes.

To achieve this dynamic adaptation, our framework sets up three collaborating modules as shown in Figure 25:

- the Run Time Component Manager (RTCM) captures low level events manifesting context changes (e.g., lighting changes); it forwards them to the Configuration Adapter which returns a

\[\text{Figure 25. Run Time Adaptation Architecture}\]

\(\text{\footnotesize \begin{itemize} \item \textbf{RTCM} (e.g., OSGI) \item \textbf{CA} (e.g., heuristics) \item \textbf{MM} (e.g., BDD based) \end{itemize}}\]

\[\text{Video surveillance processing chain} \]

\[\text{adaptation orders (insert, substitute, remove)} \]

\[\text{events} \]

\[\text{Component sub-model} \]

\[\text{Feature (de)selection} \]

\[\text{Reasoning Engine (e.g., BDD based)} \]

\[\text{Transformation rules} \]

\[\text{Variability models} \]

\[\text{\footnotesize \begin{itemize} \item \text{RTCM} \item \text{CA} \item \text{MM} \end{itemize}}\]

\[\text{\footnotesize \begin{itemize} \item \text{the Run Time Component Manager (RTCM) captures low level events manifesting context changes (e.g., lighting changes); it forwards them to the Configuration Adapter which returns a}
\end{itemize}}\]

\[\text{\footnotesize \begin{itemize} \item \text{Here the “features” correspond to selectable concepts of the systems; they can be at any abstraction level (a feature may correspond to an specification entity such as “Intrusion detection” or to a more concrete element such as “High frame rate”). The features are organized along a tree, with logical selection relations (optional, mandatory features, exclusive choices...) and some constraints that restrict the valid combinations of features (i.e., configurations)\end{itemize}}\]
new component configuration; the RTCM is then responsible for applying this configuration, that is
to tune, add, remove, or replace components, and possibly to change the workflow itself.

- the Configuration Adapter (CA) receives change events from the RTCM, translates them into the
  feature formalism, and forwards the result to the Model Manager; in return, it obtains a sub-model of
  component configurations compatible with the change; this sub-model is a compact representation
  of a set of valid configurations and the CA is responsible to select one and to instruct the RTCM to
  apply it; this selection uses some heuristics, possibly based on a cost function such as minimizing
  the number of component changes in the processing chain or maximizing the quality of service (e.g.,
  accuracy, responsiveness).

- the Model Manager (MM) manages the representation of the two specialized Feature Models
  corresponding to the specification and possible component assemblies of the current application
  together with their constraints; its role is to enforce configuration validity. It is also responsible
  of the set of rules relating run time events and (de)selection of features in both models. From the
  CA, the Model Manager receives information about incoming events; it uses the rules to select or
deselect the corresponding features; it then applies constraints, rules, and model transformations to
  infer a component sub-model that represents a subset of valid component configurations and that it
  returns to the CA.

This year, we have tested our approach on simple applications using well-known libraries (OpenCV) on
different scenarios. At the moment, 77 features and $10^5$ configurations are present in the specification model
while 51 features and $10^6$ configurations are present in the component model. Once the video surveillance
designer has selected the features required by an application, before deployment, the average number of
features to consider at runtime in the component model is less than $10^4$. The configuration spaces is reduced
by several orders of magnitude and enables the use of the other tools in the end-to-end engineering process,
whereas it would not have been possible without. Our experiments show the feasibility of such an approach
with a limited performance overhead (if any) compared to traditional run time control where ad hoc adaptation
code is hardwired and does not rely on the run time availability of an abstract representation of the application
and its context evolution [29]. The next step will be to test our approach on our SUP platform and to study
intelligent configuration selection heuristics.

6.21. The Girgit Software

Participants: Leonardo Rocha, Sabine Moisan, Jean-Paul Rigault, Sagar Sen.

Girgit is a Python based framework to build context-aware self-adaptive software systems.

Girgit is a simple and small [1] framework that allows dynamic reconfiguration of data processing chains and
accepts any set of components for the configuration as long as they have the corresponding Python wrapper.

The basic Idea of Girgit is to provide a platform to be able to make dynamic adaptive systems. It provides
a dynamic adaptive engine that can deal with event/action pairs called rules and also provides an application
programming interface to be able to use it.

As the system is designed to be interactive, any user can interact with it, be a human operator or a reasoning
engine. Rules can be pre-loaded with the configuration, and components can launch events, this allows pre-
configured rules to be loaded at launch time of the Girgit.

Girgit has been used to evaluate the performance on real time video applications and show the architecture.
Three publications [47], [45], [46] where based on the framework.

6.22. Scenario Analysis Module

Participants: Sabine Moisan, Annie Ressouche, Jean Paul Rigault.
To generate activity recognition systems we supply a scenario analysis module (SAM) to express and recognize complex events from primitive events generated by SUP or others sensors. In this framework, this year we focus on recognition algorithm improvement in order to face the problem of large number of scenario instances recognition.

The purpose of this research axis is to offer a generic tool to express and recognize activities. Genericity means that the tool should accommodate any kind of activities and be easily specialized for a particular framework. In practice, we propose a concrete language to specify activities in the form of a set of scenarios with temporal constraints between scenarios. This language allows domain experts to describe their own scenario models. To recognize instances of these models, we consider the activity descriptions as synchronous reactive systems [69] and we apply general modeling methods [62] to express scenario behaviors. This approach facilitates scenario validation and allows us to generate a recognizer for each scenario model. The SAM module thus provides users with (1) a simulation tool to test scenario behaviors; (2) a generator of a recognizer for each scenario model; (3) an exhaustive verification of safety properties relying on model checking techniques our approach allows. The latter offers also the possibility to define safety properties to prove as “observers” [63] expressed in the scenario language.

Last year we have completed SAM in order to address the life cycle of scenario instances. For a given scenario model there may exist several (possibly many) instances at different evolution states. These instances are created and deleted dynamically, according to the input event flow. The challenge is to manage the creation/destruction of this large set of scenario instances efficiently (in time and space), to dispatch events to expecting instances, and to make them evolve independently. This year, to face this challenge we first replace some operators of the language, by others having a more strict semantics. For instance, we replace the before operator whose semantics allowed that events can meet, by two operators, a strict before and a meet. Hence, the number of events a scenario instance reacts to decreases. Second, we now generate within the recognition engine, the expected events of the next step. This avoids to run the engine automatically with events that are not relevant for the recognition process.

Presently, we still rely on the existing synchronous language (Lustre [62]) to express the equational semantics of scenario models and to generate recognizers because this language offers simulation and verification means. But, to improve efficiency, we plan to build our own compiler and to generate recognizers directly from the Boolean equation systems modeling scenario models. This implies that we must supply our own simulation tool and that we interface with a model checking tool as NuSMV [57].

Now the challenge is to take into account some uncertainty on the primitive events due to input sensor errors. In the family of synchronous languages, the Lutin language 8 could be able to automate the generation of realistic input sequences of events, taking into account probabilistic distributions over primitive events. In other words, it could generate a set of input events for which a set a constraints can be verified. In complement, it offers also means to compute the real values verifying these constraints. Thus, we think to rely on Lutin to express uncertainty on primitive events and get input events to feed scenario recognition engines.

### 6.23. Multiple Services for Device Adaptive Platform for Scenario Recognition

**Participants:** Annie Ressouche, Jean-Yves Tigli.

The aim of this research axis is to federate the inherent constraints of an activity recognition platform like SUP (see section 6.19) with a service oriented middleware approach dealing with dynamic evolutions of system infrastructure. The Rainbow team (Nice-Sophia Antipolis University) proposes a component-based adaptive middleware (WComp [76], [75], [64]) to dynamically adapt and recompose assemblies of components. These operations must obey the "usage contract" of components. The existing approaches don’t really ensure that this usage contract is not violated during application design. Only a formal analysis of the component behavior models associated with a well sound modeling of composition operation may guarantee the respect of the usage contract.

---

8 [http://www-verimag.imag.fr/Lutin.html](http://www-verimag.imag.fr/Lutin.html)
The approach we adopted introduces in a main assembly, a synchronous component for each sub assembly connected with a critical component. This additional component implements a behavioral model of the critical component and model checking techniques apply to verify safety properties concerning this critical component. Thus, we consider that the critical component is validated.

When a critical component has multiple synchronous monitors corresponding to several concern managements in the application, we want to build an only synchronous model component which agrees with all these primitive synchronous monitors. To specify how output events sent by different synchronous monitors and connected to a critical component, we introduce a sound (with respect to our mathematical formalism) operation of composition under constraints of synchronous models (see figure 26). We proved that this operation preserves already separately verified properties of synchronous components. This operation is an answer to the multiple access to critical components. Actually, we supply a graphical interface to design both critical component behaviors and properties as observers in the synchronous language Lustre [62]. Then the validation of properties and the creation of the validated synchronous component is automatic [44], [53].

This year we focus on the main challenge of this approach which is to deal with the possibly very large number of constraints a user must specify. Indeed, each synchronous monitor has to tell how it combines with others, then we get a combinatorial number of constraints with respect to the number of synchronous monitors and inputs of the critical component. To be adaptive with efficiency, we must face this problem. We first introduced some default rules to avoid the user to express a large number of constraints. We also studied how Abstract Interpretation technique can help us to reduce this complexity. This approach works if we forbid some “non monotonic” constraints, but this is a strong limitation. Thus, it is still a challenge for us. This drawback is a popular challenge in adaptive middleware and some results exist relying on controller synthesis methods. We are not in the exact framework where these techniques apply, but we plan to study if we can rely on some extension of these techniques.

On another hand, we also want to complement our preservation result in studying how the proof of a global property can be decomposed into the proof of local ones. In general, this decomposition (known as assume-
guarantee paradigm) is difficult to apply but there is no communication between synchronous monitors and so the decomposition could be tractable. Moreover, some works address this problem and we can rely on them.

6.24. The Clem Workflow

Participants: Annie Ressouche, Daniel Gaffé.

This research axis concerns the theoretical study of a synchronous language LE with modular compilation and the development of a toolkit (see figure 6) around the language to design, simulate, verify and generate code for programs. The novelty of the approach is the ability to manage both modularity and causality. Indeed, only few approaches consider a modular compilation because there is a deep incompatibility between causality and modularity. Thus, relying on semantics to compile a language ensures a modular approach but requires to complete the compilation process with a global causality checking. To tackle this problem, we introduced a new way to check causality from already checked sub programs and the modular approach we infer. The equational semantics compute a Boolean equation system and we ensure both modularity and causality in computing all the partial orders valid for a system and we define a way to merge two partial orders. The algorithm which computes partial orders rely on the computation of two dependency graphs: the upstream (downstream) dependency graph computes the dependencies of each variable of the system starting from the input (output) variables. This way of compiling is the cornerstone of our approach. We defined three different approaches to compute the partial orders valid for an equation system:

1. apply PERT method : inputs (resp. outputs) have date 0 and recursively increase of dates for each vertex in the upstream (resp downstream) dependencies graph;

2. apply graph theory:
   - compute the adjacency matrix U of upstream (resp. downstream) dependencies graph;
   - the length of the maximal path from a variable v to system inputs is characterized by the maximal k such that $U^k[v, i] \neq 0$ for all inputs i.

3. apply fix point theory: the vector of earliest (resp. latest) dates can be computed as the least fix point of a monotonic increasing function.

The fix point characterization helps us to prove that the merge algorithm is correct (i.e we get the same partial orders using the merge algorithm on two previously sorted equation systems or when sorting the union of the two equation systems considered).

To be modular, we defined a technique to compose two already sorted equation systems: first, we memorize the two dependency graphs of equation systems. Second, we define two merge algorithms relying on two different techniques:

1. propagation of common variables dates adjustment;

2. fix point characterization starting with the vectors of already computed dates and considering only the variables in the dependencies (upstream and downstream) of common variables

This year we began the implementation of a separated compilation of LE programs, according to these theoretical results. We define a new intermediate format (lea) to record partially compiled module, i.e module whose Boolean equation systems may be composed of non defined variables (we called them abstract). Then we are implementing a refinement operation which replaces these abstract variables by their definition and performs adjustment of the dates. According to our theoretical results, we know that the resulting sorting is the same as with a global approach. After the termination of this separated compilation of LE programs, the challenge will be to use Clem to design a large application in the domain of smart cards. The application needs more than forty LE automata in parallel and the compiled code will have more than 500 registers and thousands variables. Only a separated compilation will work.

The Clem toolkit is completely described in [28]
5. New Results

5.1. Plausible Image Rendering

5.1.1. Filtering Solid Gabor Noise

Participants: Ares Lagae, George Drettakis.

Solid noise is a fundamental tool in computer graphics. Surprisingly, no existing noise function supports both high-quality anti-aliasing and continuity across sharp edges. Existing noise functions either introduce discontinuities of the solid noise at sharp edges, which is the case for wavelet noise and Gabor noise, or result in detail loss when anti-aliased, which is the case for Perlin noise and wavelet noise. In this project, we therefore present a new noise function that preserves continuity over sharp edges and supports high-quality anti-aliasing. We show that a slicing approach is required to preserve continuity across sharp edges, and we present a new noise function that supports anisotropic filtering of sliced solid noise. This is made possible by individually filtering the slices of Gabor kernels, which requires the proper treatment of phase. This in turn leads to the introduction of the phase-augmented Gabor kernel and random-phase Gabor noise, our new noise function. We demonstrate that our new noise function supports both high-quality anti-aliasing and continuity across sharp edges, as well as anisotropy. Fig. 3 shows several solid procedural textures generated with our new random-phase solid Gabor noise.

This work was presented at ACM SIGGRAPH 2011 in Vancouver and published in ACM Transactions on Graphics [18] (SIGGRAPH paper).

5.1.2. Image-Guided Weathering for Flow Phenomena

Participants: Carles Bosch, Pierre-Yves Laffont, George Drettakis.

The simulation of weathered appearance is essential in the realistic modeling of urban environments. With digital photography and Internet image collections, visual examples of weathering effects are readily available. These images, however, mix the appearance of the weathering phenomena with the specific local context. In [12], we have introduced a new methodology to estimate parameters of a phenomenological weathering simulation from existing imagery, in a form that allows new target-specific weathering effects to be produced. In addition to driving the simulation from images, we complement the visual result with details and colors extracted from the images. This methodology has been illustrated using flow stains as a representative case, demonstrating how a rich collection of flow patterns can be generated from a small set of exemplars. In Fig. 4, we show the major components required for this approach.
Figure 4. From left to right: given a photograph of a stain effect, we extract its stain degree map, we estimate the information required for the simulation (color, simulation parameters, and high frequency details), and run simulations on new scenes using this information, resulting in a wide variety of similar-looking stains.

5.1.3. Relighting Photographs of Tree Canopies

Participants: Marcio Cabral, George Drettakis.

We present an image-based approach to relighting photographs of tree canopies. Our goal is to minimize capture overhead; thus the only input required is a set of photographs of the tree taken at a time of day, while allowing relighting at any other time. We first analyze lighting in a tree canopy, both theoretically and using simulations. From this analysis, we observe that tree canopy lighting is similar to volumetric illumination. We assume a single-scattering volumetric lighting model for tree canopies, and diffuse leaf reflectance. To validate our assumptions, we apply our method on several synthetic renderings of tree models, for which we are all able to compute all quantities involved.

Our method first creates a volumetric representation of the tree using 10-12 images taken at a single time of day and use a single-scattering participating media lighting model - these photos are taken from different viewpoints, around the tree. An analytical sun and sky illumination model, namely the Preetham model, provides consistent representation of lighting for the captured input and unknown target times.

We relight the input image by applying a ratio of the target and input time lighting representations. We compute this representation efficiently by simultaneously coding transmittance from the sky and to the eye in spherical harmonics. We validate our method by relighting images of synthetic trees and comparing to path-traced solutions. We also present results for photographs, validating with time-lapse ground truth sequences. An example is shown in Fig. 5. This work was published in the IEEE Transactions on Computer Graphics and Visualization [15], and was in collaboration with the past members of REVES, N. Bonneel and S. Lefebvre.

5.1.4. Silhouette-aware Warping for Image-based Rendering

Participants: Gaurav Chaurasia, George Drettakis.
Mulberry tree Top row: input images and 4 target images with corresponding times of day. Middle row: 4 resulting relit images using our approach. Bottom row: $E_{in}$ and the four $E_{targ}$ images.

Figure 5.
Figure 6. Top: Overview of our approach starting from input images to rendered novel views along with novel view generated using current state of the art IBR. Bottom: Novel views rendered using our approach.
We have presented a novel solution for image-based rendering (IBR) of urban scenes based on image warping. IBR techniques allow capture and display of 3D environments using photographs. Modern IBR pipelines reconstruct proxy geometry using multi-view stereo, reproject the photographs onto the proxy and blend them to create novel views. The success of these methods depends on accurate 3D proxies, which are difficult to obtain for complex objects such as trees and cars. Large number of input images do not improve reconstruction proportionally; surface extraction is challenging even from dense range scans for scenes containing such objects. Our approach does not depend on dense accurate geometric reconstruction; instead we compensate for sparse 3D information by variational warping. In particular, we formulate silhouette-aware warps that preserve salient depth discontinuities. This improves the rendering of difficult foreground objects, even when deviating from view interpolation. We use a semi-automatic step to identify depth discontinuities and extract a sparse set of depth constraints used to guide the warp. On the technical side, our formulation breaks new ground by demonstrating how to incorporate discontinuities in variational warps. Our framework is lightweight and results in good quality IBR for previously challenging environments as shown in figure. 6 .

**Applications.** Robust image-based rendering can be used to generate photo-realistic visual content easily which can be very useful for virtual reality applications. Commercial products like Google StreetView and Microsoft PhotoSynth use rudimentary image-based rendering for large scale visualization of cities. Our work advances the state of the art by treating the hardest class of scenes.

This work is a collaboration with Prof. Dr. Olga Sorkine (ETH Zurich). The work was published in the special issue of the journal Computer Graphics Forum [16], presented at the Eurographics Symposium on Rendering 2011 in Prague, Czech Republic.

### 5.1.5. Real Time Rough Refractions

**Participant:** Adrien Bousseau.

We have proposed an algorithm to render objects made of transparent materials with rough surfaces in real-time, under environment lighting. Rough surfaces such as frosted and misted glass cause wide scattering as light enters and exits objects, which significantly complicates the rendering of such materials. We introduced two contributions to approximate the successive scattering events at interfaces, due to rough refraction: First, an approximation of the Bidirectional Transmittance Distribution Function (BTDF), using spherical Gaussians, suitable for real-time estimation of environment lighting using pre-convolution; second, a combination of cone tracing and macro-geometry filtering to efficiently integrate the scattered rays at the exiting interface of the object. Our method produces plausible results in real-time, as demonstrated by comparison against stochastic ray-tracing (Figure 7).
This work is a collaboration with Charles De Rousiers, Kartic Subr, Nicolas Holzschuch (ARTIS / INRIA Rhône-Alpes) and Ravi Ramamoorthi (UC Berkeley) as part of our CRISP associate team with UC Berkeley. The work was presented at the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games 2011 and won the best paper award.

### 5.1.6. Single view intrinsic images

**Participants:** Pierre-Yves Laffont, Adrien Bousseau, George Drettakis.

We introduced a new algorithm for decomposing photographs of outdoor scenes into *intrinsic images*, i.e. independent layers for illumination and reflectance (material color).

Extracting intrinsic images from photographs is a hard problem, typically solved with image-driven methods using numerous user indications. Recent methods in computer vision allow easy acquisition of medium-quality geometric information about a scene using multiple photographs from different views. We developed a new algorithm which allows us to exploit this noisy and somewhat unreliable information to automate and improve image-driven propagation algorithms to deduce intrinsic images. In particular, we develop a new approach to estimate cast shadow regions in the image, by refining an initial estimate available from reconstructed geometric information. We use a voting algorithm in color space which robustly identifies reflectance values for sparse reconstructed 3D points, allowing us to accurately determine visibility to the sun at these points.

---

Figure 8. Starting from multiple views of the scene (a), our method automatically decomposes photographs into three intrinsic components — the reflectance (b), the illumination due to sun (c) and the illumination due to sky (d). Each intrinsic component can then be manipulated independently for advanced image editing applications.
This information is propagated to the remaining pixels. In a final step we adapt appropriate image propagation algorithms, by replacing manual user indications by data inferred from the geometry-driven shadow and reflectance estimation. This allows us to automatically extract intrinsic images from multiple viewpoints, thus allowing many types of manipulation in images.

As illustrated on Figure 8, our method can extract reflectance at each pixel of the input photographs. The decomposition also yields separate sun illumination and sky illumination components, enabling easier manipulation of shadows and illuminant colors in image editing software.

This work won the best paper award at “les journées de l’AFIG (Association Francophone d’Informatique Graphique) 2011”, and will appear in the French journal REFIG [21].

5.1.7. Single view intrinsic images with indirect illumination

Participants: Pierre-Yves Laffont, Adrien Bousseau, George Drettakis.

Building on top of our work published in [21], we extended our intrinsic image decomposition method to handle indirect illumination, in addition to sun and sky illumination. We introduced a new algorithm to reliably identify points in shadow based on a parameterization of the reflectance with respect to sun visibility, and proposed to separate illumination components using cascaded image-based propagation algorithm. We also demonstrated direct applications of our results for the end user in widespread image editing software. This rich intrinsic image decomposition method has been submitted for publication.

5.1.8. Multi-lighting multi-view intrinsic images

Participants: Pierre-Yves Laffont, Adrien Bousseau, George Drettakis.

We compute intrinsic image decompositions using several images of the same scene under different viewpoints and lighting conditions. Such images can be easily gathered for famous landmarks, using photo sharing websites such as Flickr. Our method leverages the heterogeneity of photo collections (multiple viewpoints and lighting conditions) to guide the intrinsic image separation process. With this automatic decomposition, we aim to facilitate many image editing tasks and improve the quality of image-based rendering from photo collections.

This work is in collaboration with Frédo Durand (associate professor at MIT) and Sylvain Paris (researcher at Adobe).

5.1.9. Warping superpixels

Participants: Gaurav Chaurasia, Sylvain Duchêne, George Drettakis.

We are working on developing a novel representation of multi-view scenes in the form of superpixels. Such a representation segments the scene into semantically meaningful regions called superpixels which admit variational warps [16], thereby leveraging the power of shape-preserving warps while providing automatic silhouette and occlusion handling. This entails several contributions namely incorporating geometric priors in superpixel extraction, generating consistent warping constraints and a novel rendering pipeline that assembles warped superpixels into the novel view maintaining spatio-temporal coherence.

This work is a collaboration with Prof. Dr. Olga Sorkine (ETH Zurich).

5.1.10. State-of-the-art Report on Temporal Coherence for Stylized Animations

Participant: Adrien Bousseau.
Non-photorealistic rendering (NPR) algorithms allow the creation of images in a variety of styles, ranging from line drawing and pen-and-ink to oil painting and watercolor. These algorithms provide greater flexibility, control and automation over traditional drawing and painting. The main challenge of computer generated stylized animations is to reproduce the look of traditional drawings and paintings while minimizing distracting flickering and sliding artifacts present in hand-drawn animations. These goals are inherently conflicting and any attempt to address the temporal coherence of stylized animations is a trade-off. This survey is motivated by the growing number of methods proposed in recent years and the need for a comprehensive analysis of the trade-offs they propose. We formalized the problem of temporal coherence in terms of goals and compared existing methods accordingly (Figure 9). The goal of this report is to help uninformed readers to choose the method that best suits their needs, as well as motivate further research to address the limitations of existing methods.

This work is a collaboration with Pierre Bénard and Joëlle Thollot (ARTIS / INRIA Rhône-Alpes) and was published in the journal Computer Graphics Forum [14].

5.1.11. Improving Gabor Noise

**Participant:** Ares Lagae.

We have recently proposed a new procedural noise function, Gabor noise, which offers a combination of properties not found in existing noise functions. In this project, we present three significant improvements to Gabor noise: (1) an isotropic kernel for Gabor noise, which speeds up isotropic Gabor noise with a factor of roughly two, (2) an error analysis of Gabor noise, which relates the kernel truncation radius to the relative error of the noise, and (3) spatially varying Gabor noise, which enables spatial variation of all noise parameters. These improvements make Gabor noise an even more attractive alternative for existing noise functions. Fig. 3 shows a procedural textures generated with our new improved Gabor noise. This work was published in IEEE Transactions on Computer Graphics and Visualization [19].

5.2. Visual Perception and Audio Rendering
5.2.1. Perception of Visual Artifacts in Image-Based Rendering of Façades

Participants: Peter Vangorp, Gaurav Chaurasia, Pierre-Yves Laffont, George Drettakis.

Image-based rendering (IBR) techniques allow users to create interactive 3D visualizations of scenes by taking a few snapshots (Figure 11 (left)). However, despite substantial progress in the field, the main barrier to better quality and more efficient IBR visualizations are several types of common, visually objectionable artifacts. These occur when scene geometry is approximate or viewpoints differ from the original shots, leading to parallax distortions (Figure 11 (mid)), blurring, ghosting (Figure 11 (right)) and popping errors that detract from the appearance of the scene. We argue that a better understanding of the causes and perceptual impact of these artifacts is the key to improving IBR methods.

We present a series of psychophysical experiments in which we systematically map out the perception of artifacts in IBR visualizations of façades as a function of the most common causes. We separate artifacts into different classes and measure how they impact visual appearance as a function of the number of images available, the geometry of the scene and the viewpoint. The results reveal a number of counter-intuitive effects in the perception of artifacts.

We summarize our results in terms of the following practical guidelines for improving existing and future IBR techniques:
• When the total number of available images is small, e.g., because of storage limitations, it is preferable to use a sudden transition with its associated popping artifact rather than a gradual blending transition with its associated ghosting artifact.

• Interestingly, the depth range of the façade does not affect the perceived parallax distortions, even though it clearly does affect the objective parallax distortions. Only the intended output viewing angle should be taken into account when capturing images.

• For Google Street View™-like visualizations, a shorter cross-fading transition would improve the perceived quality.

This work is a collaboration with Roland W. Fleming (Justus-Liebig-University Giessen, Germany). The work was published in the special issue of the journal Computer Graphics Forum [20] and presented at the Eurographics Symposium on Rendering 2011.

5.2.2. Perception of Slanted, Textured Façades

Participants: Peter Vangorp, Adrien Bousseau, Gaurav Chaurasia, George Drettakis.

In large-scale urban visualizations, buildings are often geometrically represented by simple boxes textured with images of the façades. Any depth variations in the façade, such as balconies, are perceived to have distorted angles when the viewer is not at the capture camera position. The retinal hypothesis provides the most likely prediction of the magnitude of the perceived distortion. We conduct psychophysical experiments to measure the perceived distortion, thereby validating the retinal hypothesis, and to measure the threshold for detecting any distortion. The result is a prediction of the valid range of viewer motion for a given capture.

This work is a collaboration with Martin S. Banks (UC Berkeley).

5.2.3. Binocular and Dynamic Cues to Glossiness

Participants: Peter Vangorp, George Drettakis.

Recent advances in display technology have made it possible to present high quality stereoscopic imagery with accurate head tracking. This improves not only depth perception but also affects the perception of glossy materials. Previous work has shown that these conditions can increase the perceived gloss by a small amount. We conduct psychophysical experiments to measure this effect quantitatively.

This work is a collaboration with Roland W. Fleming (Justus-Liebig-University Giessen, Germany) and Martin S. Banks (UC Berkeley).

5.2.4. Sound Particles

Participants: Charles Verron, George Drettakis.

This research deals with a sound synthesizer dedicated to particle-based environmental effects, and intended to be used in interactive virtual environments. The synthesis engine is based on five physically-inspired basic elements (sound atoms) that can be parameterized and stochastically distributed in time and space. Physically-inspired controls simultaneously drive graphics particle models (e.g., distribution of particles, average particles velocity etc.) and sound parameters (e.g., distribution of sound atoms, spectral modifications etc.). The simultaneous audio/graphics controls result in an intricate interaction between the two modalities that enhances the naturalness of the scene. The approach is currently illustrated with three environmental phenomena: fire, wind, and rain.

5.2.5. Sound Synthesis for Crowds

Participants: Charles Verron, George Drettakis.

We are currently investigating new methods for synthesis of crowd sounds in virtual environments. Crowd sounds are constituted of many overlapping voices spatialized at different positions in the environment. A novel level of detail for crowd sounds is desirable: the cost of spatializing many individual voice sounds can be replaced by an efficient babble noise synthesis model. Furthermore, high-level control should be proposed to modify the crowd sounds by semantic parameters, related to the crowd emotional state (e.g., calm, angry...). This research should result in a new real-time crowd sound synthesizer with semantic controls for virtual environments.
5.3. Interaction and Design for Audiovisual Virtual Environments

5.3.1. Lighting Design for Material Depiction

Participants: Adrien Bousseau, Emmanuelle Chapoulie.

Shading, reflections and refractions are important visual features for understanding the shapes and materials in an image. While well designed lighting configurations can enhance these features and facilitate image perception (Figure 12 b), poor lighting design can lead to misinterpretation of image content (Figure 12 a).

![Figure 12](image.png)

*Figure 12. Our method (a) automatically optimizes the lighting to enhance material-specific visual features. The lighting reveals the thin and thick parts of the subsurface scattering wax candle, it accentuates the Fresnel reflections along the side of the porcelain vase and it adds strong specular highlights to emphasize the shiny chrome metal of the sculpture. Poorly designed lighting (b) diminishes these characteristic visual features of the materials. The candle appears more like solid plastic, the vase appears to be made of diffuse clay and the sculpture no longer looks like it is made of chrome.*

We have presented an automated system for optimizing and synthesizing environment maps that enhance the appearance of materials in a scene. We first identified a set of lighting design principles for material depiction. Each principle specifies the distinctive visual features of a material and describes how environment maps can emphasize those features. We then proposed a general optimization framework to solve for the environment map that best fulfill the design principles. Finally we described two techniques for transforming existing photographic environment maps to better emphasize materials. Our approach generates environment maps that enhance the depiction of a variety of materials including glass, metal, plastic, marble and velvet.

This work is a collaboration with Ravi Ramamoorthi and Maneesh Agrawala (UC Berkeley) as part of our CRISP associate team with UC Berkeley. The work was published in the special issue of the journal Computer Graphics Forum, presented at the Eurographics Symposium on Rendering 2011.

5.3.2. A Multimode Immersive Conceptual Design System for Architectural Modeling and Lighting

Participants: Marcio Cabral, Peter Vangorp, Gaurav Chaurasia, Emmanuelle Chapoulie, Martin Hachet, George Drettakis.

We developed a system which allows simple architectural design in immersive environments. The user is able to define the initial conceptual design of the model and can take into account the effects of daylight. Our system allows the manipulation of simple elements such as windows, doors and rooms while the overall model is automatically adjusted to the manipulation. The system runs on a four-sided stereoscopic, head-tracked immersive display. We also provide simple lighting design capabilities, with an abstract representation of sunlight and its effects when shining through a window. Our system provides three different modes.
of interaction: a miniature-model table mode, a fullscale immersive mode and a combination of table and immersive which we call mixed mode (see Figure 13). Our goal is to study direct manipulation for basic 3D modeling in an immersive setting, in the context of conceptual or initial design for architecture. We performed an initial pilot user test to evaluate the relative merits of each mode for a set of basic tasks such as resizing and moving windows or walls, and a basic light-matching task. The study indicates that users appreciated the immersive nature of the system, and found interaction to be natural and pleasant. In addition, the results indicate that the mean performance times seem quite similar in the different modes, opening up the possibility for their combined usage for effective immersive modeling systems for novice users.

Figure 13. (a) Table mode, (b) Immersive mode, (c) Mixed mode

5.3.3. Walking in a Cube: Novel Metaphors for Safely Navigating Large Virtual Environments in Restricted Real Workspaces
Participants: Peter Vangorp, Emmanuelle Chapoulie, George Drettakis.

Immersive spaces such as 4-sided displays with stereo viewing and high-quality tracking provide a very engaging and realistic virtual experience. However, walking is inherently limited by the restricted physical space, both due to the screens (limited translation) and the missing back screen (limited rotation). We propose three novel navigation techniques that have three concurrent goals: keep the user safe from reaching the translational and rotational boundaries; increase the amount of real walking; and finally, provide a more enjoyable and ecological interaction paradigm compared to traditional controller-based approaches. We notably introduce the “Virtual Companion”, which uses a small bird to guide the user through VEs larger than the physical space. We evaluate the three new techniques through a user study with pointing and path following tasks. The study provides insight into the relative strengths of each new technique for the three aforementioned goals.

This work is a collaboration with Gabriel Cirio, Maud Marchal, and Anatole Lécuyer (VR4I / INRIA Rennes) in the context of ARC NIEVE (Section 6.2.1) and has been accepted for publication [17].

5.3.4. Inferring Normals Over Design Sketches
Participant: Adrien Bousseau.

We are currently working on a sketch-based tool to infer normals over a 2D drawing. Our tool should allow users to apply realistic and non-photorealistic shading over the drawing, with applications in product design. This work is a collaboration with Alla Sheffer (University of British Columbia), Cloud Shao and Karan Singh (University of Toronto).
5.3.5. Using natural gestures into virtual reality immersive space

Participants: Emmanuelle Chapoulie, George Drettakis.

We are studying the use of gestures which are as natural as possible in a context of virtual reality environments. We define a scenario which is a sequence of tasks (hiding, finding, pushing, pulling, grabbing, picking up, putting down objects) that the users will perform with hands and with wands, in order to evaluate the usability of our interaction approach. Each task will be used to evaluate a specific criterion.
6. New Results

6.1. Coordination Parsing

Participants: Bruno Guillaume, Guy Perrier.

In the development of the French grammar, FRIGRAM, Joseph Le Roux and Guy Perrier have tackled the difficult problem of modelling and parsing coordination [39]. They have enriched FRIGRAM with a module expressing different syntactic constructions with coordination. An important drawback of this approach is the number of elementary constructions that have to introduced to obtain a reasonable coverage of the phenomenon.

In the continuation of his Master thesis, Valmi Dufour-Lussier with Bruno Guillaume and Guy Perrier worked on a different approach. They propose to process coordination at the parsing level as a linguistic performance issue, outside the grammar, rather than as a matter of competence [15]. They apply a specific algorithm to combine coordinated syntactic structures that were partially parsed using a coordination-less grammar, resulting in a directed acyclic parse graph in which constituent sharing appears sharply. They have experimented the algorithm within the framework of Tree-Adjoining Grammars (although it can be adapted to other formalisms) on a small subset of the Penn Treebank 2. They have shown that it is able to handle many types of coordinate constructions, including left and right node raising, argument clusters, and verb gapping.

6.2. Graph Rewriting

Participants: Bruno Guillaume, Mathieu Morey, Guy Perrier.

Guillaume Bonfante (from CARTE team), Bruno Guillaume, Mathieu Morey and Guy Perrier have improved their graph rewriting calculus, experimenting it in two directions. Taking an asynchronous perspective on the syntax-semantics interface, they have designed a modular graph rewriting system to produce underspecified semantic representations from a syntactic dependency graph [14]. They experimentally validated this approach on a set of sentences extracted from the French Treebank annotated with syntactic dependencies [27]. The results open the way for the production of underspecified semantic dependency structures from corpora annotated with syntactic dependencies and, more generally, for a broader use of modular rewriting systems for computational linguistics.

In a second application, they show how to enrich a syntactic dependency annotation of the French Treebank, using graph rewriting, in order to compute its semantic representation [13]. The rewriting system is composed of grammatical and lexical rules structured in modules. The lexical rules use a control information extracted from Dicovalence, a lexicon of French verbs 3.

6.3. ACG Type System

Participants: Philippe de Groote, Sylvain Pogodalla, Florent Pompigne.

In order to extend the flexibility and the expressiveness of the the ACG framework, [57] proposed a type-system extension. However, the formal properties of the system have to be proved. In his PhD work, Florent Pompigne is proposing alternate η-rules and commutative conversions in order to get the desirable properties. This work, currently in progress, relates to former proposals for a linear calculus with dependent types [28] and a calculus for extensionality with variants [40].

2 http://www.cis.upenn.edu/~treebank/
3 http://bach.arts.kuleuven.be/dicovalence/
6.4. Logic and Grammars

Participant: Maxime Amblard.

Maxime Amblard has presented an extension of Minimalist Categorial Grammars (MCG) to encode Chomsky’s phases in [11]. These grammars are based on Partially Commutative Logic (PCL) and encode properties of Minimalist Grammars (MG) of Stabler. The first implementation of MCG were using both non-commutative properties (to respect the linear word order in an utterance) and commutative ones (to model features of different constituents). Here, we propose to adding Chomsky’s phases with the non-commutative tensor product of the logic.

6.5. Discourse dynamics

Participants: Maxime Amblard, Sai Qian.

Sai Qian and Maxime Amblard has presented a framework which constructs an event-style discourse semantics, [17]. The discourse dynamics are encoded in continuation semantics [54] and various rhetorical relations are embedded in the resulting interpretation of the framework. They assume that discourse and sentence are distinct semantic objects, that play different roles in meaning evaluation. Moreover, two sets of composition functions, for handling different discourse relations, are introduced.

6.6. Modeling pathological discourse

Participant: Maxime Amblard.

Maxime Amblard starts a conjoint work with a psychologist Michel Musiol (IntePsy) and a philosopher Manuel Rebuschi (Archives Poincaré) about developing a formal analysis of pathological conversations involving schizophrenic speakers [18]. Such conversations give rise to manifest incongruities or ruptures that can be seen as mere contradictions by any “normal” speaker. Our construal relies both on semantic and pragmatic features of conversation. After an overview on the making of the corpus, we propose a SDRT-inspired account of pathological conversations, and we apply it to two relevant excerpts. We conclude with a short discussion about the localization of incoherencies by schizophrenics, either in semantics or in pragmatics, and its importance for our understanding of thought disorders.
6. New Results

6.1. Embedded data management

Participants: Nicolas Anciaux, Luc Bouganim, Yanli Guo, Lionel Le Folgoc, Philippe Pucheral, Shaoyi Yin.

Inspired by low cost economic models, this work draws the idea of a one-dollar database machine, with the objective to disseminate databases everywhere, up to the lightest smart objects. In contrast to traditional database machines relying on massively parallel architectures, the one-dollar database machine considers the cheapest form of computer available today: a microcontroller equipped with GBs size (external) Flash storage. Designing such a database machine is very challenging due to a combination of conflicting RAM and NAND Flash constraints. To tackle this challenge, this work proposes a new paradigm based on database serialization (managing all database structures in a pure sequential way) and stratification (restructuring them into strata when a scalability limit is reached). We show that a complete DBMS engine can be designed according to this paradigm and demonstrate the effectiveness of the approach through a performance evaluation.

This work capitalizes on previous results related to the indexing of Flash resident data [39] and has also obvious connections with the more general study we are conducting on Flash-based data management (see Section 6.2). Partial elements of this solution have been demonstrated at [28].

6.2. Flash-based Data Management

Participant: Luc Bouganim.

Bimodal flash devices. While disks have offered a stable behavior for decades, thus guaranteeing the timelessness of many database design decisions, flash devices keep on mutating. Many researchers have proposed to adapt database algorithms to existing flash devices. However, today, there is no reference DBMS design based on solid assumptions of flash devices behavior, precisely because flash device behavior varies across models, across firmware updates and possibly over time for the same model: database researchers are running after flash memory technology. In this study, we took the reverse approach and defined how flash devices should support database management. We advocated that flash devices should provide guarantees to a DBMS so that it can devise stable and efficient IO management mechanisms. Based on the characteristics of flash chips, we defined a bimodal FTL that distinguishes between a minimal mode where sequential writes, sequential reads and random reads are optimal while updates and random writes are forbidden, and a mode where updates and random writes are supported at the cost of sub-optimal IO performance. This work started at the end of 2010 and was published at CIDR’11 [19], in cooperation with the IT University of Copenhagen. DBMS/Flash device co-design considerations were the focus of a tutorial on flash devices given recently at VLDB 2011 [20].

6.3. Privacy-Preserving Data Publishing

Participants: Tristan Allard, Benjamin Nguyen, Philippe Pucheral.

While most PPDP works make the assumption of a trusted central publisher, this study advocates a decentralized way of publishing anonymized datasets. More precisely, our work concerns the proof of feasibility of adapting traditional PPDP schemes, such as $k$-anonymity, $\ell$-diversity or differential privacy to encompass the use of secure portable devices. In the applications we consider, each secure device is a data provider with weak computing capacities and weak connectivity (frequency and duration of connections are unpredictable)\(^1\). Weak

\(^1\)E.g., in the e-health context, patients may have their medical folder embedded in a secure device and connect it sporadically when they visit their physician or when they want to consult it at home.
connectivity precludes any P2P solution to the problem. A server allowing asynchronous communications between the devices becomes necessary to implement a distributed PPDP mechanism but this server does not benefit from the same trustworthiness as the participating devices. Our work aims to provide a generic method to adapt an important subclass of PPDP algorithms to this context, using both the limited secure computation capacities of each device (but taking advantage of their number) and the powerful computation abilities of an untrusted server available 24/7. Our proposal is based on a meta algorithm divided in three phases: (1) a collection phase where encrypted data is collected by the untrusted server, (2) a construction phase where the untrusted server performs a sound computation of a given privacy mechanism to generate sanitization rules and (3) a sanitization phase where the encrypted data is decrypted then sanitized by the devices to produce a final clear-text result. The last phase can be distributed using many different devices for better efficiency. In [15], [17], we showed how it is possible to transform existing anonymity mechanisms into decentralized ones using secure devices, while maintaining equivalent security guarantees against honest-but-curious and weakly malicious adversaries. In [16], we studied the (unlikely) event that some secure devices might be compromised, and can collude with the untrusted server. We provided schemes to detect the compromised devices with a probability that can be fixed as close to 1 as desired (the trade-off being the latency of the protocol).

6.4. Minimal Exposure

Participants: Nicolas Anciaux, Benjamin Nguyen.

When users request a service, the service provider usually asks for personal documents to tailor its service to the specific situation of the applicant. For example, the rate and duration of consumer’s loans are usually adapted depending on the risk based on the income, assets or past lines of credits of the borrower. In practice, an excessive amount of personal data is collected and stored. Indeed, a paradox is at the root of this problem: service providers require users to expose data in order to determine whether that data is needed or not to achieve the purpose of the service. We currently explore a reverse approach, where service providers would publicly describe the data they require to complete their task, and where the applicants would confront those descriptions with their own data to determine themselves the minimal subset of information to expose. We have first investigated solutions for simplistic tasks (e.g., evaluating a decision tree to determine the loan rate and duration a given applicant can claim), and we plan to address more complex ones (e.g., building the profile of customers, mining association rules, etc.) in the short term. The work on Minimal Exposure has just started and a first paper is under evaluation.

6.5. Experiment in the medical field

Participants: Nicolas Anciaux, Luc Bouganim, Lionel Le Folgoc, Philippe Pucheral, Alexei Troussov.

The PlugDB engine is being experimented in the field since September 2011 to implement a secure and portable medical-social folder. The objective is to improve the coordination of medical care and social services provided at home for dependent people. Details related to this experiment conducted with about 120 practitioners and patients are given in Section 7.2. While this action did not generate new academic results (though it helped us validating some previous results), it imposed us a strong investment in terms of test and optimization for our prototype and in terms of communication to promote this experiment at the regional level.
6. New Results

6.1. MLIF

TALARIS contributes to ISO TC 37 committee “Terminologies and other Language Resources”, and more specifically to the activities of its SC3 “Computer Applications in Terminology”, and SC4 “Linguistic Resources Management”. Within TC37/SC4, TALARIS is currently contributing, as project leader, to the definition and specification of the Multi Lingual Information Framework (MLIF) [ISO FDIS 24616]. MLIF is being designed with the objective of providing a common abstract model being able to generate several formats used in the framework of translation and localization. MLIF has been released as FDIS (Final Draft International Standard) and it should finally be published as an official ISO standard soon. MLIF has been extensively used within the ITEA2 META VERSE1 project. [42], [43], [12].

6.2. TEXT CLASSIFICATION

Neural clustering algorithms show high performance in the general context of the analysis of homogeneous textual datasets. We have recently proposed a new incremental growing neural gas algorithm using the cluster label maximization (IGNGF) [44] [34]. In this strategy the use of a standard distance measure for determining a winner is completely suppressed by considering the label maximization approach as the main winner selection process. One if its important advantage is that it provides the method with an efficient incremental character as it becomes independent of parameters. Although it performs better than the standard clustering methods on textual data, we have shown this year than the obtained results are not as efficient as expected whenever an analysis of very complex heterogeneous textual datasets is performed [33]. We have thus explored several variations of IGNG-F approach based on combination of distance based criteria and cluster label maximization. Our new results on all kinds of datasets, especially on the most complex heterogeneous textual datasets, clearly reflect the advantages of our new algorithm as compared to other existing algorithms and to our former adaptations [29]. Cluster quality evaluation represents a key process for all kinds of data analysis tasks, and more especially for textual data. We have recently presented different variations of unsupervised Recall/Precision and F-measures measures that cope with the defects of classical indexes, like inertia-based indexes. Our new indexes directly exploit the maximized features of the data associated to each cluster after the clustering process without prior consideration of clusters profiles. As compared to classical indexes, their main advantage is thus to be independent of the clustering methods and of their operating mode. They thus altogether permit the objective comparison of clustering methods and represent a sound technique for efficient cluster labeling. We have more especially worked this year on the large scale validation our indexes using reference labeled textual datasets [35].

We are also currently investigating to set up a platform for efficiently assisting the patents experts in the process of patents validation. Reaching such a goal has implied to develop new semi-supervised classification methods or propose in-deep adaptation of existing ones in order to establish relevant relationships between hierarchical patents classification and bibliographical references describing research covering the fields related to the different patents classes. In this context, we have successfully explored this year new classification techniques based on taboo search [14].

To cope with the current defects of existing incremental clustering methods, an alternative approach for analyzing information evolving over time consists in performing diachronic analysis. We have thus explored this year different an original technique based on this approach on texts by the use of the combination of cluster labeling with unsupervised Bayesian reasoning between cluster labels extracted from clustering model issued from different time periods. Based on a reference dataset issued from the IST-PROMTECH project, we have clearly shown that these new techniques, whilst providing a new framework for automatizing such kind of analysis, outperformed existing ones [32] [31] [30].
6.3. DIALOG

Within the Emospeech project, we developed the Emospeech Dialogue Toolkit (cf. Software section); used the Wizard of Oz infrastructure it includes to collect dialog data; and trained an interpreter and a dialog manager. The collected data comprises 591 dialogues in French collected within the context of the Mission Plastechnology serious game, 4874 utterances, 77854 words and 1321 player utterances containing 12901 word tokens and 1427 word types. We collected in average 50 conversations for each sub-dialogue in the game. Dialog length varies from 78 to 142 with an average length of 106 utterances per dialog.

6.4. GIVE

For the Generation of Instruction in Virtual Environment challenge edition 2.5 (GIVE), we developed two systems. The first system is the successor of the system that we presented to the GIVE 2 challenge (2010). We solved two known problems of this system, namely the indefinite presupposition problem and the ambiguity arising from underspecified referring expressions [26]. The GIVE 2.5 challenge proved that these improvements were efficient, and showed an increase of 21% in terms of task success (47% in GIVE 2, 68% in GIVE 2.5). The second system, developed in collaboration with the University of Cordoba is the first to our knowledge that uses a human-human corpus to provide whole utterances thanks to plan matching techniques [22], [21]. The system ranked fifth in terms of task success (58%), but second in terms of referring targets identification. The naturalness of instructions and the simplified system development makes it an interesting research track to follow.

6.5. Verb Classification

To help computer systems in the task of understanding and representing the full meaning of a text, verb classifications have been proposed which group together verbs with similar syntactic and semantic behaviour. For English verbs, VerbNet provides such a large scale classification but there are no similar French resource available. We investigated different ways both of automatically constructing such a resource; and of evaluating it.

Using Formal Concept Analysis (FCA), we developed a method for classifying verbs based on their (syntactic) subcategorisation information extracted from existing French lexical resources; and by translating the English Verbnet, we showed how to associate the obtained classes with semantic information represented by Verbnet’s thematic role sets ([27]). As a result, a VerbNet like classification for French verbs can be constructed fully automatically.

The FCA approach we pursued, first builds a classification based on verbs and verb features and second filters this classification using various metrics (e.g., concept probability, concept stability). We are currently comparing this approach with a clustering approach which makes use of detailed evaluation metrics [44] and uses probabilistic information to guide classification. First results are promising and outperform the state of the art methods in this domain [63].

One important difference between the clustering and the FCA approach we experimented with is that only the second, allows a verb to belong to several classes. Since verbs are highly ambiguous, this is an important difference. To evaluate the impact of this difference on the usability of the classifications built by each of the methods, we are currently conducting a task-based, extrinsic evaluation of both classifications by analysing their impact when used in a Semantic Role Labeling task on a French corpus.

6.6. I-FLEG

Within the Allegro project, we developed the I-FLEG game [16], [15], a virtual world in which the learner exercises French by clicking on objects and answering the questions raised by the system. The language learning exercises produced by I-FLEG are automatically generated using the GenI sentence generator from a knowledge base describing the virtual world. A preliminary evaluation of I-FLEG with school children [17] suggests that the “game” aspect increases learner motivation and that spoken output is essential in maintaining learner interest.
TEMICS Project-Team

6. New Results

6.1. Analysis and modeling for compact representation and navigation

6.1.1. Joint projection/filling method for virtual view synthesis

Participants: Christine Guillemot, Vincent Jantet.

This study is carried out in collaboration with INSA/IETR (Luce Morin). Associated with a view synthesis method, a multi-view plus depth video allows the generation of virtual views of the scene from any viewpoint. Many algorithms have thus been developed to synthesize virtual views from one or several input views video plus depth data. These rendering algorithms are either based on Image-Based Rendering (IBR) techniques or Geometry-Based Rendering (GBR) techniques, according to the amount of 3D information they use. IBR techniques require limited geometric information to synthesize intermediate views and allow the generation of photo-realistic virtual views at the expense of virtual camera freedom. GBR techniques require detailed 3D models of the scene to synthesize arbitrary viewpoints (points of view). GBR techniques are sensitive to the accuracy of the 3D model, which is difficult to estimate from real multi-view videos. Depth-Image-Based Rendering (DIBR) techniques include hybrid rendering methods between IBR and GBR techniques. DIBR methods are based on warping equations, which project a reference view onto a virtual viewpoint. Each input view is defined by a "color" (or "texture") map and a "depth" map, which associates a depth value to each image pixel.

In classical DIBR schemes, the rendering proceeds in several distinct steps, each one designed to solve a specific problem. First, the input depth map is warped onto the virtual viewpoint. The obtained warped depth map contains disocclusions, cracks and ghosting artifacts. Second, this virtual depth map is filtered with a median filter, in order to remove the cracks, and then to dilate disocclusion areas on the background side, in order to avoid ghosting artifacts during view synthesis. Third, the filtered depth map is used in a backward warping to compute the color of each pixel of the virtual view. Fourth, this resulting depth map is inpainted, to fill in disocclusion areas. Finally, this complete depth map is used by a depth-aided inpainting algorithm to fill in disocclusions in the color map. However, all these steps are inter-dependent, and errors introduced by each one are amplified by the following one. Connectivity information is lost during the first projection step, as shown in Fig. 2. Without this connectivity information, every inpainting method fails to fill in background disocclusions if the disoccluded area is surrounded by foreground objects. This case may happen each time a foreground object is not convex, and contains holes, as shown in Fig. 2 -(a). As a result, depth-aided inpainting uses wrong foreground patches to fill in background disocclusions, producing annoying artifacts, as shown in Fig. 2 -(b).

We have developed two DIBR techniques, both based on a novel forward projection technique, called the Joint Projection Filling (JPF) method [16]. The JPF method performs forward projection, using connectivity information to fill in disocclusions in a single step. The first proposed DIBR method is designed to extrapolate virtual views from a single input view plus depth video sequence. The synthesis of virtual depth maps by the JPF method avoids the use of dedicated filtering and inpainting processes and leads to synthesized depth maps of higher quality. The second proposed DIBR method is designed to interpolate intermediate views from multiple input view plus depth sequences. This interpolation method uses the Floating Texture approach to register multiple inputs view plus depth sequences before blending. The JPF method fills in disocclusion areas during the projection, to ensure that geometrical structures are well preserved. The method uses the occlusion-compatible ordering presented by McMillan, which uses epipolar geometry to select a pixel scanning order. The occlusion-compatible ordering is used to handle disocclusions gracefully. Cracks are filled in by interpolation of neighboring pixels, whereas disocclusions are only filled in by background pixels. This technique can also be used with non-rectified views, avoiding prior creation of parallax maps.
6.1.2. 2D/3D image inpainting for virtual view synthesis

Participants: Josselin Gauthier, Christine Guillemot, Mouid Keskes, Olivier Le Meur.

Inpainting methods play an important role in a wide range of applications. Removing text and advertisements (such as logos), removing undesired objects, noise reduction and image reconstruction from incomplete data are the key applications of inpainting methods. Algorithms can be classified into two categories: PDE (Partial Derivative Equation)-based schemes and examplar-based schemes. The former uses diffusion schemes in order to propagate structures in a given direction. Their drawback is the introduction of blur due to diffusion. The latter relies on the sampling and the copying of texture from the known parts of the picture.

We have proposed a novel inpainting algorithm combining the advantages of both aforementioned methods. As in Criminisi et al’s approach, the proposed method involves two steps: first, a filling order is defined to favor the propagation of structure in the isophote direction. Second, a template matching is performed in order to find the best candidates to fill in the hole. Compared to previous approaches, the main contributions concern the use of structure tensors to define the filling order instead of field gradients. The structure tensor is defined as follow:

\[ J = \sum_{i=\{R,G,B\}} \nabla I_i \nabla I_i^T \]  

(39)

\( J \) is the sum of the scalar structure tensors \( \nabla I_i \nabla I_i^T \) of each image channel \( I_i \) (\( i \in \{R,G,B\} \)). Information about local geometry can be deduced by computing the eigenvalues and eigenvectors of \( J \). The local vector geometry is computed from the structure tensor \( J \). Its eigenvectors \( v_{1,2} \) (\( v_i \in \mathbb{R}^n \)) define an oriented orthogonal basis and its eigenvalues \( \lambda_{1,2} \) define the amount of structure variation. \( v_1 \) is the orientation with the highest fluctuations (orthogonal to the image contours), and \( v_2 \) gives the preferred local orientation. This

---

eigenvector (having the smallest eigenvalue) indicates the isophote orientation. The use of structure tensor allows to retrieve a more coherent local geometry. The computation of the filling order as proposed by Criminisi et al is then replaced by a term coming from PDE-based schemes, called Coherence Enhancing Diffusion. The use of structure tensor in an exemplar-based approach leads to a more robust algorithm that visually improves the quality of the inpainted areas.

Additionally the simple template matching originally used in previous methods has been improved by using a K-nearest neighbor approach. The weights of the linear combination of the first $K$ best candidate are adjusted by taking into account that all candidate patches are not equally reliable. Note that the number $K$ is also locally adjusted in function of the local spatial complexity.

The 2D inpainting algorithm described above has been extended to deal with 3D content. In this work the goal is to synthesize novel views directly from the original images. Image-based rendering (IBR) is commonly used to render a virtual view. It generates a nearby viewpoint image by projecting a point from the reference view to the virtual view using the depth data. However, when the viewpoint is shifted, occluded regions in the original view point are disoccluded. Handling these disocclusions (holes) is a difficult problem. We propose to use an extension of the 2D inpainting method to fill in these holes. For this goal, we have modified the computation of the structure tensor by adding the depth information. Equation (2) is simply modified as follow:

\[ J = \sum_{i=(R,G,B,Z)} \nabla I_i \nabla I_i^T \]

where $Z$ represents the depth map. As previously, the tensor is used to compute the filling order. A directional term is also included in order to favor a filling direction. Specifically, when the viewpoint is shifted from left to right in the horizontal direction, occluded regions in the left image appear in the right image around the right side of the object. Therefore, it is recommended to start the filling from the right to the left. This filling is performed by a modified template matching using texture information as well as depth data. Figure 3 illustrates the inpainting quality for different approaches.

![Figure 3. Virtual synthesized view. From left to right: original view projected into the new viewpoint; disocclusions filled by Criminisi’s approach, Daribo’s approach and the proposed method.](image)

6.1.3. Computational modelling of visual attention

Participants: Josselin Gauthier, Olivier Le Meur.

6.1.3.1. Eye-movement study:

In 2011, we have investigated whether two populations of visual fixation exist in 2D context. The question is simple: do all visual fixations have the same role in the free viewing of natural scenes? Recent studies suggest that there are at least two types of visual fixations: focal and ambient fixations. The former is believed to be used to inspect local areas accurately, whereas the latter is used to obtain the context of the scene.
From a collaboration with Technicolor (P. Guillotel and C. Chamaret) and LUTIN (T. Baccino), we found new evidence to support a focal-ambient dichotomy. Our results published in the journal i-Perception [14] indicate that the determining factor to classify the visual fixations is the saccade amplitude. We proposed an automatic system to cluster visual fixations in two groups using four types of natural scene images. From this automatic classification, the terms focal saliency map and ambient saliency map have been introduced. The dependence on the low-level visual features and the time course of these two kinds of visual fixations were examined. Our results demonstrate that there is an interplay between both fixation populations and that focal fixations are more dependent on low-level visual features than ambient fixations. These results might have a strong impact on both the computational modelling of visual attention and their performance assessment.

A second study related to eye-movement dealt with the role of the binocular disparity depth cue in the deployment of visual attention. To address this point, we compared eye tracking data recorded while observers viewed natural images in 2D and 3D conditions. The influence of disparity on the inter-observers congruency, saliency, center and depth bias was first examined. Results show that visual exploration in depth layer detection task is affected by the binocular disparity. In particular, participants tend to look first at closer areas just after the stimuli onset with the introduction of disparity, and then direct their gaze to more widespread locations. Our results has been submitted in the journal Cognitive Computation.

6.1.3.2. Model of visual attention:

Since 1998 with the publication of the influential work of Itti, Kock and Niebur[11], the computational modelling of the visual attention has known an increasing interest. The former models only used the low-level visual features for getting a saliency map. They perform well in a number of cases in predicting where an observer would look at. However, to improve the quality of the prediction, it seems unavoidable to use high-level information in order to account for visual deployment.

This work aims at designing a computational model mixing low-level and high-level features. Among the different factors influencing our gaze, we have focused our works on two cues: the dominant depth and the horizon line position. The dominant depth and the spatial position of the horizon line were inferred from the low-level visual features. A training database has been set up to perform a learning. Results indicate that the proposed model outperforms state-of-the-art models [37].

From behavioual studies on eye-movement in a 3D context, we have proposed a model of visual attention able to predict saliency of 3D pictures. The method developped aims at using the depth cue, the central bias and the low-level visual features. The predicted saliency is obtained by linearly combining these cues. The weights of the linear combination are learnt from a training database and are time-dependent. This study is under revision in the journal Cognitive Computation.

6.1.3.3. Predicting the inter-observer visual congruency:

This work aims at predicting the inter-observer visual congruency (IOVC), indicating the congruence or the variability among different subjects looking at the same image [35]. Predicting this congruence is of interest for image processing applications where the visual perception of a picture matters such as website design, advertisement, etc. We proposed a computational model of the IOVC. This new model is a mixture of low-level visual features extracted from the input picture. Model’s parameters are learned by using a large eye-tracking database. In this study, we also proposed a new scheme to compute the depth of field of a picture. Finally, once the training and the feature extraction have been carried out, a score ranging from 0 (minimal congruency) to 1 (maximal congruency) is computed. A value of 1 indicates that observers would focus on the same locations and suggests that the picture presents strong locations of interest. To illustrate the interest of the proposed model, we have used it to automatically rank personalized photograph. Figure 4 illustrates the proposed approach.

6.1.4. Visual cues analysis and modelling

Participants: Safa Cherigui, Christine Guillemot.

This work is carried out in collaboration with Technicolor (D. Thoreau, Ph. Guillotel, P. Perez) and aims at designing a compression algorithm based on the concept of epitomes. An epitome is a condensed representation of an image (or a video) signal containing the essence of the textural properties of this image. Different forms of epitomes have been proposed, such as a patch-based probability model learned either from still image patches or from space-time texture cubes taken from the input video. These probability models together with appropriate inference algorithms, are useful for content analysis inpainting or super-resolution. Another family of approaches makes use of computer vision techniques, like the KLT tracking algorithm, in order to recover self similarities within and across images. In parallel, another type of approach consists in extracting epitome-like signatures from images using sparse coding and dictionary learning.

The method developed aims at tracking self-similarities within an image using a block matching (BM) algorithm [25]. The epitome is constructed from disjoint pieces of texture (“epitome charts”) taken from the original image and a transform map which contains translational parameters (see Fig. 5 - middle row). Those parameters keep track of the correspondences between each block of the input image and a block of the epitome. An Intra image compression scheme based on the epitome has been developed showing a rate saving of up to 12.

6.2. Representation and compression of large volumes of visual data

6.2.1. 3d representations for multi-view video sequences

Participants: Christine Guillemot, Vincent Jantet.

Multi-view plus depth video content represent very large volumes of input data which need to be compressed for storage and transmission to the rendering device. The huge amount of data contained in multi-view sequences indeed motivates the design of efficient representation and compression algorithms. In collaboration with INSA/IETR (Luce Morin), we have studied layered depth image (LDI) and layered depth video (LDV) representations as a possible compact representation format of multi-view video plus depth data. LDI gives compact representations of 3D objects, which can be efficiently used for photo-realistic image-based rendering (IBR) of different scene viewpoints, even with complex scene geometry. The LDI extends the 2D+Z representation, but instead of representing the scene with an array of depth pixels (pixel color with associated depth values), each position in the array may store several depth pixels, organised into layers.

Various approaches exist to construct LDI, which all organize layers by visibility. The first layer contains all pixels visible from a chosen reference viewpoint. The other layers contain pixels hidden by objects in previous layers. With classical construction solutions, each layer may contain pixels from the background and pixels from objects in a same neighbourhood, creating texture and depth discontinuities within the same layer. These
Figure 5. Epitome and reconstructed images: (top row) Original images (columns a and b); (middle) Epitomes; (bottom) Reconstructed images from epitome texture and transform map.
discontinuities are blurred during the compression process which in turn significantly reduces the rendering quality. We have thus developed a novel object-based LDI representation which is more tolerant to compression artifacts, as well as being compatible with fast mesh-based rendering techniques [34]. This representation organises LDI pixels into two separate layers (foreground and background) to enhance depth continuity (see Fig. 6). The construction of this object-based LDI makes use of a foreground-background region-growing segmentation algorithm followed by inpainting of both colour and texture images to have a complete background layer (without the holes corresponding to disocclusion areas). The costly inpainting algorithm is thus processed once, during the LDI classification, and not during each view synthesis, which helps to speed up the rendering step.

![Figure 6. Object-based LDI: (top) Foreground and background layers; (bottom) Rendering results classical and object LDI.](image)

6.2.2. From sparse to spread representations

**Participant:** Jean Jacques Fuchs.

Sparse representations, where one seeks to represent a vector on a redundant basis using the smallest number of basis vectors, appear to have numerous applications. The other extreme, where one seeks a representation that uses all the basis vectors, might be of interest if one manages to spread the information nearly equally over all of them. Minimizing the $\ell_\infty$-norm of the vector of weights is one way to find such a representation. Properties of the solution and a dedicated fast algorithm have been developed. While the application of such models in robust data coding and in improving achievable data rates over amplitude constrained channels seems to be wishful thinking, its use in indexing techniques appears to be promising. In this context, one further replaces the optimal vector by its sign vector (potentially associated with a re-evaluated scalar weight) to get a binary vector that is not only cheap to store and (somehow) easy to search for but also allows for an explicit reconstruction unlike all other Hamming embedding functions used to map real vectors into binary vectors.
6.2.3. On-line dictionary learning methods for prediction

Participants: Christine Guillemot, Mehmet Turkan.

One crucial question to the problem of sparse approximation, and hence also of prediction based on sparse approximation, is the choice of the dictionary. Various advanced dictionary learning schemes have been proposed in the literature for the sparse signal approximation problem, so that the dictionary used is well suited to the data at hand. The popular dictionary learning algorithms include the K-SVD, the Method of Optimal Directions (MOD), Sparse Orthonormal Transforms (SOT), and (Generalized) Principle Component Analysis (PCA). However, the above learning methods are often used off-line since their computational complexity, which results from the number and the dimension of training samples, makes them inappropriate for online learning. In addition, these methods are adapted to the learning of basis to be used for approximating input data vectors, but not to the problem of predicting unknown samples from noisy observed samples in a causal neighborhood.

In 2011, we have developed a method for on-line training dictionaries adapted to the prediction problem [41]. Let $A$ be the input dictionary, which is divided into two sub-dictionaries: $A_c$ and $A_t$. The goal is to have a simple on-line dictionary learning method which is adapted to the prediction problem, i.e., which will learn both sub-dictionaries so that sparse vectors found by approximating the known samples (the template) using the first sub-dictionary ($A_c$) will also lead to a good approximation of the block to be predicted when used together with the second sub-dictionary ($A_t$). When dealing with the prediction problem, the sparse signal approximation is indeed first run with a set of masked basis functions (dictionary $A_c$), the masked samples corresponding to the location of the pixels to be predicted. The principle of the approach is to first search for the linear combination of basis functions which best approximates known sample values in a causal neighborhood, and keep the same linear combination of basis functions, but this time with the unmasked samples (dictionary $A_t$) to approximate the block to be predicted. The decoder similarly runs the algorithm with the masked basis functions and taking the previously decoded neighborhood as the known support. The use of masked basis functions converts the complete approximation problem into an overcomplete approximation problem. Because of its simplicity, of the limited number of training samples it requires, can be used for online learning of dictionaries, i.e. while doing the block-wise encoding of the image. The training samples are all possible previously coded/decoded texture patches (blocks of pixels) within a search window located in a causal neighborhood of the block to be predicted.

6.2.4. Neighbor embedding methods for image prediction and inpainting

Participants: Christine Guillemot, Mehmet Turkan.

The problem of texture prediction as well as image inpainting can be regarded as a problem of texture synthesis. Given observations, or known samples in a spatial neighborhood, the goal is to estimate unknown samples of the block to be predicted or of the patch to be filled in inpainting. We have developed texture prediction methods as well as a new inpainting algorithm based on neighbor embedding techniques which come from the area of data dimensionality reduction. The methods which we have more particularly considered are Locally Linear Embedding (LLE) and Non-negative Matrix Factorization (NMF). The first step in the developed methods consists in searching, within the known part of the image, for the $K$ nearest (KNN) patches to the set of known samples in the neighborhood of the block to be predicted or of samples to be estimated in the context of inpainting. This first step can be seen as constructing a dictionary matrix by stacking in the matrix columns the vectorized K-NN texture patches. The non-negative dictionary matrix $A \in \mathbb{R}^{N \times M}$, is formed by $K$ nearest neighbors to the vector formed by the known samples in the neighborhood of the samples to be predicted. These $K$ nearest neighbors are texture patches of the same shape taken from the known part of the image. This dictionary can then be used for approximating the known samples by masking the rows of the matrix which correspond to the position of the unknown samples, solving a least squares problem under the constraint of sum-to-one of the weights in the case of LLE, or under the constraint of non-negativity of the weights for NMF. It is actually a variant of NMF since one of the components matrices is fixed (the one corresponding to the dictionary matrix) and only the matrix containing the weights of the linear approximation must then be found. The approaches are thus intended to explore the properties of the manifolds on which the
input texture patches are assumed to reside. The underlying assumption is that the corresponding uncomplete and complete patches have similar neighborhoods on some nonlinear manifolds. The new prediction methods give RD performances which are significantly better than the ones given by the H.264 Intra prediction modes, in particular for highly textured images [21], the highest gain being achieved with NMF.

A new examplar-based inpainting algorithm using neighbor embedding techniques has been developed. A new priority order has been proposed in order to inpaint first patches containing structures or contour information. The methods have also been shown to enhance the quality of inpainted images when compared to classical examplar-based solutions using simple template matching techniques to estimate the missing pixels, (see Fig. 7).

![Figure 7. Inpainting results: (left) mask of the image to be inpainted; (middle) Inpainting results with classical examplar-based inpainting; (right) Inpainting results with LLE (right).](image)

### 6.2.5. Lossless coding for medical images

**Participants:** Claude Labit, Jonathan Taquet.

Last year, we developed a hierarchical oriented prediction (HOP) algorithm, for resolution scalable lossless and near lossless compression of biomedical images. In 2011, the algorithm has been slightly improved with an iterative optimization of the predictors in order to get better results on less noisy/smooth images [39].

Recently, there have been a growing interest for the compression of an emerging imaging modality: the virtual microscopy (VM). It is used in anatomopathology and may produce huge images of more than 1 Gigabytes. We have studied the efficiency for lossless and lossy compression of our previously developed algorithms HOP and OWD (optimized wavelet decomposition) and of two extensions of OWD: near-lossless and/or region of interest (ROI) coding. The lossless results, which are slightly better than JPEG-LS and JPEG-2000 standards with about 3:1 compression ratio, show that lossless compression is not suited to VM. By compressing only the information area (ROI) which represents about 20 percents of the size of test images, 9:1 ratio could be obtained, and combined with near-lossless approach, depending on the required quality, ratio can reach 17:1 with no visual losses to more than 30:1 with some visual losses (or approximately about 6:1 for ROI only data). We have concluded that it would probably be better to use lossy or efficient quality scalable compression. Because those images have specific contents (cellular tissues for example) we have also introduced and investigated new learning based methods. We have developed an optimization process for designing multiple KLT (Karhunen-Loeve Transform) in order to get orthonormal bases that are optimal for decorrelation and quality scalability. This learning approach has been applied as an a-posteriori transform of a wavelet decomposition in order to propose transforms with no blocking artefacts. A fully quality-scalable coding algorithm allows to obtain interesting PSNR improvements compared to the optimized coding process.
of JPEG-2000. Gain is around 0.5 dB for 16:1 compression of ROI only data, and more than 1 dB for 8:1 compression ratio.

6.3. Distributed processing and robust communication

6.3.1. Loss concealment based on video inpainting

Participants: Mounira Ebdelli, Christine Guillemot, Olivier Le Meur.

In 2011, we have started developing a loss concealment scheme based on a new video exemplar-based inpainting algorithm. The developed video inpainting approach relies on new patch priority functions as well as on a motion confidence-aided neighbor embedding techniques. Neighbor embedding approaches aim at approximating input vectors (or data points) as a linear combination of their neighbors. The search of the weights of the linear combination (i.e. of the embedding) are formulated as constrained least squares problems. When using the locally linear embedding, the constraint is that the sum of the weights is equal to 1. We have also considered non-negative matrix factorization to solve the problem, in which case the constraint is that the weights and the other vector are non-negative. The motion confidence introduced in the neighbor embedding improves the robustness of the algorithm in the sense that it limits the error propagation effects which otherwise result from uncertainties on the motion information of the unknown pixels to be estimated. A new patch similarity measure which accounts for the correlation between motion information has been defined for the $K$-NN search inherent to neighbor embedding techniques. Evaluations of the algorithm in a context of video editing (object removal) are on-going. The next step will be to assess the performance of the approach in a context of loss concealment, that is to estimate unknown pixels after decoding when the corresponding transport packets have been lost on the transmission network.

6.3.2. Unequal Erasure Protection and Object Bundle Protection

Participant: Aline Roumy.

In 2011, we started a new collaboration in the framework of the Joint INRIA/Alcatel Lucent lab. In this work, carried out with V. Roca (Planete, INRIA), B. Sayadi and R. Imad (Alcatel Lucent), we proposed and analyzed a novel technique capable of providing both an unequal erasure protection service and an object bundle protection service.

Unequal Erasure Protection: When a data flow contains information of different priority levels, it is natural to try to offer an unequal protection where the high priority data benefits from a higher protection than the rest of data. In this work we focused on the "erasure channel", for instance the Internet where the UDP/IP datagram integrity is guaranteed by the physical layer FCS (or CRC) and the UDP checksum. In this context UEP refers to an Unequal Erasure Protection (rather than Error) and the FEC code being used is one of the various Application-Layer Forward Erasure Correction (AL-FEC) codes that have been designed and standardized in the past years, like Reed-Solomon, one of the LDPC variants, or Raptor(Q) codes. Offering an unequal protection in this context can be achieved by one of the following three general approaches: by using dedicated UEP-aware FEC codes, by using a dedicated UEP-aware packetization scheme, or by using an UEP-aware signaling scheme. In this work we focused on and compared the last two approaches and more precisely the well known Priority Encoding Transmission (PET) scheme that belongs to the UEP-aware packetization category and a Generalized Object Encoding (GOE) scheme, we proposed [53], that belongs to the UEP-aware signaling category. Through a careful modeling of both proposals [55], whose accuracy has been confirmed by simulations, we have demonstrated that the protection performance (i.e. erasure resiliency and average decoding delay) of both approaches are equivalent, not only asymptotically but also in finite length conditions. In fact the key differences between these approaches become apparent when applying them in practical systems. Such metrics as the simplicity of the solution, the number of packets processed, the maximum memory requirements, the number of FEC encoding and decodings, as well as the system of linear equations complexity (number of variables) are in favor of the GOE approach.
**Object Bundle Protection:** we considered the use of PET, more precisely an extension called Universal Object Delivery (UOD), and GOE in situations where one needs to send a bundle of small object (e.g. files). If both solutions can address this need, we showed that once again the GOE scheme is highly recommendable for practical realizations. This is mostly due to the lack of flexibility of the PET/UOD approach. For instance the limited size of a packet creates an upper bound to the number of objects that can be considered together (e.g. UOD limits this number to 255), the symbol size has a coarse granularity (e.g. UOD requires symbols to be multiple of 4 bytes when used with RaptorQ codes) which can create rounding problems with certain sets of objects (i.e. the actual packet size may be significantly shorter than the target, and/or the actual code rate significantly different than its target). GOE has no such constraints. In particular GOE offers the possibility to adjust the packet interleaving to the use-case and channel erasure features. One can easily trade robustness in front of long erasure bursts for very short decoding delays of high priority objects and low memory requirements, which can be a key asset in case of small, lightweight terminals or timely delivery services. This feature may be sufficiently important to justify by itself the use of a GOE FEC Scheme [55].

6.3.3. **Distributed compressed sensing**

**Participants:** Aline Roumy, Velotiaray Toto-Zarasoa.

This work has been performed in collaboration with E. Magli and G. Coluccia (Politecnico di Torino) in the framework of the FP7 IST NOE NEWCOM++ (Jan. 2008 - Apr. 2011). A new lossy compression scheme for distributed and sparse sources under a low complexity encoding constraint has been proposed in [26]. This problem naturally arises in wireless sensor networks. Indeed, nodes of a sensor network may acquire temperature readings over time. The temperature may vary slowly over time, and hence consecutive readings have similar values. However, they also have inter-sensor correlation, as the sensors may be in the same room, in which the temperature is rather uniform. The question hence arises of how to exploit intra- and inter-sensor correlations without communication between the sensors and with a low complexity acquisition process in order to save energy consumption at the sensor. Therefore, we consider continuous, correlated, distributed and sparse (in some domain) sources and perform lossyuniversal compression under a low encoding complexity constraint.

In order to meet the low complexity encoding constraint, the encoding stage is performed by a lossy distributed compressed sensing (CS). More precisely, the proposed architecture is based on the joint use of CS to capture memory of a signal, and DSC to take advantage of inter-sensor correlations. First, we showed that the resilience of CS to quantization error also holds in the distributed setup. Moreover, the optimal number of measurements can be chosen as the one guaranteeing (close-to-)perfect reconstruction. In addition, using joint decoding, dequantization and reconstruction techniques allows to boost performance even further. The joint use of CS and DSC allows to save 1.18 bit per source sample for the same PSNR quality w.r.t. the non-distributed CS scheme. Compared to the DSC scheme (without CS), we observe a gain increasing with the rate for the same PSNR quality. All these results makes the proposed scheme an attractive choice for environments such as sensor networks, in which the devices performing acquisition and processing are severely constrained in terms of energy and computations.

6.3.4. **Super-resolution as a communication tool**

**Participants:** Marco Bevilacqua, Christine Guillemot, Raul Martinez-Noriega, Aline Roumy.

In 2011, we started a new collaboration in the framework of the Joint INRIA/Alcatel Lucent lab. In this work, carried out with M.-L. Alberi (Alcatel Lucent), we proposed a novel technique capable of producing a high-resolution (HR) image from a single low-resolution (LR) image. This method that belongs to the class of single-image super-resolution (SR), offers the promise of overcoming the inherent limitations of the video acquisition and transmission systems. More precisely, one can think of sending a low resolution video to adapt to the complexity constraint of the encoder and/or the bandwidth limitation of the network, while the decoder reconstructs a high-resolution video.
As a first step toward the more ambitious goal of compressing video through SR, we proposed a novel method for single-image super-resolution based on a neighbor embedding technique. Each low-resolution input patch is approximated by a linear combination of nearest neighbors taken from a dictionary. This dictionary stores low-resolution and corresponding high-resolution patches taken from natural images and is thus used to infer the HR details of the super-resolved image. The entire neighbor embedding procedure is carried out in a feature space. Features which are either the gradient values of the pixels or the mean-subtracted luminance values are extracted from the LR input patches, and from the LR and HR patches stored in the dictionary. The algorithm thus searches for the $K$ nearest neighbors of the feature vector of the LR input patch and then computes the weights for approximating the input feature vector. The so-obtained weights are finally used to compute a linear combination of the corresponding HR patches, which yields the super-resolved image. The use of a positive constraint for computing the weights of the linear approximation is shown to have a more stable behavior than the use of sum-to-one constraint and lead to significantly higher PSNR values for the super-resolved images.
TEXMEX Project-Team

6. New Results

6.1. Advanced algorithms of data analysis, description

6.1.1. Advanced description techniques

6.1.1.1. Image joint description and compression

**Participant:** Ewa Kijak.

*This is a joint work with the TEMICS project-team (J. Zepeda and C. Guillemot).*

In the context of ANR project ICOS-HD ended at december 2010, in collaboration with Christine Guillemot from TEMICS, we investigated sparse representations methods for local image description. We have developed methods for learning dictionaries to be used for sparse signal representations. These methods lead to dictionaries which have been called Iteration-Tuned Dictionaries (ITDs), Basic ITD (BITD), Tree-Structured ITD (TSITD) and Iteration-Tuned and Aligned Dictionaries (ITAD). All three proposed ITD schemes (BITD, TSITD and ITAD) have been shown to outperform the state-of-the-art learned dictionaries in terms of PSNR versus sparsity. The performance of these dictionaries has also been assessed for both compression and de-noising applications. ITAD in particular has been used to produce a new image codec that outperforms JPEG2000 for a fixed image class and leads in 2011 to two new publications [49], [20].

6.1.1.2. Bag-of-colors

**Participant:** Hervé Jégou.

*This is joint work with Christian Wengert (Kooaba) and Matthijs Douze (INRIA LEAR and SED project-teams).*

This work investigates [48] the use of color information when used within a state-of-the-art large scale image search system. We introduce a simple color signature generation procedure, used either to produce global or local descriptors. As a global descriptor, it outperforms several state-of-the-art color description methods, in particular the bag-of-words method based on color SIFT. As a local descriptor, our signature is used jointly with SIFT descriptors (no color) to provide complementary information.

6.1.1.3. Aggregating local image descriptors into compact codes

**Participant:** Hervé Jégou.

*This is joint work with Matthijs Douze (INRIA LEAR and SED project-teams), Patrick Pérez (Technicolor), Florent Perronnin (Xérox Research Center Europe) and Cordelia Schmid (INRIA LEAR).*

This work [19] addresses the problem of large-scale image search and consolidates and extends results from a previous work [78]. Different ways of aggregating local image descriptors into a vector are compared, and the Fisher vector is shown to achieve better performance than the reference bag-of-visual words approach for any given vector dimension. We then jointly optimize dimensionality reduction and indexing in order to obtain a precise vector comparison as well as a compact representation. The evaluation shows that the image representation can be reduced to a few dozen bytes. Searching a 100 million image dataset takes about 250 ms on one processor core.

6.1.2. Browsing multimedia databases

**Participant:** Laurent Amsaleg.

*This is a joint work with Björn Pór Jónsson and Grímur Tómasson from the School of Computer Science, Reykjavik University, Iceland.*
Since the introduction of personal computers, personal collections of digital media have been growing ever larger. It is therefore increasingly important to provide effective browsing tools for such collections. We have proposed a multi-dimensional model for media browsing, called ObjectCube, based on the multi-dimensional model commonly used in OLAP applications. We implemented a prototype of a media browser based on the ObjectCube model. We then ran evaluations of its performance using three different underlying data stores and photo collections of up to one million photos.

6.1.3. Advanced data analysis techniques

6.1.3.1. Factorial analysis and output display for text and textual streams mining

Participant: Annie Morin.

Textual data can be easily transformed in frequency tables and any method working on contingency tables can be used to process them. Besides, with the important amount of available textual data, we need to find convenient ways to process the data and to get invaluable information. It appears that the use of factorial correspondence analysis allows us to get most of the information included in the data. We start exploring temporal changes in textual data and mainly focus on the visualization of results: we try to detect the topics if they have not already been identified and to study the evolution of the previous vocabulary inside a topic through time. In fact, as with economical datasets, we try to find seasonal components and cycling components in the documents and to characterize these components.

6.1.3.2. Intensive use of SVM for text and image mining

Participants: François Poulet, Thanh Nghi Doan.

Support Vector Machines (SVM) and kernel methods are known to provide accurate models but the learning task usually needs a quadratic program, so this task for very large datasets requires a large memory capacity and a long time. We have developed new algorithms. The first versions of the algorithms were based on a CPU distributed software program, then we have used GP-GPU (General Purpose GPU) versions to significantly improve the algorithm speed (130 times faster than the CPU one, 2500 times faster than libSVM, SVMPerf or CB-SVM). We have extended the least squares SVM algorithm (LS-SVM) to adapt the algorithm to datasets having a very large number of dimensions and have applied boosting to LS-SVM for datasets having simultaneously a very large number of vectors and dimensions on standard computers. In image classification, the usual frameworks involve three steps: feature extraction, building codebook by feature quantization and training the classifier with a standard classification algorithm (eg. SVM). However, task complexity becomes very large when applying this approach on large scale datasets like the ImageNet dataset containing more than 14 million images and 21,000 classes. The complexity is both about the time needed to perform each task and the memory and disk usage (eg. 11TB are needed to store the SIFT descriptors computed on the full datasets). Efficient algorithms must be used into these three steps: - obviously, the descriptors computed for one image are independant of the other image ones, so they can be computed in a parallel way, - the quantization step usually uses a k-means algorithm, we have developed different versions of parallel k-means algorithms to use on GPU or a cluster of CPUs, - for the learning task, we have developed a parallel version of LibSVM. The first results on the ten largest classes of ImageNet dataset are promising [55], we have developed a fast and efficient framework for large scale image classification.

6.1.4. Security of media

6.1.4.1. Security of content based image retrieval

Participants: Thanh Toan Do, Ewa Kijak, Laurent Amsaleg, Teddy Furon.

Over the years, the level of maturity reached by content-based retrieval systems (CBRSs) has significantly increased. CBRSs have so far been used in very friendly settings where cultural enrichments are paramount. CBRSs are also used in quite different settings where the control, the surveillance and the filtering of multimedia information are central, such as for copyright enforcement systems. While an abundant literature assesses that today’s CBRSs are robust against general-purpose attacks, we address in this work the security of content-based retrieval systems. Because of our expertise, we focus on security of content-based image retrieval, where images are described by SIFT descriptors and indexed by NV-Tree. We proved in one
preliminary study that a real system fails to match a specifically attacked image and its quasi-copy, breaking its otherwise excellent copyright protection performances. After proposing specific attacks that aim to disturb the descriptor detection stage by both prevent some key-points of being detected and create new ones [75], [74], we pursue the work by considering attacks dedicated to the description computation stage.

6.1.4.2. Estimation of the false alarm probability in watermarking and fingerprinting

Participant: Teddy Furon.

A key issue in watermarking and fingerprinting applications is to satisfy the requirement on the probability of false detection or false accusation. Assume commercial contents are encrypted and watermarked and that future consumer electronics storage devices have a watermark detector. These devices refuse to record a watermarked content since it is copyrighted material. The probability of false alarm is the probability that the detector considers an original piece of content (which has not been watermarked) as protected. The movie that a user shot during his holidays could be rejected by his storage device. This absolutely non user-friendly behavior really scares consumer electronics manufacturers.

In fingerprinting, users’ identifiers are embedded in purchased contents. When this content is found in an illegal place (e.g. a P2P network), the copyright holders decode the hidden message, find an identifier, and thus they can trace the traitor, i.e. the customer who has illegally broadcast his copy. However, the task is not that simple because dishonest users might collude. For security reason, anti-collusion codes have to be employed. Yet, these solutions have a non-zero probability of error (defined as the probability of accusing an innocent). This probability should be, of course, extremely low, but it is also a very sensitive parameter: anti-collusion codes get longer (in terms of the number of bits to be hidden in content) as the probability of error decreases. Fingerprint designers have to strike a trade-off, which is hard to conceive when only rough estimation of the probability of error is known. The major issue for fingerprinting algorithms is the fact that embedding large sequences implies also assessing reliability on a huge amount of data, which may be practically unachievable without using rare event analysis.

In collaboration with the team-projects ASPI and ALEA, we developed a novel strategy for simulating rare events and an associated Monte Carlo estimation of tail probabilities. Our method uses a system of interacting particles and exploits a Feynman-Kac representation of that system to analyze their fluctuations. Our precise analysis of the variance of a standard multilevel splitting algorithm reveals an opportunity for improvement. This leads to a novel method that relies on adaptive levels and produces, in the limit of an idealized version of the algorithm, estimates with optimal variance. Some numerical results show performance close to the idealized version of our technique for these practical applications. This work has been published in the journal Statistics and computing [13]. Algorithms for estimating extreme probabilities and quantiles are implemented as a Matlab package.

6.1.4.3. New decoders for fingerprinting

Participant: Teddy Furon.

So far, the accusation process of a Tardos fingerprinting code is based on single decoders which compute a score per user. Users with the highest score or whose scores is above a threshold are then deemed guilty. In the past years, we have contributed to this approach bringing two improvements; the ‘learn and match’ strategy aims at estimating the collusion process and using the matched score function; a rare event analysis translates this score into a more meaningful probability of being guilty. A fast implementation computes the scores of one million of users within 0.2 second on a regular laptop. Therefore, contrary to common belief, although a single decoder is exhaustive with a linear complexity in $O(n)$, it is not slow.

This fast implementation allows us to propose iterative decoders. A first idea is that conditioning by the identities of some colluders bring more discrimination power to the score function. The first iteration is thus a single decoder, users we are extremely confident to accuse are enrolled as side information. The next iteration computes new scores for the remaining users etc. A second idea is that information theory proves that a joint decoder computing scores for pairs, triplets, or in general $t$-tuples is more powerful than single decoders working with scores for single users. However, nobody did try them for large scale setups since the number of $t$-tuples is in $O(n^t)$. We propose in a first iteration to use a single decoder, to prune out users who are
definitively innocents (because their scores are low) and keeping \(O(\sqrt{n})\) individual suspects. The second iteration is a joint decoding working on pairs of users etc. Iteratively, we prune out enough users such that it is manageable to run a joint decoder on bigger \(t\)-tuples. This work has been done under a collaboration of TEMICS, and published in a series of conference papers [37], [38], [36]. A journal version has been submitted to IEEE Trans. on Information Forensics and Security. A Tardos code software suite (generation of code, collusion attacks, accusation algorithms) is available as a C package.

6.1.4.4. Protocols for fingerprinting

**Participant:** Teddy Furon.

A key assumption of the fingerprinting schemes developed so far is that the colluders may know their own codewords but they ignore the codeword of any other innocent user. Otherwise, the collusion can very easily forge a pirated content framing an innocent user because it contains a sequence close enough to his/her codeword. This puts a lot of pressure on the versioning mechanism which creates the personal copy of the content in accordance to a codeword. For instance, suppose that the versioning is done in the user’s setup box, the unique codeword being loaded into this device at the manufacture. If the code matrix ends up in the hands of an untrustworthy employee, then the whole fingerprinting system is pulled down. This is one argument of the motivation for designing cryptographic protocols for the construction, the versioning and the accusation. We have proposed a new asymmetric fingerprinting protocol dedicated to the state-of-the-art Tardos codes. We believe that this is the first such protocol, and that it is practically efficient. The construction of the fingerprints and their embedding within pieces of content is based on oblivious transfer and do not need a trusted third party. Note, however, that during the accusation stage, a trusted third party, like a Judge, is necessary like in any asymmetric fingerprinting scheme we are aware of. This works has been done in collaboration with the team-project TEMICS, Lab-STICCC Telecom Bretagne and University College London, and presented at Information Hiding [22]. Ana Charpentier defended her PhD. thesis in October 2011 [72].

6.1.4.5. Reconstructing an image from its local descriptors

**Participant:** Hervé Jégou.

We show [47] that an image can be approximately reconstructed based on the output of a black-box local description software such as those classically used for image indexing. Our approach consists first in using an off-the-shelf image database to find patches which are visually similar to each region of interest of the unknown input image, according to associated local descriptors. These patches are then warped into input image domain according to interest region geometry and seamlessly stitched together. Final completion of still missing texture-free regions is obtained by smooth interpolation. As demonstrated in our experiments, visually meaningful reconstructions are obtained just based on image local descriptors like SIFT, provided the geometry of regions of interest is known. The reconstruction allows most often the clear interpretation of the semantic image content. As a result, this work raises critical issues of privacy and rights when local descriptors of photos or videos are given away for indexing and search purpose.

6.2. Multi-dimensional indexing and clustering

6.2.1. Improved NV-tree

**Participant:** Laurent Amsaleg.

_This is a joint work with Björn Pír Jónsson from the School of Computer Science, Reykjavik University, Iceland and with Herwig Lejsek, Videntifier Technologies, Iceland._

We have further improved the NV-Tree (Nearest Vector Tree) indexing techniques. It addresses the specific, yet important, problem of efficiently and effectively finding the approximate \(k\)-nearest neighbors within a collection of a few billion high-dimensional data points. The NV-Tree is a very compact index, as only six bytes are kept in the index for each high-dimensional descriptor. It thus scales extremely well when indexing large collections of high-dimensional descriptors. The NV-Tree efficiently produces results of good quality, even at such a large scale that the indices cannot be kept entirely in main memory any more. We have demonstrated this with extensive experiments using a collection of 2.5 billion SIFT (Scale Invariant Feature Transform) descriptors. Additional experiments involving more than 30 billion SIFT descriptors show results are still of a good quality and that disks are handled as efficiently as they can be.
6.2.2. **Indexation of time series**  
**Participants:** Laurent Amsaleg, Romain Tavenard.

Dynamic Time Warping (DTW) is the most popular approach for evaluating the similarity of time series, but its computation is costly. Therefore, simple functions lower bounding DTW distances have been designed, accelerating searches by quickly pruning sequences that could not possibly be best matches. The tighter the bounds, the more they prune and the better the performance. Designing new functions that are even tighter is difficult because their computation is likely to become complex, canceling the benefits of their pruning. It is possible, however, to design simple functions with a higher pruning power by relaxing the no false dismissal assumption, resulting in approximate lower bound functions. We have discovered how very popular approaches accelerating DTW such as LB_Keogh and LB_PAA can be made more efficient via approximations. The accuracy of approximations can be tuned, ranging from no false dismissal to potential losses when aggressively set for great response time savings. At very large scale, indexing time series is mandatory. These approximate lower bound functions can be used with iSAX. Furthermore, we have also observed that a k-means-based quantization step for iSAX gives significant performance gains.

6.2.3. **Improved image indexing with asymmetric Hamming embedding**  
**Participants:** Patrick Gros, Mihir Jain, Hervé Jégou.

We have proposed [28] an improved asymmetric Hamming Embedding scheme for large scale image search based on local descriptors. The comparison of two descriptors relies on an vector-to-binary code comparison, which limits the quantization error associated with the query compared with the original Hamming Embedding method. The approach is used in combination with an inverted file structure that offers high efficiency, comparable to that of a regular bag-of-features retrieval systems, and consistently improves the search quality over the symmetric version on the two datasets used for the evaluation.

6.2.4. **Compression techniques for nearest neighbor search**  
**Participants:** Laurent Amsaleg, Teddy Furon, Hervé Jégou, Romain Tavenard.

Part of this work on this topic was done in cooperation with Matthijs Douze and Cordelia Schmid (INRIA/LB).  

6.2.4.1. **Re-ranking with source coding**

An extension of our previous work on source coding techniques for high-dimensional indexing has been proposed [29]. The goal is to index a large set of vectors, as large as 1 billion vectors, with limited CPU and memory usage. Based on the product quantization-based indexing technique [18], we show that it is interesting to add an additional level of processing to refine the estimated distances. It consists in quantizing the difference vector between a point and the corresponding centroid. When combined with an inverted file, this gives three levels of quantization. Experiments performed on SIFT and GIST image descriptors show excellent search accuracy outperforming three state-of-the-art approaches. Compared with the original work [18], the proposed re-ranking technique is shown to obtain better trade-off with respect to memory, efficiency and search quality.

6.2.4.2. **Anti-sparse coding for approximate nearest neighbor search**

Following recent works on Hamming Embedding techniques, we propose [67] a binarization method that aim at addressing the problem of nearest neighbor search for the Euclidean metric by mapping the original vectors into binary vectors ones, which are compact in memory, and for which the distance computation is more efficient.

Our method is based on the recent concept of anti-sparse coding, which exhibits here excellent performance for approximate nearest neighbor search. Unlike other binarization schemes, this framework allows, up to a scaling factor, the explicit reconstruction from the binary representation of the original vector. We also show that random projections which are used in Locality Sensitive Hashing algorithms, are significantly outperformed by regular frames for both synthetic and real data if the number of bits exceeds the vector dimensionality, i.e., when high precision is required.
6.2.5. Architecture-aware indexing techniques for solid state disks

Participants: Laurent Amsaleg, Gylfi Gudmundsson.

This is a joint work with Björn Pór Jónsson from the School of Computer Science, Reykjavik University, Iceland.

The scale of multimedia data collections is expanding at a very fast rate. In order to cope with this growth, the high-dimensional indexing methods used for content-based multimedia retrieval must adapt gracefully to secondary storage. Recent progress in storage technology, however, means that algorithm designers must now cope with a spectrum of secondary storage solutions, ranging from traditional magnetic hard drives to state-of-the-art solid state disks. We have analyzed the impact of storage technology on a simple, prototypical high-dimensional indexing method for large scale query processing. We found that while the algorithm implementation deeply impacts the performance of the indexing method, the setup of the underlying storage technology is equally important.

6.3. New techniques for linguistic information acquisition and use

6.3.1. NLP for document description

6.3.1.1. Semantic annotation of multimedia documents based on textual data

Participants: Ali Reza Ebadat, Vincent Claveau, Pascale Sébiliot, Ewa Kijak.

This work is done in the framework of the Quaero project (see below).

On this subject, TEXMEX is implied in three tasks of the Quaero project.

The first task concerns the extraction of terminology from documents. The objective of this work is to study the development and the adaptation of methods to automate the acquisition and the structuring of terminologies. In this context, in 2011, we have undergone a new evaluation of terminology extraction systems. Here again, our system, relying on TermoStat (see previous reports) ranked first for the tracks in which we participated.

We have also continued our work the use of morphology for biomedical terminologies. This approach relies on the decomposition of terms into morphemes and the translation of these morphemes into Japanese (kanji) sub-words. The kanji characters thus offer a semantic way to access the semantics of the morpheme and allow us to detect semantic relations between them. We have tested this approach on more languages and have proved its relevance for information retrieval problems.

The second task aims at extracting semantic and ontological relations from documents. Indeed, detecting semantic and ontological relations in texts is a key to describe a domain and thus manipulate cleverly documents. In 2011, we developed a new relation extraction system based on k-nearest-neighbors and language modeling. It has been tested in the framework of the Quaero evaluation campaign and ranked first or second for all tracks. We have also developed a clustering technique for named entities. It relies on new representation schemes called bag-of-vectors (or bag-of-bags-of-features), which perform better than the classical bag-of-word approach.

The last task directly deals with the semantic annotation of multimedia documents based on textual data, for, very often, many textual or language-related data can be found in multimedia documents or come along such documents. For example, a TV-broadcast, contains speech that can be transcribed, Electronic Program Guide and standard program guide information, closed captions, associated websites, etc. All these sources offers a way to exploit complementary information that can be used to semantically annotate multimedia document. During this year, we finished the development of a football multimedia corpus. It contains the video of several matches, the speech transcripts, associated textual data from specialized websites. All these media have been manually annotated in terms of events, named entities, specialized relations (fouls, replacements, etc) and other relevant information. This corpus will be distributed under LGPL-LR license.

6.3.1.2. Text recognition in videos

Participants: Khaoula Elagouni, Pascale Sébiliot.
This work is done in the context of a joint TeXMEX/Orange Ph.D. thesis supported by a CIFRE grant with Orange Labs.

We aim at helping multimedia content understanding by obtaining benefit from textual clues embedded in digital video data. In 2011, we proposed an Optical Character Recognition-based method to recognize natural scene texts in images, avoiding the conventional character segmentation step. The text image is scanned with multi-scale windows and a robust recognition model is applied on each window, relying on a neural classification approach, to identify non valid characters and recognize valid ones. A graph model is used to represent spatial constraint between recognition results, and to determine the best sequence of characters. Some linguistic knowledge is also incorporated in the graph to remove errors due to recognition confusions. The method was evaluated on the ICDAR 2003 database of scene text images and outperforms state-of-the-art approaches. This work will be presented at DAS2012.

6.3.1.3. DEFT evaluation campaign participation

Participants: Vincent Claveau, Christian Raymond.

Christian Raymond and Vincent Claveau participated to DEFT (http://deft2011.limsi.fr/). Two tasks were proposed: the first one was called "the diachronic variation task" whose objective was to identify the writing year of some OCR newspapers from 1801 to 1944. The second one was a abstract/article pairing task. Their approaches based on boosting and k-nearest neighbors was ranked first on the difficult diachronic task.

6.3.2. Oral and textual information retrieval

6.3.2.1. Graded-inclusion-based Information retrieval systems

Participants: Vincent Claveau, Laurent Ughetto.

Our work on this topic is done in close collaboration with Olivier Pivert from the PILGRIM project-team of IRISA Lannion.

Databases (DB) querying mechanisms, and more particularly the division of relations was at the origin of the Boolean model for Information Retrieval Systems (IRSs). This model has rapidly shown its limitations and is no more used in Information retrieval (IR). Among the reasons, the Boolean approach do not allow to represent and use the relative importance of terms indexing the documents or representing the queries. However, this notion of importance can be captured by the division of fuzzy relations. This division, modeled by fuzzy implications, corresponds to graded inclusions. Theoretical work conducted by the PILGRIM project-team have shown the interest of this operator in IR.

Our first work was to investigate the use of graded inclusions to model the information retrieval process. In this framework, documents and queries are represented by fuzzy sets, which are paired with operations like fuzzy implications and T-norms. Through different experiments, we have shown that only some among the wide range of fuzzy operations are relevant for information retrieval. When appropriate settings are chosen, it is possible to mimic classical systems, thus yielding results rivaling those of state-of-the-art systems. These positive results have validated the proposed approach, while negative ones have given some insights on the properties needed by such a model.

More recently, the links between our fuzzy model and other classical IR models have been studied. It has been shown that our fuzzy implication-based model can be interpreted as several classical models: an Extended Boolean Model, a Logical Model, a Vector Space Model or a Language Model in IR.

6.3.2.2. Information retrieval in TV streams using automatic speech recognition

Participants: Guillaume Gravier, Patrick Gros, Julien Fayolle, Fabienne Moreau, Christian Raymond.
Automatic speech recognition outputs are by nature incomplete and uncertain, so much that lexical indexes of speech are not sufficient to overcome the errors due to out-of-vocabulary words and to most of the named entities, consisting in important semantic information from the discourse. Using if necessary a phonetic index is a solution to retrieve partially the mis-recognized words but at the price of a lower precision because the phonetic representation is also noisy. We proposed this year (still to be submitted) an indexation method which jointly model lexical and phonetic levels with finite-state transducers, offering the possibility to take a lexical path or a phonetic path between two synchronization nodes. The edges are weighted by a vector of features (edition scores, confidence measures, durations) that will be used in a supervised manner to estimate the reliability of the returned result at the search step. The experiments have shown the complementary of lexical-phonetic representations and their contribution for a task of spoken utterance retrieval using named entity queries.

6.4. New processing tools for audiovisual documents

6.4.1. TV stream structuring

6.4.1.1. Repetition detection-based TV structuring

Participants: Vincent Claveau, Guillaume Gravier, Patrick Gros, Emmanuelle Martienne, Abir Ncibi.

We work on the issue of structuring large TV streams. More precisely, we focus on the problem of labeling the segments of a stream according to their types (e.g., programs, commercial breaks, sponsoring, etc). Contrary to existing techniques, we wanted to take into account the sequential aspect of the data, and thus we used Conditional Random Fields (CRF), a classifier which has proved useful to handle sequential data in other domains like computational linguistics or computational biology. During this year, we proved the relevance of CRF in the framework of TV segments labeling. We conducted different experiments, either on manually or automatically segmented streams, with different label granularities, and demonstrated that this approach rivals existing ones. The use of this model for semi-supervised and unsupervised learning are under study.

6.4.2. Program structuring

6.4.2.1. Audiovisual models for event detection in videos

Participants: Guillaume Gravier, Patrick Gros, Cédric Penet.

This work was performed in close collaboration with Technicolor as external partner.

Following our work on the detection of audio concepts related to violence in movie soundtracks [58], we developed a system for the detection of violent scenes in movies, combining multimodal features. We investigated multimodal fusion strategies and temporal integration exploiting Bayesian networks as a joint distribution model. Several strategies for learning the structure of the Bayesian networks were compared, resulting in a complete system for violence detection. The system was evaluated on the Violent Scenes Detection task of the MediaEval 2011 international evaluation [42] that we co-organized with Technicolor and the University of Geneva [62]. A fair amount of time was dedicated this year to the organization of the evaluation campaign which includes defining the task and metrics, supervising the annotation, recruiting participants, analyzing the results and organizing the corresponding workshop session.

6.4.2.2. Unsupervised multimedia content mining

Participants: Guillaume Gravier, Anh Phuong Ta.

This work on audio content discovery was partially carried out in collaboration with Armando Muscariello and Frédéric Bimbot from the Metiss project-team.

As an alternative to supervised approaches for multimedia content analysis, where predefined concepts are searched for in the data, we investigate content discovery approaches where knowledge emerge from the data. Following this general philosophy, we pursued work on motif discovery in audio and video content.
Audio motif discovery is the task of finding out, without any prior knowledge, all pieces of signals that repeat, eventually allowing variability. In 2011, we extended our recent work on seeded discovery to near duplicate detection and spoken document retrieval from examples. First, we proposed algorithmic speed ups for the discovery of near duplicate motifs (low variability) in large (several days long) audio streams, exploiting subsampling strategies [39]. Second, we investigated the use of previously proposed efficient pattern matching techniques to deal with motif variability in speech data [40] in a different setting, that of spoken document retrieval from an audio example. We demonstrated the potential of model-free approaches for efficient spoken document retrieval on a variety of data sets, in particular in the framework of the Spoken Web Search task of the MediaEval 2011 international evaluation [41].

Video structure is often enforced through editing rules which result in a set of shots defining an event that repeats throughout the video with a high visual and audio similarity. Typical such shots are anchor persons and close-up on guests in talk-shows. We recently proposed an unsupervised multimodal approach to discover such events exploiting audio and visual consistency between two sets of independent nested clusters, one for each modality [21]. In 2011, we extended the approach in two directions. First, we improved the selection of consistent audio and visual clusters and the unsupervised selection of positive and negative examples exploiting redundancy between nested clusters. Second, we extended the method to discover several audio-visually consistent events rather than a single one in our previous work, thus enabling the use of unsupervised mining as a pre-processing step for video structure analysis.

6.4.2.3. Topic segmentation with vectorization and morpho-mathematics

**Participant:** Vincent Claveau.

*Our work on this topic is done in close collaboration with Sébastien Lefèvre from the SEASIDE project-team of IRISA Vannes.*

Segmenting a program into topics is an important step for fine-grained structuring of TV streams. Based on our work on vectorization (see previous reports), we have developed a new segmentation technique using speech transcripts. Making an analogy with image segmentation, we have adapted the watershed transform to handle these textual data and more precisely the distances computed by vectorization between possible segments.

This method has been tested on different TV collections (news, reports) as well as more usual text collection used for segmentation evaluation. In every cases, our technique has outperformed any state-of-the-art approaches.

6.4.3. Using speech to describe and structure video

**Participants:** Camille Guinaudeau, Guillaume Gravier, Ludivine Kuznik, Bogdan Ludusan, Pascale Sébillot.

Speech can be used to structure and organize large collections of spoken documents (videos, audio streams, etc) based on semantics. This is typically achieved by first transforming speech into text using automatic speech recognition (ASR), before applying natural language processing (NLP) techniques on the transcripts. Our research focuses firstly on the adaptation of NLP methods designed for regular texts to account for the specific aspects of automatic transcripts. In particular, we investigate a deeper integration between ASR and NLP, i.e., between the transcription phase and the semantic analysis phase.

In 2011, we mostly focused on robust transcription, hierarchical topic segmentation and collection structuring. On the one hand, we investigated the use of broad phonetic landmarks and syllable prominence to improve large vocabulary speech recognition by guiding the Viterbi search process. Several mechanisms to incorporate landmarks into the search space were studied. Significant improvements were observed on radio broadcast news data in the French language. On the other hand, we pursued our work on unsupervised topic adaptation, focusing on the automatic selection of out-of-vocabulary words combining phonetic and morpho-syntactic criteria.
Linear topic segmentation has been widely studied for textual data and recently adapted to spoken contents. However, most documents exhibit a hierarchy of topics which cannot be recovered using linear segmentation. We investigated hierarchical topic segmentation of TV programs exploiting the spoken material. Recursively applying linear segmentation methods is one solution but fails at the lowest levels of the hierarchy when small segments are targeted, in particular when transcription errors jeopardize lexical cohesion. We proposed new probabilistic measures of the lexical cohesion to emphasize the contribution of words that appear only locally, thus attenuating the impact of words which contributed to the segments at an upper level of the hierarchy [11].

Finally, we initiated work in collaboration with INA on structuring a large collection of news reports. The idea is to automatically create links and threads between reports in several months of broadcast news shows, based either on the documentary records of the shows and/or on the automatic transcripts. As preliminary step towards this goal, we investigated distances between documentary records in an information retrieval setting so as to construct a nearest neighbor graph. The next step consists in exploiting graph clustering methods. Our research in speech for TV content structuring was illustrated through the Texmix demonstration (see Section 5.2) which exploits most of our achievements in the field, including transcription, topic segmentation and collection structuring.
6. New Results

6.1. Physical modelling and simulation

6.1.1. Modal analysis for haptic manipulation of deformable models

Participants: Zhaoguang Wang, Georges Dumont [contact].

Real-time interaction between designer and deformable mock-up in VR (Virtual Reality) environment is a natural and promising manner to evaluate designing feasibility. Using finite element method (FEM) for solving this issue leads to high fidelity simulation but to simulation rates that do not meet the requirements (1000Hz) of real time haptic applications. We have proposed a two-stage method based on linear modal analysis. In this method, different modal subspaces, related to use scenarios, are pre-computed offline. These data are then combined online with respect to a simulation division scheme to obtain real time deformations of the parts with respect to the modal response. Two main features are developed in the method. First, we apply an adapted meshing method during the pre-computation process. This method allows to automatically switch between different modal subspaces depending on the interaction region. Second, we divide the real time deformation computation into two separate modules by extracting sub-matrixes from the precomputed modal matrices. This separates the haptic simulation loop from the whole deformation computation an thus preserves the haptic response. This work was presented in WINVR 2011 conference [31] is accepted for publication [8] and was the subject of the PhD Thesis of Zhaoguang Wang, that was defended in june 2011 [3].

6.1.2. Real-time mechanical simulation of brittle fracture

Participants: Loëiz Glondu, Georges Dumont [contact], Maud Marchal [contact].

Simulating brittle fracture of stiff bodies is now commonplace in computer graphics. However, simulating the deformations undergone by the bodies in a realistic way remains computationally expensive. Thus, physically-based simulation of brittle fracture in real-time is still challenging for interactive applications. We are currently working on a new physically-based approach for simulating realistic brittle fracture in real-time. Our method is composed of two main original parts: (1) a fracture initiation model based on modal analysis and a new contact force model and (2) a fracture propagation model based on a novel physically-based algorithm (Figure 2). First results of this method have been published in [32].

Adding physical properties to objects within a virtual world can not generally be handled in real-time during a simulation. For that reason, it is still difficult nowadays to physically simulate fragments of fractured objects or parts of teared/cut objects. We have proposed a method for handling the real-time physical simulations of arbitrary objects that are represented by their surface mesh. Our method is based on a pre-computed shape database in which physical data are stored for a wide variety of objects. When a query object needs to be physically simulated in the virtual world, a similarity search is performed inside the database and the associated physical data are extracted. Our approach proposes to compare three different similarity search methods that fit with our real-time needs. Our results show that our approach has a great potential for the physical simulation of arbitrary objects in interactive applications. These results have been published in the Eurographics International Workshop on Virtual Reality Interaction and Physical Simulation (Vriphys) [21].

6.1.3. Collision detection in large scale environments with High Performance Computing

Participants: Quentin Avril, Valérie Gouranton [contact], Bruno Arnaldi.
Virtual reality environments are becoming increasingly large and complex and real-time interaction level is becoming difficult to stably insure. Indeed, because of their complexity, detailed geometry and specific physical properties, these large scale environments create a critical computational bottleneck on physical algorithms. Our work focused on the first step of the physical process: the collision detection. These algorithms can sometimes have a quadratic complexity. Solving and simplifying the collision detection problem is integral to alleviating this bottleneck. Hardware architectures have undergone extensive changes in the last few years that have opened new ways to relieve this computational bottleneck. Multiple processor cores offer the ability to execute algorithms in parallel on one single processor. At the same time, graphics cards have gone from being a simple graphical display device to a supercomputer. These supercomputers now enjoy attention from a specialized community dealing solely with physical simulation. To perform large scale simulations and remain generic on the runtime architecture, we proposed unified and adaptive mapping models between collision detection algorithms and the runtime architecture using multi-core and multi-GPU architectures. We have developed innovative and effective solutions to significantly reduce the computation time in large scale environments while ensuring the stability and reproducibility of results (cf. Figure 3). We proposed a new pipeline of collision detection with a granularity of parallelism on multicore processors or multi-GPU platforms[11]. It enables simultaneous execution of different stages of the pipeline and a parallel internal to each of these steps. This was the subject of the PhD thesis of Quentin Avril [1].

6.1.4. Assessment of inverse dynamics method for muscle activity analysis

Participants: Georges Dumont [contact], Charles Pontonnier.

The use of virtual reality tools for ergonomics applications is a very important challenge. In order to improve the design of workstations, an estimation of the muscle forces involved in the work tasks has to be done. Several methods can lead to these muscle forces. In this study, we try to assess the level of confidence for results obtained with an inverse dynamics method from real captured work tasks. The chosen tasks are meat cutting tasks, well known to be highly correlated to musculoskeletal troubles appearance in the slaughter industry.

The experimental protocol consists in recording three main data during meat cutting tasks, and analyse their variation when some of the workstation design parameters are changing [26].

1. External (cutting)force data: for this purpose, a 3D instrumented knife has been designed in order
Figure 3. Simulation of moving objects with varying size. Our approach enables to perform the Broad phase step in interactive time using optimized spatial brute force algorithm. (Left: 2,000 Objects - Right: 50,000 Objects.)

to record the force applied by the subject during the task;
2. Motion Capture data: for this purpose, we use a motion capture system with active markers (Visualeyez II, Phoenix Technologies, Canada);
3. EMG data: several muscle activities are recorded using electromyographic electrodes, in order to compare these activities to the ones obtained from the inverse dynamics method.

Then the motion is replayed in the AnyBody modeling system (AnyBody, Aalborg, Denmark) in order to obtain muscle forces generated during the motion. A trend comparison is then done [27], comparing recorded and computed muscle activations. Results show that most of the computed activations are qualitatively close from the recorded ones (similar shapes and peaks), but quantitative comparison leads to major differences between recorded and computed activations (the trend followed by the recorded activations in regard of a workstation design parameter, such as the table height, is not obtained with the computed activations). We currently explore those results to see if the fact that co-contraction of single joints muscles is badly estimated by classical inverse dynamics method can be a reason of this issue. We also work on the co-contraction simulation in order to improve the results [28].

This work has been done in collaboration with the Center for Sensory-motor Interaction (SMI, Aalborg University, Aalborg, Denmark), particularly Mark de Zee (Associate Professor) and Pascal Madeleine (Professor). Charles Pontonnier spent a 9 months post-doctoral fellowship at SMI from December 2010 to August 2011.

6.2. Multimodal immersive interaction

6.2.1. Brain-Computer Interaction based mental state

Participants: Anatole Lécuyer [contact], Bruno Arnaldi, Laurent George, Yann Renard.

In [20], presented at IEEE EMBS conference, we have explored the use of electrical biosignals measured on scalp and corresponding to mental relaxation and concentration tasks in order to control an object in a video game as illustrated in Figure 4. To evaluate the requirements of such a system in terms of sensors and signal processing we compared two designs. The first one used only one scalp electroencephalographic (EEG) electrode and the power in the alpha frequency band. The second one used sixteen scalp EEG electrodes and machine learning methods. The role of muscular activity was also evaluated using five electrodes positioned on the face and the neck.
Results show that the first design enabled 70% of the participants to successfully control the game, whereas 100% of the participants managed to do it with the second design based on machine learning. Subjective questionnaires confirm these results: users globally felt to have control in both designs, with an increased feeling of control in the second one. Offline analysis of face and neck muscle activity shows that this activity could also be used to distinguish between relaxation and concentration tasks. Results suggest that the combination of muscular and brain activity could improve performance of this kind of system. They also suggest that muscular activity has probably been recorded by EEG electrodes.

In [19], presented in the 5th International Brain-Computer Interface Conference, we introduce the concept of Brain-Computer Interface (BCI) inhibitor, which is meant to standby the BCI until the user is ready, in order to improve the overall performance and usability of the system. BCI inhibitor can be defined as a system that monitors user’s state and inhibits BCI interaction until specific requirements (e.g. brain activity pattern, user attention level) are met.

We conducted a pilot study to evaluate a hybrid BCI composed of a classic synchronous BCI system based on motor imagery and a BCI inhibitor (Figure 5). The BCI inhibitor initiates the control period of the BCI when requirements in terms of brain activity are reached (i.e. stability in the beta band).

Preliminary results with four participants suggest that BCI inhibitor system can improve BCI performance.

6.2.2. Navigating in virtual worlds using a Brain-Computer Interface

Participants: Anatole Lécuyer [contact], Jozef Legény.

When a person looks at a light flickering at a constant frequency, we can observe a corresponding electrical signal in their EEG. This phenomenon, located in the occipital area of the brain is called Steady-State Visual-Evoked Potential (SSVEP).
In [7] we introduce a novel paradigm for a controller using SSVEP. Compared to the state-of-the-art implementations which use static flickering targets, we have used animated and moving objects. In our example applications we have used animated butterflies flying in front of the user as shown in Figure 6. A study has revealed that, at the cost of decreased performance, this controller increases the personal feeling of presence.

These results show that integrating visual SSVEP stimulation into the environment is possible and that further study is necessary in order to improve the performance of the system.

6.2.3. Walking-in-place in virtual environments

Participants: Anatole Lécuyer [contact], Maud Marchal [contact], Léo Terziman, Bruno Arnaldi, Franck Multon.

The Walking-In-Place interaction technique was introduced to navigate infinitely in 3D virtual worlds by walking in place in the real world. The technique has been initially developed for users standing in immersive setups and was built upon sophisticated visual displays and tracking equipments. We have proposed to revisit the whole pipeline of the Walking-In-Place technique to match a larger set of configurations and apply it notably to the context of desktop Virtual Reality. With our novel "Shake-Your-Head" technique, the user has the possibility to sit down, and to use small screens and standard input devices for tracking. The locomotion simulation can compute various motions such as turning, jumping and crawling, using as sole input the head movements of the user (Figure 7).

In a second study [29] we analyzed and compared the trajectories made in a Virtual Environment with two different navigation techniques. The first is a standard joystick technique and the second is the Walking-In-Place (WIP) technique. We proposed a spatial and temporal analysis of the trajectories produced with both techniques during a virtual slalom task. We found that trajectories and users’ behaviors are very different across the two conditions. Our results notably showed that with the WIP technique the users turned more often and navigated more sequentially, i.e. waited to cross obstacles before changing their direction. However, the users were also able to modulate their speed more precisely with the WIP. These results could be used to optimize the design and future implementations of WIP techniques. Our analysis could also become the basis of a future framework to compare other navigation techniques.

6.2.4. Improved interactive stereoscopic rendering: SCVC

Participants: Jérôme Ardouin, Anatole Lécuyer [contact], Maud Marchal [contact], Eric Marchand.
Frame cancellation comes from the conflict between two depth cues: stereo disparity and occlusion with the screen border. When this conflict occurs, the user suffers from poor depth perception of the scene. It also leads to uncomfortable viewing and eyestrain due to problems in fusing left and right images.

In [10], presented at the IEEE 3DUI conference, we propose a novel method to avoid frame cancellation in real-time stereoscopic rendering. To solve the disparity/frame occlusion conflict, we propose rendering only the part of the viewing volume that is free of conflict by using clipping methods available in standard real-time 3D APIs. This volume is called the Stereo Compatible Volume (SCV) and the method is named Stereo Compatible Volume Clipping (SCVC).

Black Bands, a proven method initially designed for stereoscopic movies is also implemented to conduct an evaluation. Twenty two people were asked to answer open questions and to score criteria for SCVC, Black Bands and a Control method with no specific treatment. Results show that subjective preference and user’s depth perception near screen edge seem improved by SCVC, and that Black Bands did not achieve the performance we expected.

At a time when stereoscopic capable hardware is available from the mass consumer market, the disparity/frame occlusion conflict in stereoscopic rendering will become more noticeable. SCVC could be a solution to recommend. SCVC’s simplicity of implementation makes the method able to target a wide range of rendering software from VR application to game engine.

6.2.5. Six degrees-of-freedom haptic interaction

Participants: Anatole Lécuyer [contact], Maud Marchal [contact], Gabriel Cirio.

Haptic interaction with virtual objects is a major concern in the virtual reality field. There are many physically-based efficient models that enable the simulation of a specific type of media, e.g. fluid volumes, deformable and rigid bodies. However, combining these often heterogeneous algorithms in the same virtual world in order to simulate and interact with different types of media can be a complex task.

In [5], published at IEEE Transactions on visualization and Computer Graphics, we propose a novel approach that allows real-time 6 Degrees of Freedom (DoF) haptic interaction with fluids of variable viscosity. Our haptic rendering technique, based on a Smoothed-Particle Hydrodynamics (SPH) physical model, provides a realistic haptic feedback through physically-based forces. 6DoF haptic interaction with fluids is made possible thanks to a new coupling scheme and a unified particle model, allowing the use of arbitrary-shaped rigid bodies. Particularly, fluid containers can be created to hold fluid and hence transmit to the user force feedback coming from fluid stirring, pouring, shaking or scooping. We evaluate and illustrate the main features of our approach through different scenarios, highlighting the 6DoF haptic feedback and the use of containers.
The Virtual Crepe Factory [14] illustrates this approach for 6DoF haptic interaction with fluids. It showcases a 2-handed interactive haptic scenario: a recipe consisting in using different types of fluid in order to make a special pancake also known as "crepe". The scenario (Figure 8) guides the user through all the steps required to prepare a crepe: from the stirring and pouring of the dough to the spreading of different toppings, without forgetting the challenging flipping of the crepe. With the Virtual Crepe Factory, users can experience for the first time 6DoF haptic interactions with fluids of varying viscosity.

Figure 8. A complete use-case of our approach: a virtual crepe preparation simulator. The user manipulates a bowl (left hand, left haptic device) and a pan (right hand, right haptic device).

In [15], presented at the IEEE Virtual Reality Conference, we propose the first haptic rendering technique for the simulation and the interaction with multistate (Figure 9) media, namely fluids, deformable bodies and rigid bodies, in real-time and with 6DoF haptic feedback. The shared physical model (SPH) for all three types of media avoids the complexity of dealing with different algorithms and their coupling. We achieve high update rates while simulating a physically-based virtual world governed by fluid and elasticity theories, and show how to render interaction forces and torques through a 6DoF haptic device.

Figure 9. 6DoF haptic interaction in a medical scenario. Fluid blood pours from the deformable intestine when the user penetrates it with the rigid probe.

6.2.6. Joyman: a human-scale joystick for navigating in virtual worlds

Participants: Maud Marchal [contact], Anatole Lécuyer, Julien Pettré.
We have proposed a novel interface called Joyman (Figure 10), designed for immersive locomotion in virtual environments. Whereas many previous interfaces preserve or stimulate the users' proprioception, the Joyman aims at preserving equilibrioception in order to improve the feeling of immersion during virtual locomotion tasks. The proposed interface is based on the metaphor of a human-scale joystick. The device has a simple mechanical design that allows a user to indicate his virtual navigation intentions by leaning accordingly. We have also proposed a control law inspired by the biomechanics of the human locomotion to transform the measured leaning angle into a walking direction and speed - i.e., a virtual velocity vector. A preliminary evaluation was conducted in order to evaluate the advantages and drawbacks of the proposed interface and to better draw the future expectations of such a device.

This principle of this new interface was published at international conference IEEE 3DUI [25] and a patent has been filed for the interface. A demonstration of this interface was proposed at ACM Siggraph Asia Emerging Technologies [33].

![Figure 10. Prototype of the "Joyman"

6.2.7. Interactions within 3D virtual universes

**Participants:** Thierry Duval [contact], Valérie Gouranton [contact], Bruno Arnaldi, Laurent Aguerreche, Cédric Fleury, Thi Thuong Huyen Nguyen.

Our work focuses upon new formalisms for 3D interactions in virtual environments, to define what an interactive object is, what an interaction tool is, and how these two kinds of objects can communicate together. We also propose virtual reality patterns to combine navigation with interaction in immersive virtual environments.

We have worked upon generic interaction tools for collaboration, based on multi-point interaction. In that context we have studied the efficiency of one instance of our Reconfigurable Tangible Device, the RTD-3, for collaborative manipulation of 3D objects compared to state of the art metaphors [9]. We have setup an experiment for collaborative distant co-manipulation (figure 1) of a clipping plane inside for remotely analyzing 3D scientific data issued from an earthquake simulation.

6.3. Collaborative work in CVE’s

6.3.1. The immersive interactive virtual cabin (IIVC)

**Participants:** Thierry Duval [contact], Valérie Gouranton [contact], Alain Chauffaut, Bruno Arnaldi, Cédric Fleury.

We are still improving the architecture of our Immersive Interactive Virtual Cabin to improve the user’s immersion with all his real tools and so to make the design and the use of 3D interaction techniques easier, and to make possible to use them in various contexts, either for different kinds of applications, or with different kinds of physical input devices.
The IIVC is now fully implemented in our two VR platforms: OpenMASK 5.1 and Collaviz 7.1.3.

6.3.2. Generic architecture for 3D interoperability

**Participants:** Thierry Duval [contact], Valérie Gouranton, Cédric Fleury, Rozenn Bouville Berthelot, Bruno Arnaldi.

Our goal is to allow software developers to build 3D interactive and collaborative environments without bothering with the 3D graphics API they are using. This work is the achievement of the IIVC software architecture. We have proposed PAC-C3D (Figure 11), a new software architectural model for collaborative 3D applications, in order to provide a higher abstraction for designing 3D virtual objects, and in order to provide interoperability, making it possible to share a virtual universe between heterogeneous 3D viewers [17],[16].

![Figure 11. The PAC-C3D software architectural model makes interoperability possible between heterogeneous 3D viewers](image)

We also study how to offer interoperability between virtual objects that are loaded in the same virtual environment but that are described using different formats. This is why we have proposed a generic architecture for enabling interoperability between 3D formats (Figure 12), the Scene Graph Adapter [12]. Our SGA is now able to allow events coming from a 3D format to act upon data provided in another format, such as X3D events operating on Collada data [4].

6.4. Immersia Virtual Reality room

**Participants:** Georges Dumont [contact], Alain Chauffaut [contact], Ronan Gaugne [contact], Rémi Félix, Marwan Badawi, Bruno Arnaldi, Thierry Duval, Valérie Gouranton.

The team was the first in France to host a large-scale immersive virtual reality equipment known as Immersia. This platform, with full visual and sound immersion, is dedicated to real-time, multimodal (vision, sound, haptic, BCI) and immersive interaction. The Immersia platform is a key node of the European transnational VISIONAIR infrastructure and will be open in 2012 to the access of foreign research projects. It will accommodate experiments using interactive and collaborative virtual-reality applications that have multiple local or remote users. Our new wall has four faces: a front, two sides and a ground. Dimensions are 9.6 m wide, 2.9 m deep and 3.1 m height. The visual reproduction system combines eight Barco Galaxy NW12 projectors and three Barco Galaxy 7+ projectors. Visual images from Barco projectors are rendered on glass screens. They are adjusted for the user’s position, and this, together with their high resolution and homogeneous coloring, make them very realistic. The ART localization system, constituted of 16 ART-track2 cameras, enables real
Figure 12. Our architecture allows the loading of any 3D graphics format simultaneously in any available rendering engine. The scene graph adapter is an interface that adapts a scene graph (SG) of a given format into a renderer scene graph and which also allows the rendering part to request this scene graph.

objects to be located within the U-shape. Sound rendering is provided by a Yamaha processor, linked either to Genelec speakers with 10.2 format sound or Beyer Dynamic headsets with 5.1 virtual format sound, controlled by the user’s position.
6. New Results

6.1. Multimedia Models and Formats

6.1.1. SMIL timesheets

With the advent of HTML5 and its support in most popular browsers, HTML is becoming an important multimedia language. Video and audio can now be embedded in HTML pages without worrying about the availability of plugins. However, a major issue is to specify the dynamic behavior of documents (user interactions, timing and synchronization with continuous contents). This is done usually by writing (often complex) scripts, which require programming skills from the authors.

To address this issue, we have created the timesheets.js library, a scheduler that allows HTML documents to be animated and synchronized in a purely declarative way. This work is based on the SMIL Timing and Synchronization module and the SMIL Timesheets specification, with a few extensions.

The library is implemented in JavaScript, which makes it usable in any browser. Authors can specify the dynamic behavior of HTML5 (+CSS3) documents. They can thus develop multimedia applications without writing a single line of JavaScript. Timesheets can also be used with other XML document languages, such as SVG for instance. This approach was validated in a class with students learning web multimedia.

6.1.2. Multimedia content adaptation

Multimedia documents may have to be played on multiple devices such as mobile phones, tablets, desktop computers, set-top boxes, etc. Usage and platform diversity requires documents to be adapted according to execution contexts, sometimes unpredictable at design time. In a joint work with project-team Exmo, we have designed a semantic adaptation framework for multimedia documents. This framework captures the semantics of document composition and transforms the relations between media objects according to adaptation constraints.

6.2. XML Processing

Mature results about XML processing were obtained along three main directions: the formalization and implementation for checking the impact of schema evolution on validation and queries; logical extensions supporting a notion of counting and the shuffle operator in trees; and the decision of a subtyping relation for a very expressive type algebra supporting a notion of polymorphism.

In addition, preliminary results were obtained on the definition of a rigorous logical framework for the static analysis of semantic web languages, on the static analysis of cascading style sheets, and on the equipment of an IDE with new static analysis features for XQuery.

6.2.1. Impact of XML schema evolution

In the ever-changing context of the web, XML schemas continuously change in order to cope with the natural evolution of entities they describe. Schema changes have important consequences. First, existing documents valid with respect to the original schema are no longer guaranteed to fulfill the constraints described by the evolved schema. Second, the evolution also impacts programs manipulating documents whose structure is described by the original schema.

We have proposed a unifying framework for determining the effects of XML Schema evolution both on the validity of documents and on queries [2]. The system is very powerful in analyzing various scenarios in which forward/backward compatibility of schemas is broken, and in which the result of a query may not be anymore what was expected. Specifically, the system offers a predicate language which allows one to formulate properties related to schema evolution. The system then relies on exact reasoning techniques to perform a fine-grained analysis. This yields either a formal proof of the property or a counter-example that can be used for debugging purposes. The system has been fully implemented and tested with real-world use cases, in particular with the main standard document formats used on the web, as defined by W3C. The system identifies precisely compatibility relations between document formats. In case these relations do not hold, the system can identify queries that must be reformulated in order to produce the expected results across successive schema versions.

6.2.2. Counting in trees

A major challenge of query language design is the combination of expressivity with effective static analyses such as query containment. In the setting of XML, documents are seen as finite trees, whose structure may additionally be constrained by type constraints such as those described by an XML schema. We have considered the problem of query containment in the presence of type constraints for a class of regular path queries extended with counting and interleaving operators [1]. The counting operator restricts the number of occurrences of children nodes satisfying a given logical property. The interleaving operator provides a succinct notation for describing the absence of order between nodes satisfying a logical property. We have proposed a logic supporting these operators, which can be used to solve common query reasoning problems such as satisfiability and containment of queries in exponential time [4].

6.2.3. Typing higher-order programs

We have considered a type algebra equipped with recursive, product, function, intersection, union, and complement types together with type variables and universal quantification over them. We have defined the subtyping relation between such type expressions, and have proved its decidability[9].

This has solved an open problem that was attracting a considerable research effort. The novelty, originality and strength of our solution reside in introducing a logical modeling for the semantic subtyping framework. We have modeled semantic subtyping in a tree logic and use a satisfiability-testing algorithm in order to decide subtyping. We have shown how the subtyping relation can be decided in EXPTIME. We have reported on practical experiments made with a full implementation of the system. This has provided a powerful polymorphic type system aiming at maintaining full static type-safety of functional programs that manipulate trees, even with higher-order functions, which is particularly useful in the context of XML.

6.2.4. Detection of inconsistent paths and dead code in XML IDEs

One of the challenges in web software development is to help achieving a good level of quality in terms of code size and runtime performance, for increasingly popular domain specific languages such as XQuery. We have presented an IDE equipped with static analysis features for assisting the programmer [8]. These features are capable of identifying and eliminating dead code automatically. The tool is based on newly developed formal programming language verification techniques, which are now mature enough to be introduced in the process of software development.

6.2.5. Static analysis of semantic web languages

We work with the Exmo project-team on the static analysis of semantic web languages such as RDF, OWL and SPARQL by investigating modal logics over graphs. We seek to build a rigorous logical reasoning framework based on $\mu$-calculus adapted for the web semantic languages [7] [11]. In particular, we studied the containment problem for SPARQL queries: determining whether, for any graph, the answers to a query are contained in those of another query. Our approach consists in encoding RDF graphs as transition systems and queries as $\mu$-calculus formulas and then reducing the containment problem to testing satisfiability in the logic.
6.2.6. Static analysis of style sheets

Developing and maintaining cascading style sheets (CSS) is an important issue to web developers as they suffer from the lack of rigorous methods. Most existing means rely on validators that check syntactic rules, and on runtime debuggers that check the behavior of a CSS style sheet on a particular document instance. However, the aim of most style sheets is to be applied to an entire set of documents, usually defined by some schema. To this end, a CSS style sheet is usually written w.r.t. a given schema. While usual debugging tools help reducing the number of bugs, they do not ultimately allow to prove properties over the whole set of documents to which the style sheet is intended to be applied.

We have proposed a novel approach to fill this lack \cite{14} by analyzing CSS style sheets using the same logic and compile-time verification technique we use for other XML problems. We have developed an original tool based on our XML Reasoning Solver (see section \ref{5.2}). The tool is capable of statically detecting a wide range of errors (such as empty CSS selectors and semantically equivalent selectors), as well as proving properties related to sets of documents (such as coverage of styling information), in the presence or absence of schema information. This new tool can be used in addition to existing runtime debuggers to ensure a higher level of quality of CSS style sheets.

6.3. Multimedia Authoring

6.3.1. C2M project

The C2M project (see section \ref{7.2.2}) aims at developing industrial solutions that allow multimedia developers to achieve mass production with high quality results. It uses the SCENARI platform for document production and we have proposed a solution for dealing with multimedia content in such a framework \cite{16}. Indeed, automatic tools are not always sufficient for generating high quality documents; manual editing of documents in their publishing format is often necessary to tune a number of details.

Our approach consists in providing a post-editing service to allow authors to adjust their multimedia presentations directly on the final form of documents. The first step is to provide a web rendering engine based on the latest advances in web standards, as described in section \ref{6.1.1}. The second step consists in designing web-aware authoring tools based on this library, thus providing authors with direct editing services for producing high quality multimedia documents while preserving the advantages of using an XML production workflow. We have developed a prototype of this authoring tool in which all editing templates are described with XUL (XML User interface Language) and XBL (XML Binding Language) elements that we have defined for handling time-based content and widgets (timeContainer, timeNode, timeLine, etc.).

With such a solution, we can combine two worlds: a semantic-oriented authoring approach, as provided by an XML document workflow, and a direct web-based editing system. The first guarantees homogeneous rendering while the latter enables direct adjustments on the final form of the document. Bridging these two worlds is made possible by using declarative web languages (namely HTML5, SMIL and CSS) and implementing their timing part in the browser (with the Timesheets.js library, see section \ref{6.1.1}). The authoring components are directly mapped to the document structures.

6.3.2. On-line editing of multimedia web content

In cooperation with EPFL (Lausanne) we have continued to explore the concept of template-driven editing for XML multimedia contents (see section \ref{3.3.2}). This year, we have carried out more experiments with very different types of contents, including structured documents, factual data, and multimedia objects \cite{17}.

These experiments have been done with the AXEL library developed by EPFL, based on our joint work on template languages. AXEL is an innovative client-side authoring tool that runs in the browser for editing XML documents, driven by an XTiger template. It allows average web users to easily edit XML content on web servers with no specific knowledge of XML. Our experiments have shown that the template-based approach significantly enhances the ability for web users to directly feed various applications with structured content.
6.4. Augmented Environments

A large part of the research on augmented environments specializes in the use of visual media. In WAM, we focus on the use of audio media and we put a strong emphasis on mobility.

We have developed the first indoor augmented reality audio navigation application running on personal AR devices such as mobile phones. The main idea behind the development of this application was a joint use of three concepts:

- Continuous localization by using embedded sensors together with physiological models of walking and assumptions about walking in structured indoor environments.
- OpenStreetMap Indoor Mapping used for map-aided positioning, assistive routing for visually impaired people, and environmental queries through audio panoramics.
- Guidance and navigation through AR audio, both 3D and environmental, with mixing of synthetic and natural sounds and support for timely audio information presentation.

We have demonstrated that these concepts are inter-dependent [12], and that bringing them together is a way to find new solutions to problems which are difficult to tackle when looking at them separately. These three concepts are implemented with web technologies we use XML languages and XML tools for interactive audio, building modeling, and personal navigation module configuration. This enables easy authoring of sound objects or audio icons used for building sonification [15], indoor navigation maps and panoramics, and walking models. Adaptability of navigation to preferences of people is based on the concept of audio stylesheets for OpenStreetMap data rendering, XML configuration of the Pedestrian Dead Reckoning module, and assistive routing specification.

We have developed two mobile browsers and a framework for generic navigation:

1. The Mixed Reality Browser [10] that we have developed can display PoI content either remotely through panoramics with spatialized audio, or on-site by walking to the corresponding place. MRB is the only browser of geolocalized data to use a declarative XML format for PoIs, panoramics, 3D audio and to be based on HTML5 both for the iconic and full information content of PoIs. MRB can be used for any kind of augmented reality visits. A cultural heritage visit of Grenoble (see section 7.1.2) has been realized with the tourist office of Grenoble and the CCSTI (Centre de Culture Scientifique Technique et Industrielle de Grenoble).

2. The Pedestrian Way Browser that we have developed can be used for indoor-outdoor navigation with assistive audio technology for visually impaired people on pedestrian ways with precise geospatial description. Its main characteristic is to be based entirely on the OpenStreetMap XML format for the representation of the route. We anticipate that in the context of the european project Venturi (see section 7.3.1), we will have a convergence of the MRB and PWB, allowing visually impaired people to undertake cultural heritage visit. An demonstration showing the use of the PWB in a structured outdoor environment is available online: http://www.youtube.com/watch?v=h2b8yfCauZ8

3. We have created an extensible client-server framework named TARA which allows navigation on an OpenStreetMap XML graph (indoor and outdoor) by computing routes in real-time. User preferences, like stairs versus lift, are supported through a ponderation of paths in the routing algorithm. The client is an HTML5 running in the browser on mobiles and desktop computers. It can therefore be used for simulation, to test or learn a route before the navigation in the real world. The user interface is based on three modalities, touch, audio and visual and can be operated by visually impaired people through VoiceOver using only touch and audio. Localization through embedded or external sensors is not mandatory as step by step instructions can be accessed through touch modality. The server is a full REST server (Sling-Apache) giving priority to the representation of geospatial resources and allowing environmental queries through the use of XQuery.
6. New Results

6.1. 3D object and scene modeling, analysis, and retrieval

6.1.1. Quantitative image analysis for archeology

Participants: Bryan Russell, Jean Ponce, Josef Sivic, Helene Dessales [ENS Archeology laboratory].

Figure 1. (a) Example photographs captured of the Pompeii site (563 photographs are used in total), (b) Rendered viewpoints of the recovered 3D model. Notice the fine-level details that are captured by the model.

Accurate indexing and alignment of images is an important problem in computer vision. A successful system would allow a user to retrieve images with similar content to a query image, along with any information associated with the image. Prior work has mostly focused on techniques to index and match photographs depicting particular instances of objects or scenes (e.g. famous landmarks, commercial product labels, etc.). This has allowed progress on tasks, such as the recovery of a 3D reconstruction of the depicted scene. However, there are many types of images that cannot be accurately aligned. For instance, for many locations there are drawings and paintings made by artists that depict the scene. Matching and aligning photographs, paintings, and drawings is extremely difficult due to various distortions that can arise. Examples include perspective and caricature distortions, along with errors that arise due to the difficulty of drawing a scene by hand.

In this project, we seek to index and align a database of images, paintings, and drawings. The focus of our work is the Championnet house in the Roman ruins at Pompeii, Italy. Given an alignment of the images, paintings, and drawings, we wish to explore tasks that are of interest to archaeologists and curators who wish to study and preserve the site. Example applications include: (i) digitally restoring paintings on walls where the paintings have disappeared over time due to erosion, (ii) geometrically reasoning about the site over time through the drawings, (iii) indexing and searching patterns that exist throughout the site.
Figure 2. Final alignment between the paintings and 3D model. For each example, left: painting; middle: 3D model contours projected onto painting; right: synthesized viewpoint from 3D model using recovered camera parameters. For the examples in (a-c), note how the final alignment is close to the painting. Our system handles paintings that depict the 3D structure of the scene over time and span different artistic styles and mediums (e.g. water colors, cross-hatching, copies of originals on engravings). Notice how the site changes over time, with significant structural changes (e.g. the wall murals decay over time, the columns change). Example failure cases are shown in (d,e).
Recently, we have addressed the problem of automatically aligning historical architectural paintings with 3D models obtained using multi-view stereo technology from modern photographs. This is a challenging task because of the variations in appearance, geometry, color and texture due to environmental changes over time, the nonphotorealistic nature of architectural paintings, and differences in the viewpoints used by the painters and photographers. Our alignment procedure consists of two novel aspects: (i) we combine the gist descriptor with the view-synthesis/retrieval of Irshara et al. to obtain a coarse alignment of the painting to the 3D model, and (ii) we have developed an ICP-like viewpoint refinement procedure, where 3D surface orientation discontinuities (folds and creases) and view-dependent occlusion boundaries are rendered from the automatically obtained and noisy 3D model in a view-dependent manner and matched to gPB contours extracted from the paintings. We demonstrate the alignment of XIXth Century architectural watercolors of the Casa di Championnet in Pompeii with a 3D model constructed from modern photographs using the PMVS public-domain multi-view stereo software. Figure 1 shows some of the captured photographs and snapshots of the 3D reconstruction of the site. Notice that the 3D reconstruction captures much detail of the walls and structures. Example painting to 3D model alignments are shown in figure 2.

This work resulted in a workshop publication [16].

6.1.2. Visual localization by linear combination of image descriptors

Participants: Josef Sivic, Akihiko Torii [Tokyo Institute of Technology], Tomas Pajdla [CTU in Prague].

In this work, we seek to predict the GPS location of a query image given a database of images localized on a map with known GPS locations. The contributions of this work are three-fold: (1) we formulate the image-based localization problem as a regression on an image graph with images as nodes and edges connecting close-by images; (2) we design a novel image matching procedure, which computes similarity between the query and pairs of database images using edges of the graph and considering linear combinations of their feature vectors. This improves generalization to unseen viewpoints and illumination conditions, while reducing the database size; (3) we demonstrate that the query location can be predicted by interpolating locations of matched images in the graph without the costly estimation of multi-view geometry. We demonstrate benefits of the proposed image matching scheme on the standard Oxford building benchmark, and show localization results on a database of 8,999 panoramic Google Street View images of Pittsburgh.

This work resulted in a publication [18].

6.2. Category-level object and scene recognition

6.2.1. Task-Driven Dictionary Learning

Participants: Julien Mairal, Jean Ponce, Francis Bach [INRIA SIERRA].

Modeling data with linear combinations of a few elements from a learned dictionary has been the focus of much recent research in machine learning, neuroscience and signal processing. For signals such as natural images that admit such sparse representations, it is now well established that these models are well suited to restoration tasks. In this context, learning the dictionary amounts to solving a large-scale matrix factorization problem, which can be done efficiently with classical optimization tools. The same approach has also been used for learning features from data for other purposes, e.g., image classification, but tuning the dictionary in a supervised way for these tasks has proven to be more difficult. In this paper, we present a general formulation for supervised dictionary learning adapted to a wide variety of tasks, and present an efficient algorithm for solving the corresponding optimization problem. Experiments on handwritten digit classification, digital art identification, nonlinear inverse image problems, and compressed sensing demonstrate that our approach is effective in large-scale settings, and is well suited to supervised and semi-supervised classification, as well as regression tasks for data that admit sparse representations.

This work has resulted in a publication [4].

6.2.2. Ask the locals: multi-way local pooling for image recognition

Participants: Y-Lan Boureau, Jean Ponce, Nicolas Le Roux [INRIA SIERRA], Francis Bach [INRIA SIERRA], Yann LeCun [New York University].
Invariant representations in object recognition systems are generally obtained by pooling feature vectors over spatially local neighborhoods. But pooling is not local in the feature vector space, so that widely dissimilar features may be pooled together if they are in nearby locations. Recent approaches rely on sophisticated encoding methods and more specialized codebooks (or dictionaries), e.g., learned on subsets of descriptors which are close in feature space, to circumvent this problem. In this work, we argue that a common trait found in much recent work in image recognition or retrieval is that it leverages locality in feature space on top of purely spatial locality. We propose to apply this idea in its simplest form to an object recognition system based on the spatial pyramid framework, to increase the performance of small dictionaries with very little added engineering. State-of-the-art results on several object recognition benchmarks show the promise of this approach.

This work has resulted in a publication [7].

6.2.3. A Graph-matching Kernel for Object Categorization

**Participants:** Olivier Duchenne, Armand Joulin, Jean Ponce.

This paper addresses the problem of category-level image classification. The underlying image model is a graph whose nodes correspond to a dense set of regions, and edges reflect the underlying grid structure of the image and act as springs to guarantee the geometric consistency of nearby regions during matching. A fast approximate algorithm for matching the graphs associated with two images is presented. This algorithm is used to construct a kernel appropriate for SVM-based image classification, and experiments with the Caltech 101, Caltech 256, and Scenes datasets demonstrate performance that matches or exceeds the state of the art for methods using a single type of features.

This work has resulted in an ICCV 2011 publication [9] (oral presentation).

6.2.4. A Tensor-Based Algorithm for High-Order Graph Matching

**Participants:** Olivier Duchenne, Jean Ponce, Francis Bach [INRIA SIERRA], Inso Kweon [KAIST, Korea].

This paper addresses the problem of establishing correspondences between two sets of visual features using higher-order constraints instead of the unary or pairwise ones used in classical methods. Concretely, the corresponding hypergraph matching problem is formulated as the maximization of a multilinear objective function over all permutations of the features. This function is defined by a tensor representing the affinity between feature tuples. It is maximized using a generalization of spectral techniques where a relaxed problem is first solved by a multi-dimensional power method, and the solution is then projected onto the closest assignment matrix. The proposed approach has been implemented, and it is compared to state-of-the-art algorithms on both synthetic and real data.

This work has resulted in a PAMI publication [2].

6.2.5. Clusterpath: an algorithm for clustering using convex fusion penalties

**Participants:** Armand Joulin, Toby Hocking [INRIA SIERRA], Francis Bach [INRIA SIERRA], Jean-Philippe Vert [Mines ParisTech].

We present a new clustering algorithm by proposing a convex relaxation of hierarchical clustering, which results in a family of objective functions with a natural geometric interpretation. We give efficient algorithms for calculating the continuous regularization path of solutions, and discuss relative advantages of the parameters. Our method experimentally gives state-of-the-art results similar to spectral clustering for non-convex clusters, and has the added benefit of learning a tree structure from the data.

This work has resulted in an publication [10].

6.2.6. An MRF model for binarization of natural scene text

**Participants:** Karteek Alahari, Anand Mishra [IIT India], C.V. Jawahar [IIT India].
Scene text recognition has gained significant attention from the computer vision community in recent years. Recognizing text in the wild is a challenging problem, even more so than the recognition of scanned documents. In this work, we focus on the problem of cropped word recognition. We present a framework that exploits both bottom-up and top-down cues. The bottom-up cues are derived from individual character detections from the image. We build a Conditional Random Field model on these detections to jointly model the strength of the detections and the interactions between them. We impose top-down cues obtained from a lexicon-based prior, i.e. language statistics, on the model. The optimal word represented by the text image is obtained by minimizing the energy function corresponding to the random field model.

We show very significant improvements in accuracies on two challenging public datasets, namely Street View Text (over 15%) and ICDAR 2003 (over 10%).

This work has resulted in an publication [12].

6.2.7. Strongly-supervised deformable part model for object detection
Participants: Hossein Azizpour [KTH Stockholm], Ivan Laptev, Stefan Carlsson [KTH Stockholm].

Deformable part models achieve state-of-the-art performance for object detection while relying on the greedy initialization during training. The goal of this paper is to investigate limitations of such initialization and to improve the model for the case when part locations are known at the training time. To this end, we deploy part-level supervision and demonstrate improved detection results when learning models with manually-initialized part locations. We further explore the benefits of the strong supervision and learn model structure by minimizing the variance among adjacent model parts. Our method can simultaneously handle samples with and without part-level annotation making benefit even from a fraction of fully-annotated training samples. Experimental results are reported for the detection of six animal classes in PASCAL VOC 2007 and 2010 datasets. We demonstrate significantly improved performance of our model compared to the state-of-the-art LSVM object detector and poselet detector. Example learnt models are shown in figure 3.

This work has resulted in a submission to CVPR 2012.

6.2.8. Exploiting Photographic Style for Category-Level Image Classification by Generalizing the Spatial Pyramid
Participant: Jan van Gemert [University of Amsterdam].

This paper investigates the use of photographic style for category-level image classification. Specifically, we exploit the assumption that images within a category share a similar style defined by attributes such as colorfulness, lighting, depth of field, viewpoint and saliency. For these style attributes we create correspondences across images by a generalized spatial pyramid matching scheme. Where the spatial pyramid groups features spatially, we allow more general feature grouping and in this paper we focus on grouping images on photographic style. We evaluate our approach in an object classification task and investigate style differences between professional and amateur photographs. We show that a generalized pyramid with style-based attributes improves performance on the professional Corel and amateur Pascal VOC 2009 image datasets.

This work has resulted in a publication [20].

6.2.9. Generalized Fast Approximate Energy Minimization via Graph Cuts: Alpha-Expansion Beta-Shrink Moves
Participants: Karteek Alahari, Mark Schmidt [INRIA SIERRA].

We present alpha-expansion beta-shrink moves, a simple generalization of the widely-used alpha beta-swap and alpha-expansion algorithms for approximate energy minimization. We show that in a certain sense, these moves dominate both alpha beta-swap and alpha-expansion moves, but unlike previous generalizations the new moves require no additional assumptions and are still solvable in polynomial-time. We show promising experimental results with the new moves, which we believe could be used in any context where alpha-expansions are currently employed.
Figure 3. Illustration of deformable part models trained for the dog class. Left: dog detection with the proposed supervised part model. Right: dog detection with the original LSVM model of Felzenszwalb et al. Our model adapts to the different object appearance while LSVM model attempts to explain both samples using the same deformable HOG template.

This work has resulted in a publication [17].

6.3. Image restoration, manipulation and enhancement

6.3.1. Non-uniform Deblurring for Shaken Images

Participants: Oliver Whyte, Josef Sivic, Andrew Zisserman, Jean Ponce.

We argue that blur resulting from camera shake is mostly due to the 3D rotation of the camera, causing a blur that can be significantly non-uniform across the image. However, most current deblurring methods model the observed image as a convolution of a sharp image with a uniform blur kernel. We propose a new parametrized geometric model of the blurring process in terms of the rotational velocity of the camera during exposure. We apply this model in the context of two different algorithms for camera shake removal: the first uses a single blurry image (blind deblurring), while the second uses both a blurry image and a sharp but noisy image of the same scene. We show that our approach makes it possible to model and remove a wider class of blurs than previous approaches, and demonstrate its effectiveness with experiments on real images.

The project resulted in a publication [5].

6.3.2. Deblurring shaken and partially saturated images

Participants: Oliver Whyte, Josef Sivic, Andrew Zisserman.

We address the problem of deblurring images degraded by camera shake blur and saturated or over-exposed pixels. Saturated pixels are a problem for existing non-blind deblurring algorithms because they violate the assumption that the image formation process is linear, and often cause significant artifacts in deblurred outputs. We propose a forward model that includes sensor saturation, and use it to derive a deblurring algorithm properly treating saturated pixels. By using this forward model and reasoning about the causes of artifacts
in the deblurred results, we obtain significantly better results than existing deblurring algorithms. Further we propose an efficient approximation of the forward model leading to a significant speed-up. Example result is shown in figure 4.

The project resulted in a publication [19].

Figure 4. Deblurring saturated images. Note that the ringing around saturated regions, visible in columns (b) and (c) is removed by our method (d), without causing any loss in visual quality elsewhere.

6.3.3. Dictionary Learning for Deblurring and Digital Zoom

**Participants:** Florent Couzinie, Julien Mairal, Jean Ponce, Francis Bach [INRIA SIERRA].

This work proposes a novel approach to image deblurring and digital zooming using sparse local models of image appearance. These models, where small image patches are represented as linear combinations of a few elements drawn from some large set (dictionary) of candidates, have proven well adapted to several image restoration tasks. A key to their success has been to learn dictionaries adapted to the reconstruction of small image patches. In contrast, recent works have proposed instead to learn dictionaries which are not only adapted to data reconstruction, but also tuned for a specific task. We introduce here such an approach to deblurring and digital zoom, using pairs of blurry/sharp (or low-/high-resolution) images for training, as well as an effective stochastic gradient algorithm for solving the corresponding optimization task. Although this learning problem is not convex, once the dictionaries have been learned, the sharp/high-resolution image can be recovered via convex optimization at test time. Experiments with synthetic and real data demonstrate the effectiveness of the proposed approach, leading to state-of-the-art performance for non-blind image deblurring and digital zoom.

This work has resulted in a publication [1].

6.3.4. Sparse Image Representation with Epitomes

**Participants:** Louise Benoit, Julien Mairal, Jean Ponce, Francis Bach [INRIA SIERRA].

Sparse coding, which is the decomposition of a vector using only a few basis elements, is widely used in machine learning and image processing. The basis set, also called dictionary, is learned to adapt to specific data. This approach has proven to be very effective in many image processing tasks. Traditionally, the dictionary is an unstructured "flat" set of atoms. In this work, we study structured dictionaries which are
obtained from an epitome, or a set of epitomes. The epitome is itself a small image, and the atoms are all the
patches of a chosen size inside this image. This considerably reduces the number of parameters to learn and
provides sparse image decompositions with shift invariance properties. We propose a new formulation and an
algorithm for learning the structured dictionaries associated with epitomes, and illustrate their use in image
denoising tasks.

This work has resulted in a CVPR’11 publication [6].

6.3.5. Proximal Methods for Hierarchical Sparse Coding

Participants: Julien Mairal, Rodolphe Jenatton [INRIA SIERRA], Guillaume Obozinski [INRIA SIERRA],
Francis Bach [INRIA SIERRA].

Sparse coding consists in representing signals as sparse linear combinations of atoms selected from a
dictionary. We consider an extension of this framework where the atoms are further assumed to be embedded
in a tree. This is achieved using a recently introduced tree-structured sparse regularization norm, which has
proven useful in several applications. This norm leads to regularized problems that are difficult to optimize,
and in this paper, we propose efficient algorithms for solving them. More precisely, we show that the proximal
operator associated with this norm is computable exactly via a dual approach that can be viewed as the
composition of elementary proximal operators. Our procedure has a complexity linear, or close to linear, in
the number of atoms, and allows the use of accelerated gradient techniques to solve the tree-structured sparse
approximation problem at the same computational cost as traditional ones using the l1-norm. Our method
is efficient and scales gracefully to millions of variables, which we illustrate in two types of applications:
first, we consider fixed hierarchical dictionaries of wavelets to denoise natural images. Then, we apply our
optimization tools in the context of dictionary learning, where learned dictionary elements naturally self-
organize in a prespecified arborescent structure, leading to better performance in reconstruction of natural
image patches. When applied to text documents, our method learns hierarchies of topics, thus providing a
competitive alternative to probabilistic topic models.

This work has resulted in a publication [3].

6.4. Human activity capture and classification

6.4.1. Track to the future: Spatio-temporal video segmentation with long-range motion cues

Participants: Jose Lezama, Karteek Alahari, Ivan Laptev, Josef Sivic.

Video provides rich visual cues such as motion and appearance but also much less explored long-range tem-
poral interactions among objects. We aim to capture such interactions and to construct powerful intermediate-
level video representation for subsequent recognition. Motivated by this goal, we seek to obtain spatio-
temporal oversegmentation of the video into regions that respect object boundaries and, at the same time,
associate object pixels over many video frames. The contributions of this paper are twofold. First, we develop
an efficient spatio-temporal video segmentation algorithm, that naturally incorporates long-range motion cues
from the past and future frames in the form of clusters of point tracks with coherent motion. Second, we
device a new track clustering cost-function that includes occlusion reasoning, in the form of depth ordering
constraints, as well as motion similarity along the tracks. We evaluate the proposed approach on a challenging
set of video sequences of office scenes from feature length movies.

This work resulted in a publication [11].

6.4.2. Density-aware person detection and tracking in crowds

Participants: Mikel Rodriguez, Ivan Laptev, Josef Sivic, Jean-Yves Audibert [INRIA SIERRA].
We address the problem of person detection and tracking in crowded video scenes. While the detection of individual objects has been improved significantly over the recent years, crowd scenes remain particularly challenging for the detection and tracking tasks due to heavy occlusions, high person densities and significant variation in people’s appearance. To address these challenges, we propose to leverage information on the global structure of the scene and to resolve all detections jointly. In particular, we explore constraints imposed by the crowd density and formulate person detection as the optimization of a joint energy function combining crowd density estimation and the localization of individual people. We demonstrate how the optimization of such an energy function significantly improves person detection and tracking in crowds. We validate our approach on a challenging video dataset of crowded scenes. The proposed approach is illustrated in figure 5.

This work has resulted in a publication [14].

![Figure 5. Individual head detections provided by state-of-the-art object detector (Felzenswalb et al. 2009) (bottom-left; green: true positives; red: false positives) are improved significantly by our method (bottom-right; yellow: new true positives) using the crowd density estimate (top-right) obtained from the original frame (top-left).]

### 6.4.3. Data-driven Crowd Analysis in Videos

**Participants:** Mikel Rodriguez, Josef Sivic, Ivan Laptev, Jean-Yves Audibert [INRIA SIERRA].

In this work we present a new crowd analysis algorithm powered by behavior priors that are learned on a large database of crowd videos gathered from the Internet. The algorithm works by first learning a set of crowd behavior priors off-line. During testing, crowd patches are matched to the database and behavior priors are transferred. We adhere to the insight that despite the fact that the entire space of possible crowd behaviors is infinite, the space of distinguishable crowd motion patterns may not be all that large. For many individuals in a crowd, we are able to find analogous crowd patches in our database which contain similar patterns of behavior that can effectively act as priors to constrain the difficult task of tracking an individual in a crowd. Our algorithm is data-driven and, unlike some crowd characterization methods, does not require us to have seen the test video beforehand. It performs like state-of-the-art methods for tracking people having common crowd behaviors and outperforms the methods when the tracked individual behaves in an unusual way.

This work has resulted in a publication [15].

### 6.4.4. Learning person-object interactions for action recognition in still images

**Participants:** Vincent Delaitre, Josef Sivic, Ivan Laptev.
In this work, we investigate a discriminatively trained model of person-object interactions for recognizing common human actions in still images. We build on the locally order-less spatial pyramid bag-of-features model, which was shown to perform extremely well on a range of object, scene and human action recognition tasks. We introduce three principal contributions. First, we replace the standard quantized local HOG/SIFT features with stronger discriminatively trained body part and object detectors. Second, we introduce new person-object interaction features based on spatial co-occurrences of individual body parts and objects. Third, we address the combinatorial problem of a large number of possible interaction pairs and propose a discriminative selection procedure using a linear support vector machine (SVM) with a sparsity inducing regularizer. Learning of action-specific body part and object interactions bypasses the difficult problem of estimating the complete human body pose configuration. Benefits of the proposed model are shown on human action recognition in consumer photographs, outperforming the strong bag-of-features baseline. The proposed model is illustrated in figure 6.

This work has resulted in a publication [8].

Figure 6. Representing person-object interactions by pairs of body part (cyan) and object (blue) detectors. To get a strong interaction response, the pair of detectors (here visualized at positions pi and pj) must fire in a particular relative 3D scale-space displacement (given by the vector v) with a scale-space displacement uncertainty (deformation cost) given by diagonal 3x3 covariance matrix C (the spatial part of C is visualized as a yellow dotted ellipse). Our image representation is defined by the max-pooling of interaction responses over the whole image, solved efficiently by the distance transform.

6.4.5. People Watching: Human Actions as a Cue for Single View Geometry

Participants: David Fouhey [CMU], Vincent Delaitre, Abhinav Gupta [CMU], Ivan Laptev, Alexei Efros [CMU], Josef Sivic.

We present an approach which exploits the coupling between human actions and scene geometry. We investigate the use of human pose as a cue for single-view 3D scene understanding. Our method builds upon recent advances in still-image action recognition and pose estimation, to extract functional and geometric constraints about the scene from people detections. These constraints are then used to improve state-of-the-art single-view 3D scene understanding approaches. The proposed method is validated on a collection of single-viewpoint time-lapse image sequences as well as a dataset of still images of indoor scenes. We demonstrate that observing people performing different actions can significantly improve estimates of scene geometry and 3D layout. The main idea of this work is illustrated in figure 7.
This work is in submission to CVPR 2012.

Figure 7. What can human actions tell us about the 3D structure of the scene? Quite a lot, actually. Consider the two person detections and their estimated pose in (a). They were detected in a time-lapse sequence of one of the three scenes (b-d). Can you guess which one? Most people can easily see that it is (b). Even though this is only a static image, the actions and the pose of the disembodied figures reveal a lot about the geometric structure of the scene. The pose of the left figure reveals a horizontal surface right under its pelvis, which ends abruptly at the knees. The right figure’s pose reveals a ground plane under its feet as well as a likely horizontal surface near the hand location. In both cases we observe a strong physical and functional coupling that exists between people and the 3D geometry of the scene. Our aim in this work is to exploit this coupling.

6.4.6. Joint pose estimation and action recognition in image graphs

Participants: K. Raja [INRIA Rennes], Ivan Laptev, Patrick Perez [Technicolor], L. Osei [INRIA Rennes].

Human analysis in images and video is a hard problem due to the large variation in human pose, clothing, camera view-points, lighting and other factors. While the explicit modeling of this variability is difficult, the huge amount of available person images motivates for the implicit, datadriven approach to human analysis. In this work we aim to explore this approach using the large amount of images spanning a subspace of human appearance. We model this subspace by connecting images into a graph and propagating information through such a graph using a discriminatively trained graphical model. We particularly address the problems of human pose estimation and action recognition and demonstrate how image graphs help solving these problems jointly. We report results on still images with human actions from the KTH dataset.

This work has resulted in a publication [13].

6.5. Creation of the SIERRA project-team

6.5.1. From WILLOW alone to WILLOW and SIERRA

The WILLOW team officially started in the Spring of 2007. From the start, it was clear that machine learning was a key ingredient to new breakthroughs, and our activities have steadily grown in this area. In three short years, WILLOW has grown into a mature group of about 30 people, and it divides its activities between computer vision, machine learning, and the cross-pollination of the two fields, with video as one of the core research areas. We have been very successful, with many publications in all the major international conferences and leading journals in both areas, but we are a large group with very diverse interests, ranging from camera
geometry to statistics, and from image retrieval to bioinformatics applications of structured sparse coding. With the creation of the SIERRA project-team, the core machine learning activities of WILLOW have been transferred to the new group.

The two teams continue collaborating with each other (they remain co-located at the INRIA site in central Paris), but have a sharper focus on their respective computer vision and machine learning activities.

6.5.2. SIERRA

The SIERRA project-team was created by the INRIA on January 1st 2011 and is headed by Francis Bach, who received in 2009 a Jr. ERC grant.
6. New Results

6.1. Data and Metadata Management

6.1.1. Uncertain Data Management

Participants: Reza Akbarinia, Patrick Valduriez, Guillaume Verger.

Data uncertainty in scientific applications can be due to many different reasons: incomplete knowledge of the underlying system, inexact model parameters, inaccurate representation of initial boundary conditions, inaccuracy in equipments, etc. For instance, in the monitoring of plant contamination, sensors generate periodically data which may be uncertain. Instead of ignoring (or correcting) uncertainty, which may generate major errors, we need to manage it rigorously and provide support for querying.

In [46], we address the problem of aggregate queries that return possible sum values and their probabilities. This kind of query which, we call ALL-SUM, is also known as sum probability distribution. The results of ALL-SUM can be used for answering many other type of queries over probabilistic data. In general, the problem of ALL-SUM query execution is NP-complete. We propose pseudo-polynomial algorithms that are efficient in many practical applications, e.g. when the aggr attribute values are small integers or real numbers with small precision, i.e. small number of digits after decimal point. These cases cover many practical attributes, e.g. temperature, blood pressure, needed human resources per patient in medical applications.

We have started to develop a probabilistic database prototype, called ProbDB (Probabilistic Database), on top of an RDBMS. ProbDB divides a query into two parts: probabilistic and deterministic (i.e. non probabilistic). The deterministic part is executed by the underlying RDBMS, and the rest of work is done by our probabilistic query processing algorithms that are executed over the data returned by the RDBMS. In [51], we demonstrated the efficient execution of aggregate queries with the first version of ProbDB.

6.1.2. Metadata Integration

Participants: Zohra Bellahsène, Rémi Coletta, Duy Hoa Ngo.

Due to the various types of heterogeneity of ontologies, ontology matching must exploit many features of ontology elements in order to improve matching quality. For this purpose, numerous similarity metrics have been proposed to deal with ontology semantics at different levels: elements level, structural level and instance level.

Elements level metrics can be categorized in three groups: (1) terminological, (2) structural and (3) semantic. Metrics of the first group exploit text features such as names, labels and comments to compute the similarity score between entities. Whereas metrics of the last two groups exploit the hierarchy and semantic relationship features. Our approach consists in first using terminological metrics. Then, during the matching process, mappings discovered by terminological metrics are used as input mappings to other metrics of the second and third groups. Obviously, the more precise results terminological metrics are, the more accurate results structural and semantic metrics have.

However, finding a good combination of different metrics is very difficult and time consuming. We proposed YAM++ (not Yet Another Matcher), an approach that uses machine learning to combine similarity metrics. Our main contributions are: the definition of new metrics dealing with terminological and context profile features of entities in ontologies [37], and the use of a decision tree model to combine similarity metrics [38].

To improve matching quality of YAM++, we exploit instances accompanying ontologies. We then apply similarity flooding propagation algorithm to discover more semantic mappings. At the 2011 competition of the Ontology Alignment Evaluation Initiative (http://oaei.ontologymatching.org), YAM++ achieved excellent results: first position on the Conference track and second position on the Benchmark track [39].
6.2. Data and Process Sharing

6.2.1. Social-based P2P Data Sharing

Participants: Hinde Bouziane, Michèle Cart, Esther Pacitti, Didier Parigot, Guillaume Verger.

This work focuses on P2P content recommendation for on-line communities. In [20], we propose P2Prec, a recommendation service for P2P content sharing systems that exploits users’ social data. Given a query, P2Prec finds peers that can recommend high quality documents that are relevant for the query. A document is relevant to a query if it covers the same topics. It is of high quality if relevant peers have rated it highly. P2Prec finds relevant peers through a variety of mechanisms including advanced content-based and collaborative filtering. The topics each peer is interested in are automatically calculated by analyzing the documents the peer holds. Peers become relevant for a topic if they hold a certain number of highly rated documents on this topic. To efficiently disseminate information about peers’ topics and relevant peers, we proposed new semantic-based gossip protocols. In our experimental evaluation, using the TREC09 dataset, we showed that using semantic gossip increases recall by a factor of 1.6 compared to well-known random gossiping. Furthermore, P2Prec has the ability to get reasonable recall with acceptable query processing load and network traffic. P2Prec was demonstrated in [31] and [47].

In [30], we exploit social relationships between users as a parameter to increase the trust of recommendation. We propose a novel P2P recommendation approach (called F2Frec) that leverages content and social-based recommendation by maintaining a P2P and friend-to-friend network. This network is used as a basis to provide useful and high quality recommendations. Based on F2Frec, we propose new metrics, such as usefulness and similarity (among users and their respective friend network). We define our proposed metrics based on users’ topic of interest and relevant topics that are automatically extracted from the contents stored by each user. Our experimental evaluation, using the TREC09 dataset and Wiki vote social network, shows the benefits of our approach compared to anonymous recommendation. In addition, we show that F2Frec increases recall by a factor of 8.8 compared with centralized collaborative filtering.

6.2.2. Satisfaction-based Query Replication

Participant: Patrick Valduriez.

In a large-scale Internet-based distributed, participants (consumers and providers) who are willing to share data are typically autonomous, i.e. they may have special interests towards queries and other participants’ data. In this context, a way to avoid a participant to voluntarily leave the system is satisfying its interests when allocating queries. However, participants’ satisfaction may also be negatively affected by the failures of other participants. Query replication can deal with providers failures, but, it is challenging because of autonomy: it cannot only quickly overload the system, but also dissatisfy participants with uninteresting queries. Thus, a natural question arises: should queries be replicated? If so, which ones? and how many times?

In [25], we answer these questions by revisiting query replication from a satisfaction and probabilistic point of view. We propose a new algorithm, called S b QR, that decides on-the-fly whether a query should be replicated and at which rate. As replicating a large number of queries might overload the system, we propose a variant of our algorithm, called S b QR+. The idea is to voluntarily fail to allocate as many replicas as required by consumers for low critical queries so as to keep resources for high critical queries during query-intensive periods. Our experimental results demonstrate that our algorithms significantly outperform the baseline algorithms from both the performance and satisfaction points of view. We also show that our algorithms automatically adapt to the criticality of queries and different rates of participant failures.

6.2.3. View Selection in Scientific Data Warehousing

Participants: Zohra Bellahsène, Rémi Coletta, Imen Mami.
Scientific data generate large amounts of data which have to be collected and stored for analytical purpose. One way to help managing and analyzing large amounts of data is data warehousing, whereby views over data are materialized. However, view selection is an NP-hard problem because of many parameters: query cost, view maintenance cost and storage space. In [36], we propose a new solution based on constraint programming, which has proven efficient at solving combinatorial problems. This allows using a constraint programming solver to set up the search space by identifying a set of views that minimizes the total query cost. We address view selection under two cases: (1) only the total view maintenance cost needs be minimized, assuming unlimited storage space (meaning that it is not a critical resource anymore); (2) both storage space and maintenance cost must be minimized. We implemented our approach and compared it with a randomized method (i.e., genetic algorithm). We experimentally show that our approach provides better performance resulting from evaluating the quality of the solutions in terms of cost savings. Furthermore, our approach scales well with the query workload.

6.2.4. Scientific Workflow Management

Participants: Ayoub Ait Lahcen, Eduardo Ogasawara, Didier Parigot, Patrick Valduriez.

Scientific workflows have emerged as a basic abstraction for structuring and executing scientific experiments in computational environments. In many situations, these workflows are computationally and data-intensive, thus requiring execution in large-scale parallel computers. However, parallelization of scientific workflows remains low-level, ad-hoc and laborintensive, which makes it hard to exploit optimization opportunities.

To address this problem, we propose in [23] an algebraic approach (inspired by relational algebra) and a parallel execution model that enable automatic optimization of scientific workflows. With our scientific workflow algebra, data is uniformly represented by relations and workflow activities are mapped to operators that have data aware semantics. Our workflow execution model is based on the concept of activity activation, which enables transparent distribution and parallelization of activities;

We conducted a thorough validation of our approach using both a real oil exploitation application and synthetic data scenarios. The experiments were run in Chiron, a data-centric scientific workflow engine implemented at UFRJ to support our algebraic approach. Our experiments demonstrate performance improvements of up to 226% compared to an ad-hoc workflow implementation. This work was done in the context of the Equipe Associée Sarava and the CNPq-INRIA project DatLuge.

In the context of SON, we also proposed a declarative workflow language based on service/activity rules [41]. This language makes it possible to infer a dependency graph for SON applications that provides for automatic parallelization.

6.3. Scalable Data Analysis

6.3.1. Massive Graph Management

Participant: Patrick Valduriez.

Traversing massive graphs as efficiently as possible is essential for many scientific applications. Many common operations on graphs, such as calculating the distance between two nodes, are based on the Breadth First Search (BFS) traversal. However, because of the exhaustive exploration of all the nodes and edges of the graph, this operation might be very time consuming. A possible solution is partitioning the graph among the nodes of a shared-nothing parallel system. However, partitioning a graph and keeping the information regarding the location of vertices might be unrealistic for massive graphs because of much inter-node communication. In [28], we propose ParallelGDB, a new graph database system based on specializing the local caches of any node in this system, providing a better cache hit ratio. ParallelGDB uses a random graph partitioning, avoiding complex partition methods based on the graph topology, that usually require managing extra data structures. This proposed system provides an efficient environment for distributed graph databases.

6.3.2. Top-k Query Processing in Unstructured P2P Systems

Participants: Reza Akbarinia, William Kokou Dedzoe, Patrick Valduriez.
Top-k query processing techniques are useful in unstructured P2P systems to avoid overwhelming users with too many results and provide them with the best ones. However, existing approaches suffer from long waiting times, because top-k results are returned only when all queried peers have finished processing the query. As a result, response time is dominated by the slowest queried peer. We proposed to revisit the problem of top-k query processing.

In [29] we address the problem of reducing user waiting time of top-k query processing, in the case of unstructured P2P systems with overloaded peers. We propose a new algorithm, called QUAT, in which each peer maintains a semantic description of its local data and the semantic descriptions of its neighborhood (i.e. the semantic descriptions of data owned locally by its direct neighbors and data owned locally by these neighbors direct neighbors). These semantic descriptions allow peers to prioritize the queries that can provide high quality results, and to forward them in priority to the neighbors that can provide high quality answers. We validated our solution through a thorough experimental evaluation using a real-world dataset. The results show that QUAT significantly outperforms baseline algorithms by returning faster the final top-k results to users.

6.3.3. Top-k Query Processing Over Sorted Lists

Participants: Reza Akbarinia, Esther Pacitti, Patrick Valduriez.

The problem of answering top-k queries can be modeled as follows. Suppose we have \( m \) lists of \( n \) data items such that each data item has a local score in each list and the lists are sorted according to the local scores of their data items. Each data item has an overall score computed based on its local scores in all lists using a given scoring function. Then, the problem is to find the \( k \) data items whose overall scores are the highest. This problem model is a general model for top-k queries in many centralized, distributed and P2P applications. For example, in IR systems one of the main problems is to find the top-k documents whose aggregate rank is the highest wrt. some given keywords. To answer this query, the solution is to have for each keyword a ranked list of documents, and return the \( k \) documents whose aggregate rank in all lists are the highest.

In [16], we propose an extension of our best position algorithms (BPA) which had been proposed for top-k query processing over sorted lists model. The BPA algorithms have been shown to be more efficient than the well known TA Algorithm. We propose several techniques using different data structures for managing best positions that are crucial for efficient execution of top-k algorithms. We also provide a complete discussion on the instance optimality of TA algorithm (TA was considered so far as optimal over any database of sorted lists). We illustrate that, the existence of deterministic algorithms such as BPA shows that if we are aware of positions of seen data, then one of the main arguments used for proving the instance optimality of TA is invalidated. Therefore, in this case the proof of TA's instance optimality is incorrect, and must be revisited.

6.3.4. Satellite Image Mining

Participant: Florent Masseglia.

Satellite Image Time Series (SITS) provide us with precious information on land cover evolution. By studying SITS, we can both understand the changes of specific areas and discover global phenomena that spread over larger areas. Changes that can occur throughout the sensing time can spread over very long periods and may have different start time and end time depending on the location, which complicates the mining and the analysis of series of images. In [45], we propose a frequent sequential pattern mining method for SITS analysis. Designing such a method called for important improvements on the data mining principles. First, the search space in SITS is multi-dimensional (the radiometric levels of different wavelengths correspond to infra-red, red, etc.). Furthermore, the non evolving regions, which are the vast majority and overwhelm the evolving ones, challenge the discovery of these patterns. Our framework enables discovery of these patterns despite these constraints and characteristics. We introduce new filters in the mining process to yield important reductions in the search space by avoiding consecutive occurrences of similar values in the sequences. Then, we propose visualization techniques for results analysis (where modified regions are highlighted). Experiments carried out on a particular dataset showed that our method allows extracting repeated, shifted and distorted temporal behaviors. The flexibility of this method makes it possible to capture complex behaviors from multi-source, noisy and irregularly sensed data.
6.3.5. **Distributed Approximate Similarity Join**

**Participant:** Alexis Joly.

Efficiently constructing the KNN-graph of large and high dimensional feature datasets is crucial for many data intensive applications involving feature-rich objects, such as image features, text features or sensor's features. In this work we investigate the use of high dimensional hashing methods for efficiently approximating the full kNN graph of large collections, in particular, in distributed environments. We first analyzed and experimented what seems to be the most intuitive hashing-based approach: constructing several Locality Sensitive Hashing (LSH) tables in parallel and computing the frequency of all emitted collisions. We show that balancing issues of classical LSH functions strongly affect the performance of this approach. On the other side, we show that using an alternative data-dependent hashing function (RMMH), that we introduced recently [34], can definitely change that conclusion. The main originality of RMMH hash function family is that it is based on randomly trained classifiers, allowing to learn random and balanced splits of the data instead of using random splits of the feature space as in LSH. We show that the hash tables constructed through RMMH are much more balanced and that the number of emitted collisions can be strongly reduced without degrading quality. In the end, our hashing-based filtering algorithm of the all-pairs graph is two orders of magnitude faster than the one based on LSH. An efficient distributed implementation of the method was implemented within the MapReduce framework (and is the basis of the SimJoin prototype). This work is done in the context of the supervision of a PhD student working at INRIA Imedia (Riadh Mohamed Trad).

6.3.6. **Visual objects mining**

**Participant:** Alexis Joly.

State-of-the-art content-based object retrieval systems have demonstrated impressive performance in very large image datasets. These methods, based on fine local descriptions and efficient matching techniques, can detect accurately very small rigid objects with unambiguous semantics such as logos, buildings, manufactured objects, posters, etc. Mining such small objects in large collections is however difficult. Constructing a full local matching graph with a naïve approach would indeed require to probe all candidate query leading to an intractable algorithm complexity. In this work, we first introduce an adaptive weighted sampling scheme, starting with some prior distribution and iteratively converging to unvisited regions [35]. We show that the proposed method allows to discover highly interpretable visual words while providing excellent recall and image representativity. We then focused on mining visual objects on top of the discovered visual words. We therefore developed an original shared nearest-neighbors clustering method, working directly on the generated bi-partite graph. This work is in the context of the supervision of two PhD students, one working jointly with INA and INRIA and who will join the Zenith team next year (Pierre Letessier), one working at INRIA Imedia (Amel Hamzaoui).

6.3.7. **Visual-based plant species identification from crowdsourced data**

**Participant:** Alexis Joly.

Inspired by citizen sciences, the main goal of this work is to speed up the collection and integration of raw botanical observation data, while providing to potential users an easy and efficient access to this botanical knowledge. We therefore designed and developed an original crowdsourcing web application dedicated to the access of botanical knowledge through automated identification of plant species by visual content. Technically, the first side of the application deals with content-based identification of plant leaves. Whereas state-of-the-art methods addressing this objective are mostly based on leaf segmentation and boundary shape features, we developed a new approach based on local features and large-scale matching. This approach obtained the best results within ImageCLEF 2011 plant identification benchmark [48]. The second side of the application deals with interactive tagging and allows any user to validate or correct the automatic determinations returned by the system. Overall, this collaborative system allows to enrich automatically and continuously the visual botanical knowledge and thus to increase progressively the accuracy of the automated identification. A demo of the developed application was presented at the ACM Multimedia conference [33]. This work is done in collaboration with INRIA Imedia and with the botanists of the AMAP UMR team (CIRAD). It is also closely related to a citizen science project around plant’s identification that we develop with the support of the TelaBotanica social network.