Activity Report 2011

Section New Results
<table>
<thead>
<tr>
<th>ALGORITHMS, PROGRAMMING, SOFTWARE AND ARCHITECTURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. CAMUS Team ............................................</td>
</tr>
<tr>
<td>2. CARAMEL Project-Team ................................</td>
</tr>
<tr>
<td>3. CARTE Project-Team ..................................</td>
</tr>
<tr>
<td>4. CASSIS Project-Team ..................................</td>
</tr>
<tr>
<td>5. PAREO Project-Team ...................................</td>
</tr>
<tr>
<td>6. TRIO Project-Team ....................................</td>
</tr>
<tr>
<td>7. VEGAS Project-Team ...................................</td>
</tr>
<tr>
<td>8. VERIDIS Team .........................................</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>APPLIED MATHEMATICS, COMPUTATION AND SIMULATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>9. CALVI Project-Team ..................................</td>
</tr>
<tr>
<td>10. CORIDA Project-Team ................................</td>
</tr>
<tr>
<td>11. TOSCA Project-Team ................................</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COMPUTATIONAL SCIENCES FOR BIOLOGY, MEDICINE AND THE ENVIRONMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>12. BIGS Project-Team ...............................................</td>
</tr>
<tr>
<td>13. CORTEX Project-Team ...............................................</td>
</tr>
<tr>
<td>14. MASAIE Project-Team ...............................................</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NETWORKS, SYSTEMS AND SERVICES, DISTRIBUTED COMPUTING</th>
</tr>
</thead>
<tbody>
<tr>
<td>15. ALGORILLE Project-Team .......................................</td>
</tr>
<tr>
<td>16. MADYNES Project-Team .........................................</td>
</tr>
<tr>
<td>17. SCORE Team .....................................................</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PERCEPTION, COGNITION, INTERACTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>18. ALICE Project-Team ..................</td>
</tr>
<tr>
<td>19. MAGRIT Project-Team ..................</td>
</tr>
<tr>
<td>20. MAIA Project-Team ....................</td>
</tr>
<tr>
<td>21. ORPAILLEUR Project-Team ..............</td>
</tr>
<tr>
<td>22. PAROLE Project-Team .................</td>
</tr>
<tr>
<td>23. Sémagramme Team ........................</td>
</tr>
<tr>
<td>24. TALARIS Project-Team ................</td>
</tr>
</tbody>
</table>
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
Algorithmics, Programming, Software and Architecture - New Results - Team CAMUS

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 programing 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 an 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](image-url)
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 $O(\log^8 q)$ to $O(\log^5 q)$. 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.
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. 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 database 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 $O(2^n \log n)$. 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 experimentations.
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 formulae, 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 Fourneret, 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, Elizabeeta 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.
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\(^1\) 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\(^2\), 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\(^3\) 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.

---


\(^2\)available on the Quarteft subversion repository: [https://gforge.enseeiht.fr/projects/quarteft/](https://gforge.enseeiht.fr/projects/quarteft/)

\(^3\)Object Constraint Language, [http://www.omg.org/spec/OCL/2.2/](http://www.omg.org/spec/OCL/2.2/)
TRIO Project-Team

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 PLPlatform 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 PLPlatform 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-PEOPE 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 [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 [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 [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 intelectual property issues. This problem may be avoided by estimating statistically either the WCET of a task [18] or the response time of each task [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 [12], we have proposed feasibility tests for tasks with probabilistic execution times [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/](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 Simonet-Lion, Liliana Cucu-Grosjean, Ammar Oulamara, Luca Santinelli.

TIMMO-2-USE ([http://timmo-2-use.org/](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 ATESTST2 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. 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 Groebner 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 naive 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 $S$ of $O(n^2/\log n)$ points with at most one bend per edge and with the vertices and the bend points in $S$. 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 [33] has been reduced by about 90%. The backend has been integrated in TLAPS and has been presented at a workshop [19].

6.5. Model checking within SimGrid
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. [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 [13]. Cristián Rosa successfully defended his PhD thesis [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. A new version of PlusCal
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. 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 $O\left(\min\left(\frac{\Delta x}{\Delta t}, 1\right)\Delta x^p + \Delta t^2\right)$ 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 $\theta$ 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 $(r, \theta)$ 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 comparisons 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 Cadarache).

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 tokomak 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].
CORIDA Project-Team

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-Lebou-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 \( \dot{y} = Ay + Bu \) where \( A : \mathcal{D}(A) \rightarrow X \) is 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 \( K \) 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 a 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.
TOSCA Project-Team

6. New Results

6.1. Probabilistic numerical methods, stochastic modelling and applications

Participants: Mireille Bossy, Nicolas Champagnat, Julia Charrier, Julien Claisse, Madalina Deaconu, Samuel Herrmann, James Inglis, Pierre-Emmanuel Jabin, Antoine Lejay, Sylvain Maire, Sebastian Niklitschek Soto, Nicolas Perrin, Denis Talay, Etienne Tanrè, 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 \( \mathcal{D} \) is bounded with smooth boundary, they constructed a confined primitive of Brownian motion in \( \mathcal{D} \) 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.

- 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 MANedy project.

- S. Herrmann and E. Tanrè 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 MANedy 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 particles 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 (MERE 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], 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], 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. 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 replica of \((X, Y)\) 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. 

(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:

$$
\begin{align*}
\text{d}S_t &= [\alpha - k Q_t] S_t \text{d}t + \varepsilon S_t \text{d}W^1_t, \\
\text{d}Q_t &= [d - m Q_t - k Q_t S_t + k b e^{-\mu Q_t} Q_{t-\zeta} S_{t-\zeta}] \text{d}t + \varepsilon Q_t \text{d}W^2_t.
\end{align*}
$$

(1)
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 $\epsilon\,dW_t$ 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, Auraelie, 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 $R_r$, $r \in \mathbb{R}$; $\mathbb{R}$ is supposed to be finite. These reactions involve species indexed by a set $S = 1, \ldots, M$, 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 $\lambda_r(X)$ which depends on the state of the system, described by $X$ and corresponds to a change $X \rightarrow X + \gamma_r$, $\gamma_r \in \mathbb{Z}^M$.

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 $x = (x_C, X_D)$.
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 antiperistence, 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:

$$X_t^\theta = a + \theta \int_0^t b(X_u) \, du + B_t,$$

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:

$$X_t^\theta = a + \int_0^t b(\theta; X_u) \, du + \int_0^t \sigma(\theta; X_u) \, dB_t,$$  \hspace{1cm}(2)

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:

\[
\forall c > 0, (X(c^E x))_{x \in \mathbb{R}^d} \overset{d}{=} c \{X(x)\}_{x \in \mathbb{R}^d}, \quad \text{where} \quad c^E := \exp(E \ln(c)).
\]

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 harmonizable 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

\[
S_h^N = \left\{ \sum_{n_1=1}^{N_1} \ldots \sum_{n_d=1}^{N_d} \frac{X_n(h_n^N)}{N^n}; \quad t \in [0, +\infty)^d \right\}
\]

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. 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 feed-back 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**

**Participants:** Frédéric Alexandre, Laurent Bougrain, Octave Boussaton, Axel Hutt, Baptiste Payan, Maxime Rio, Carolina Saavedra, Christian Weber.

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].
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.

![Public transport flow in Rio](image)

*Figure 1. Transportation system in Rio*

We have considered a reduced system represented by Figure 2. The outputs of the model compare favorably with the actual data (Figure 3).

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. 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 pre-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 issue 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 continue 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: a 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 a 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. 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-P2P3 Project funded by GIS-3SGS4.

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.

\(^3\) Approche Collaborative pour la Détection d’Attaques dans les réseaux Pair à Pair

\(^4\) Groupement 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.

In parallel, 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 an 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 Fesler, 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 Cyber 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 taces 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 ans 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.
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. Alignment 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 realtime 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.
ALICE Project-Team

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, Rhaeleb 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 tesselation [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 Bolcheva 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


- **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, Rhaleb 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).
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 fulfil 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 developped in our team. This year she dealt with the hand-eye coordination between the different devices. Moreover, a 3D laser pointer has being 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 \cite{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 \cite{9}.

We used the same task analysis technique to design an inguinal hernia repair simulator \cite{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

  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, we design a compiled planning task which introduces constructs allowing 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.

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 SED6 (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 IAVs 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 have 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. 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 concerned with the mining of complex hydroecological data using 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 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 into 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

Participants: Mehwish Alam, Yasmine Assess, Sid-Ahmed Benabderrahmane, Emmanuel Bresso, Thomas
Bourquard, Adrien Coulet, Sébastien Da Silva, Marie-Dominique Devignes, Anisah Ghoorah, Renaud Grisoni,
Mehdi Kaytoue, Jean-François Kneib, Florence Le Ber, Bernard Maigret, Jean-François Mari, Lazaros
Mavridis, Amedeo Napoli, Violeta Pérez-Nuено, Dave Ritchie, Malika Smail-Tabbone, Vishwesh Venkatra-
man.

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 Mavrdis, Violeta Pérez-Nueno, Dave Ritchie, Malika Smaïl-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 Chaires 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 Smaïl-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
involves 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 http://hex.loria.fr. A public GPU-powered server has also been created (http://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 KDBOCK system as part of her doctoral thesis project. KDBOCK 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 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 \cite{104}. The aim of this project is to investigate the use of Voronoi fingerprints \cite{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 \textit{Hex} \cite{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 \cite{60}. This approach was also used to predict the stability of engineered protein structures for another recent CAPRI target \cite{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 \cite{110}. The main disease target in this project is the acquired immune deficiency syndrome (AIDS), caused by the human immuno-deficiency virus (HIV) \cite{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 \cite{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 \cite{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

**Participants:** Julien Cojan, Valmi Dufour-Lussier, Inaki Fernandez, Emmanuelle Gaillard, Laura Infante-Blanco, Florence Le Ber, Jean Lieber, Amedeo Napoli, Emmanuel Nauer, Yannick Toussaint.
The Taaable project (http://taable.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, an 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 $ALC$ [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].
PAROLE Project-Team

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 adapation 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 join 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 goals: ADT (Action of Technological Development) Handicom [piqueardkipffer: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 linguistic, 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 ISCT 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%

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 Mesbahi.
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 or 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 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
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 \( N(\mu, \sigma^2) \) 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 alignment 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 language) 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. 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 coordinative 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 $\eta$-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].

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.
TALARIS Project-Team

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 METAVERSE1 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 that 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 IGNGF 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 GenL 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.