Activity Report 2011

Section Application Domains
1. AMIB Project-Team (section vide) ................................................................. 4
2. AVIZ Project-Team .......................................................................................... 5
3. BYMOORE Exploratory Action (section vide) .................................................. 6
4. COMETE Project-Team ..................................................................................... 7
5. COMMANDS Project-Team ............................................................................. 8
6. DAHU Project-Team ......................................................................................... 9
7. DEFI Project-Team ........................................................................................... 10
8. DIGIPLANTE Team (section vide) ................................................................... 12
9. DISCO Team ..................................................................................................... 13
10. GALEN Team .................................................................................................. 15
11. GECO Team .................................................................................................... 16
12. GEOMETRICA Project-Team ........................................................................... 20
13. GRAND-LARGE Project-Team ......................................................................... 21
14. HIPERCOM Project-Team ................................................................................ 24
15. IN-SITU Project-Team ..................................................................................... 26
16. LEO Team ....................................................................................................... 27
17. MAXPLUS Project-Team .................................................................................. 28
18. MEXICO Project-Team .................................................................................... 32
19. PARIETAL Project-Team .................................................................................. 34
20. PARSIFAL Project-Team .................................................................................. 35
21. PROVAL Project-Team ..................................................................................... 37
22. REGULARITY Team ......................................................................................... 38
23. SECSI Project-Team ....................................................................................... 42
24. SELECT Project-Team ..................................................................................... 43
25. TANC Project-Team ......................................................................................... 45
26. TAO Project-Team .......................................................................................... 46
27. TYPICAL Project-Team ................................................................................... 47
AMIB Project-Team (section vide)
4. Application Domains

4.1. Application Domains

AVIZ develops active collaboration with users from various application domains, making sure it can support their specific needs. By studying similar problems in different domains, we can begin to generalize our results and have confidence that our solutions will work for a variety of applications.

Our current application domains include:

- *Genealogy*, in cooperation with North Carolina State University;
- *Biological research*, in cooperation with Institut Pasteur;
- *Digital Libraries*, in cooperation with the French National Archives and the Wikipedia community;
- *Open Data*, in cooperation with Google Open Data;
- *Agrifood Process Modeling*, in cooperation with the DREAM project;
BYMOORE Exploratory Action (section vide)
4. Application Domains

4.1. Security and privacy

Participants: Mário Sergio Ferreira Alvim Junior, Miguel Andrés, Nicolás Bordenabe, Konstantinos Chartzikokolakis, Jérémy Dubreil, Catuscia Palamidessi.

The aim of our research is the specification and verification of protocols used in mobile distributed systems, in particular security protocols. We are especially interested in protocols for information hiding.

Information hiding is a generic term which we use here to refer to the problem of preventing the disclosure of information which is supposed to be secret or confidential. The most prominent research areas which are concerned with this problem are those of secure information flow and of privacy.

Secure information flow refers to the problem of avoiding the so-called propagation of secret data due to their processing. It was initially considered as related to software, and the research focussed on type systems and other kind of static analysis to prevent dangerous operations. Nowadays the setting is more general, and a large part of the research effort is directed towards the investigation of probabilistic scenarios and threats.

Privacy denotes the issue of preventing certain information to become publicly known. It may refer to the protection of private data (credit card number, personal info etc.), of the agent’s identity (anonymity), of the link between information and user (unlinkability), of its activities (unobservability), and of its mobility (untraceability).

The common denominator of this class of problems is that an adversary can try to infer the private information (secrets) from the information that he can access (observables). The solution is then to obfuscate the link between secrets and observables as much as possible, and often the use randomization, i.e. the introduction of noise, can help to achieve this purpose. The system can then be seen as a noisy channel, in the information-theoretic sense, between the secrets and the observables.

We intend to explore the rich set of concepts and techniques in the fields of information theory and hypothesis testing to establish the foundations of quantitative information flow and of privacy, and to develop heuristics and methods to improve mechanisms for the protection of secret information. Our approach will be based on the specification of protocols in the probabilistic asynchronous π-calculus, and the application of model-checking to compute the matrices associated to the corresponding channels.
4. Application Domains

4.1. Introduction

Commands is a team with a strong commitment in tackling real-life applications in addition to theoretical challenges. This shows in our long history of contracts with industrial partners. In the recent years, we have mainly contributed to the following fields of application.

4.2. Aerospace applications

In the framework of a long-term partnership with the Cnes, and more recently Astrium, we have studied trajectory optimization for space launcher problems. This kind of problems typically involves hard constraints (thermal flux, mechanical efforts) and inexact models (atmosphere, aerodynamic forces). The two main achievements were to study when singular arcs may occur, and to show the effectiveness of a HJB approach on a reduced model. Singular arcs are flight phases with a non-maximal thrust, induced by a tradeoff between speed and atmospheric drag; they cause difficulties of both theoretical and practical nature. The latter point is the first step in the process of applying global methods to this class of difficult problems.

4.3. Trading applications

In a partnership with Total, we have studied problems dealing with the trading of Liquefied Natural Gas. We have computed maximizing revenue policies, by combining the Stochastic Dual Dynamic Programming approach (SDDP) with a quantization method for the noise that enters in prices. We have also given partial results for the case of integer decision.

4.4. Energy applications

With Renault, we have studied problems of energy management for hybrid vehicles. Hybrid vehicles include an auxiliary thermal (gas) engine that is used as a range extender for the main electric propulsion. We are interested in determining the optimal policies for energy management, taking into account some stochastic uncertainties, as well as execution delay and decision lags.
4. Application Domains

4.1. Application Domains

Databases are pervasive across many application fields. Indeed, most human activities today require some form of data management. In particular, all applications involving the processing of large amounts of data require the use of a database. Increasingly complex Web applications and services also rely on DBMS, and their correctness and robustness is crucial.

We believe that the automated solutions that Dahu aims to develop for verifying such systems will be useful in this context.
4. Application Domains

4.1. Radar and GPR applications

Conventional radar imaging techniques (ISAR, GPR, ...) use backscattering data to image targets. The commonly used inversion algorithms are mainly based on the use of weak scattering approximations such as the Born or Kirchhoff approximation leading to very simple linear models, but at the expense of ignoring multiple scattering and polarization effects. The success of such an approach is evident in the wide use of synthetic aperture radar techniques.

However, the use of backscattering data makes 3-D imaging a very challenging problem (it is not even well understood theoretically) and as pointed out by Brett Borden in the context of airborne radar: “In recent years it has become quite apparent that the problems associated with radar target identification efforts will not vanish with the development of more sensitive radar receivers or increased signal-to-noise levels. In addition it has (slowly) been realized that greater amounts of data - or even additional “kinds” of radar data, such as added polarization or greatly extended bandwidth - will all suffer from the same basic limitations affiliated with incorrect model assumptions. Moreover, in the face of these problems it is important to ask how (and if) the complications associated with radar based automatic target recognition can be surmounted.” This comment also applies to the more complex GPR problem.

Our research themes will incorporate the development, analysis and testing of several novel methods, such as sampling methods, level set methods or topological gradient methods, for ground penetrating radar application (imaging of urban infrastructures, landmines detection, underground waste deposits monitoring, ...) using multistatic data.

4.2. Biomedical imaging

Among emerging medical imaging techniques we are particularly interested in those using low to moderate frequency regimes. These include Microwave Tomography, Electrical Impedance Tomography and also the closely related Optical Tomography technique. They all have the advantage of being potentially safe and relatively cheap modalities and can also be used in complementarity with well established techniques such as X-ray computed tomography or Magnetic Resonance Imaging.

With these modalities tissues are differentiated and, consequently, can be imaged, based on differences in dielectric properties (some recent studies have proved that dielectric properties of biological tissues can be a strong indicator of the tissues functional and pathological conditions, for instance, tissue blood content, ischemia, infarction, hypoxia, malignancies, edema and others). The main challenge for these functionalities is to build a 3-D imaging algorithm capable of treating multi-static measurements to provide real-time images with highest (reasonably) expected resolutions and in a sufficiently robust way.

Another important biomedical application is brain imaging. We are for instance interested in the use of EEG and MEG techniques as complementary tools to MRI. They are applied for instance to localize epileptic centers or active zones (functional imaging). Here the problem is different and consists into performing passive imaging: the epileptic centers act as electrical sources and imaging is performed from measurements of induced currents. Incorporating the structure of the skull is primordial in improving the resolution of the imaging procedure. Doing this in a reasonably quick manner is still an active research area, and the use of asymptotic models would offer a promising solution to fix this issue.
4.3. Non-destructive testing and parameter identification

One challenging problem in this vast area is the identification and imaging of defaults in anisotropic media. For instance, this problem is of great importance in aeronautic constructions due to the growing use of composite materials. It also arises in applications linked with the evaluation of wood quality, like locating knots in timber in order to optimize timber-cutting in sawmills, or evaluating wood integrity before cutting trees. The anisotropy of the propagative media renders the analysis of diffracted waves more complex since one cannot only rely on the use of backscattered waves. Another difficulty comes from the fact that the micro-structure of the media is generally not well known a priori.

Our concern will be focused on the determination of qualitative information on the size of defaults and their physical properties rather than a complete imaging which for anisotropic media is in general impossible. For instance, in the case of homogeneous background, one can link the size of the inclusion and the index of refraction to the first eigenvalue of so-called interior transmission problem. These eigenvalues can be determined from the measured data and a rough localization of the default. Our goal is to extend this kind of idea to the cases where both the propagative media and the inclusion are anisotropic. The generalization to the case of cracks or screens has also to be investigated.

In the context of nuclear waste management, many studies are conducted on the possibility of storing waste in a deep geological clay layer. To assess the reliability of such a storage without leakage, it is necessary to have a precise knowledge of the porous media parameters (porosity, tortuosity, permeability, etc.). The large range of space and time scales involved in this process requires a high degree of precision as well as tight bounds on the uncertainties. Many physical experiments are conducted in situ which are designed for providing data for parameters identification. For example, the determination of the damaged zone (caused by excavation) around the repository area is of paramount importance since microcracks yield drastic changes in the permeability. Level set methods are a tool of choice for characterizing this damaged zone.

4.4. Diffusion MRI

- Detecting physiological and pathological conditions that are accompanied by higher or lower than normal diffusion MRI signal attenuation. Examples: immediately after stroke, there is a large drop in the measured apparent diffusion coefficient; demyelinating diseases of the central nervous system have been indicated by higher than normal radial diffusivity.
- Evaluating cancer treatment by quantifying tumor cellularity based on diffusion MRI measurements. Tumor cellularity is shown to be inversely correlated to measured diffusivity.
DIGIPLANTE Team (section vide)
4. Application Domains

4.1. Modeling and analysis of Acute Myeloid Leukemia

In collaboration with the BANG project-team at INRIA Paris-Rocquencourt, the DRACULA team at INRIA Grenoble - Rhône-Alpes, the COMMANDS project team at INRIA Saclay-Île-de-France, INSERM, Cordeliers Research Center and St Antoine Hospital, Paris, we consider the modeling and control of Acute Myeloid Leukemia (AML).

The main goal of this project is the theoretical optimization of drug treatments used in AML, with experimental validation in cell cultures, aiming at proposing efficient therapeutic strategies in clinic.

We work on an discrete maturity-structured model of hematopoiesis introduced in [98]. In this model, several generations of cells are considered and, for the first time, the cell cycle duration is assumed to be distributed. At each level, the population of immature cells are divided into two subpopulations: proliferating and non proliferating cells. Physiological phenomena of re-introduction from the non proliferative into the proliferative subpopulation is modeled in the team as a nonlinear dynamical interconnection between the two sub-populations, and input-output tools seem to be useful in this context [35].

4.2. Control of continuous bioreactors

We study problems of coexistence or regulation of species of micro-organisms in bio-reactors called chemostats.

In [37], we have studied a competition model between an arbitrary number of species in a chemostat with one limiting substrate and including both monotone and non-monotone growth functions, distinct removal rates and variable yields. The dilution rate and the substrate input concentration were chosen as positive constants. We have shown that only the species with the lowest break-even concentration survives, provided that additional technical conditions on the growth functions and yields are satisfied. The proof relies on the construction of a Lyapunov function

In [62], we studied chemostat models in which the species compete for two or more limiting substrates. First we considered the case where the nutrient flow and species removal rates and input nutrient concentrations are all given positive constants. In that case, we use Brouwer degree theory to give conditions guaranteeing that the models admit globally asymptotically stable componentwise positive equilibrium points, from all componentwise positive initial states. Then we used the results to develop stabilization theory for controlled chemostats with two or more limiting nutrients. For cases where the dilution rate and input nutrient concentrations can be selected as controls, we prove that many different componentwise positive equilibria can be made globally asymptotically stable. This significantly extends the existing control results for chemostats with one limiting nutrient. We demonstrate our methods in simulations.

4.3. PVTOL Aircraft

In [20] and [51], we applied the technique of backstepping and of construction of strict Lyapunov functions to solve a tracking problem for the celebrated aircraft model PVTOL (Planar Vertical Takeoff and Landing). It is a benchmark dynamics for an aircraft moving in a vertical plane that contains the important features needed to design controllers for real aircraft. The controllers are the thrust out of the bottom and the rolling moment controller. The main challenges are that the thrust controller must remain nonnegative and that the system is underactuated. We overcame these challenges through a change of variables that transforms the PVTOL tracking dynamics into a chain of three subsystems and then applying asymptotic strict Lyapunov function methods and bounded backstepping. Relative to the PVTOL model literature, the significance of our PVTOL work was (a) the global boundedness of our controllers in the decoupled coordinates, (b) their applicability to
cases where the velocity measurements are not available, by using an observer, (c) the positive lower bound on the thrust controller, (d) our allowing a very general class of reference trajectories, and (e) our use of ISS to certify good performance under actuator errors, which would not be possible using LaSalle invariance or nonstrict Lyapunov functions.
4. Application Domains

4.1. Application Domains

- **Large Scale Urban Modeling**: The use of satellite imaging along with range data towards large scale image-driven reconstruction. The aim is to produce scalable representations of 3D models that are compact, modular and able to provide realistic 3D representations of real visual data.

- **Objet Recognition**: The use annotated data-bases towards learning class-specific visual and geometric object characteristics to perform recognition.

- **MR & Muscular Diseases**: The use of MR and Diffusion Tensor Imaging are investigated in collaboration with the Henri Mondor University Hospital and Institut of Myology towards automatic quantification of muscular mass loss and non-invasive biopsy. The aim is to provide tools that could be used to automatically analyze MR imaging and extract useful clinical measurements (Institut of Myology), and assess the potential impact of diffusion tensor imaging towards automatic quantification either of muscular diseases progression.

- **MR Brain Imaging towards Low-Gliomas Tumor Brain Understanding**: The use of contrast enhanced imaging is investigated in collaboration with the Montpellier University Hospital towards better understanding of low-gliomas positioning, automatic tumor segmentation/identification and longitudinal (tumor) growth modeling.
4. Application Domains

4.1. Quantum control

The issue of designing efficient transfers between different atomic or molecular levels is crucial in atomic and molecular physics, in particular because of its importance in those fields such as photochemistry (control by laser pulses of chemical reactions), nuclear magnetic resonance (NMR, control by a magnetic field of spin dynamics) and, on a more distant time horizon, the strategic domain of quantum computing. This last application explicitly relies on the design of quantum gates, each of them being, in essence, an open loop control law devoted to a prescribed simultaneous control action. NMR is one of the most promising techniques for the implementation of a quantum computer.

Physically, the control action is realized by exciting the quantum system by means of one or several external fields, being them magnetic or electric fields. The resulting control problem has attracted increasing attention, especially among quantum physicists and chemists (see, for instance, [78], [84]). The rapid evolution of the domain is driven by a multitude of experiments getting more and more precise and complex (see the recent review [37]). Control strategies have been proposed and implemented, both on numerical simulations and on physical systems, but there is still a large gap to fill before getting a complete picture of the control properties of quantum systems. Control techniques should necessarily be innovative, in order to take into account the physical peculiarities of the model and the specific experimental constraints.

The area where the picture got clearer is given by finite dimensional linear closed models.

- **Finite dimensional** refers to the dimension of the space of wave functions, and, accordingly, to the finite number of energy levels.
- **Linear** means that the evolution of the system for a fixed (constant in time) value of the control is determined by a linear vector field.
- **Closed** refers to the fact that the systems are assumed to be totally disconnected from the environment, resulting in the conservation of the norm of the wave function.

The resulting model is well suited for describing spin systems and also arises naturally when infinite dimensional quantum systems of the type discussed below are replaced by their finite dimensional Galerkin approximations. Without seeking exhaustiveness, let us mention some of the issues that have been tackled for finite dimensional linear closed quantum systems:

- controllability [19],
- bounds on the controllability time [15],
- STIRAP processes [89],
- simultaneous control [61],
- optimal control ([57], [28], [39]),
- numerical simulations [67].

Several of these results use suitable transformations or approximations (for instance the so-called rotating wave) to reformulate the finite-dimensional Schrödinger equation as a sub-Riemannian system. Open systems have also been the object of an intensive research activity (see, for instance, [20], [58], [79], [34]).
In the case where the state space is infinite dimensional, some optimal control results are known (see, for instance, [24], [35], [54], [25]). The controllability issue is less understood than in the finite dimensional setting, but several advances should be mentioned. First of all, it is known that one cannot expect exact controllability on the whole Hilbert sphere [88]. Moreover, it has been shown that a relevant model, the quantum oscillator, is not even approximately controllable [80], [70]. These negative results have been more recently completed by positive ones. In [26], [27] Beauchard and Coron obtained the first positive controllability result for a quantum particle in a 1D potential well. The result is highly nontrivial and is based on Coron’s return method (see [43]). Exact controllability is proven to hold among regular enough wave functions. In particular, exact controllability among eigenfunctions of the uncontrolled Schrödinger operator can be achieved. Other important approximate controllability results have then been proved using Lyapunov methods [69], [74], [55]. While [69] studies a controlled Schrödinger equation in $\mathbb{R}$ for which the uncontrolled Schrödinger operator has mixed spectrum, [74], [55] deal mainly with general discrete-spectrum Schrödinger operators.

In all the positive results recalled in the previous paragraph, the quantum system is steered by a single external field. Different techniques can be applied in the case of two or more external fields, leading to additional controllability results [46], [31].

The picture is even less clear for nonlinear models, such as Gross–Pitaevski and Hartree–Fock equations. The obstructions to exact controllability, similar to the ones mentioned in the linear case, have been discussed in [52]. Optimal control approaches have also been considered [23], [36]. A comprehensive controllability analysis of such models is probably a long way away.

4.2. Neurophysiology

At the interface between neurosciences, mathematics, automatics and humanoid robotics, an entire new approach to neurophysiology is emerging. It arouses a strong interest in the four communities and its development requires a joint effort and the sharing of complementary tools.

A family of extremely interesting problems concerns the understanding of the mechanisms supervising some sensorial reactions or biomechanics actions such as image reconstruction by the primary visual cortex, eyes movement and body motion.

In order to study these phenomena, a promising approach consists in identifying the motion planning problems undertaken by the brain, through the analysis of the strategies that it applies when challenged by external inputs. The role of control is that of a language allowing to read and model neurological phenomena. The control algorithms would shed new light on the brain’s geometric perception (the so-called neurogeometry [76]) and on the functional organization of the motor pathways.

- A challenging problem is that of the understanding of the mechanisms which are responsible for the process of image reconstruction in the primary visual cortex V1.

The visual cortex areas composing V1 are notable for their complex spatial organization and their functional diversity. Understanding and describing their architecture requires sophisticated modeling tools. At the same time, the structure of the natural and artificial images used in visual psychophysics can be fully disclosed only using rather deep geometric concepts. The word “geometry” refers here to the internal geometry of the functional architecture of visual cortex areas (not to the geometry of the Euclidean external space). Differential geometry and analysis both play a fundamental role in the description of the structural characteristics of visual perception.

A model of human perception based on a simplified description of the visual cortex V1, involving geometric objects typical of control theory and sub-Riemannian geometry, has been first proposed by Petitot ( [77]) and then modified by Citti and Sarti ( [42]). The model is based on experimental observations, and in particular on the fundamental work by Hubel and Wiesel [51] who received the Nobel prize in 1981.
In this model, neurons of V1 are grouped into orientation columns, each of them being sensitive to visual stimuli arriving at a given point of the retina and oriented along a given direction. The retina is modeled by the real plane, while the directions at a given point are modeled by the projective line. The fiber bundle having as base the real plane and as fiber the projective line is called the bundle of directions of the plane.

From the neurological point of view, orientation columns are in turn grouped into hypercolumns, each of them sensitive to stimuli arriving at a given point, oriented along any direction. In the same hypercolumn, relative to a point of the plane, we also find neurons that are sensitive to other stimuli properties, such as colors. Therefore, in this model the visual cortex treats an image not as a planar object, but as a set of points in the bundle of directions of the plane. The reconstruction is then realized by minimizing the energy necessary to activate orientation columns among those which are not activated directly by the image. This gives rise to a sub-Riemannian problem on the bundle of directions of the plane.

- Another class of challenging problems concern the functional organization of the motor pathways.
  
  The interest in establishing a model of the motor pathways, at the same time mathematically rigorous and biologically plausible, comes from the possible spillovers in neurophysiology. It could help to design better control strategies for robots and artificial limbs, rendering them capable to move more progressively and smoothly and also to react to exterior perturbations in a flexible way. An underlying relevant societal goal (clearly beyond our domain of expertise) is to clarify the mechanisms of certain debilitating troubles such as cerebellar disease, chorea and Parkinson’s disease.

  A key issue in order to establish a model of the motor pathways is to determine the criteria underlying the brain’s choices. For instance, for the problem of human locomotion (see [22]), identifying such criteria would be crucial to understand the neural pathways implicated in the generation of locomotion trajectories.

  A nowadays widely accepted paradigm is that, among all possible movements, the accomplished ones satisfy suitable optimality criteria (see [87] for a review). One is then led to study an inverse optimal control problem: starting from a database of experimentally recorded movements, identify a cost function such that the corresponding optimal solutions are compatible with the observed behaviors.

  Different methods have been taken into account in the literature to tackle this kind of problems, for instance in the linear quadratic case [56] or for Markov processes [75]. However all these methods have been conceived for very specific systems and they are not suitable in the general case. Two approaches are possible to overcome this difficulty. The direct approach consists in choosing a cost function among a class of functions naturally adapted to the dynamics (such as energy functions) and to compare the solutions of the corresponding optimal control problem to the experimental data. In particular one needs to compute, numerically or analytically, the optimal trajectories and to choose suitable criteria (quantitative and qualitative) for the comparison with observed trajectories. The inverse approach consists in deriving the cost function from the qualitative analysis of the data.

4.3. Switched systems

Switched systems form a subclass of hybrid systems, which themselves constitute a key growth area in automation and communication technologies with a broad range of applications. Existing and emerging areas include automotive and transportation industry, energy management and factory automation. The notion of hybrid systems provides a framework adapted to the description of the heterogeneous aspects related to the interaction of continuous dynamics (physical system) and discrete/logical components.

The characterizing feature of switched systems is the collective aspect of the dynamics. A typical question is that of stability, in which one wants to determine whether a dynamical system whose evolution is influenced by a time-dependent signal is uniformly stable with respect to all signals in a fixed class ([63]).
The theory of finite-dimensional hybrid and switched systems has been the subject of intensive research in the last decade and a large number of diverse and challenging problems such as stabilizability, observability, optimal control and synchronization have been investigated (see for instance [85], [64]).

The question of stability, in particular, because of its relevance for applications, has spurred a rich literature. Important contributions concern the notion of common Lyapunov function: when there exists a Lyapunov function that decays along all possible modes of the system (that is, for every possible constant value of the signal), then the system is uniformly asymptotically stable. Conversely, if the system is stable uniformly with respect to all signals switching in an arbitrary way, then a common Lyapunov function exists [65]. In the linear finite-dimensional case, the existence of a common Lyapunov function is actually equivalent to the global uniform exponential stability of the system [71] and, provided that the admissible modes are finitely many, the Lyapunov function can be taken polyhedral or polynomial [29], [30], [44]. A special role in the switched control literature has been played by common quadratic Lyapunov functions, since their existence can be tested rather efficiently (see [45] and references therein). Algebraic approaches to prove the stability of switched systems under arbitrary switching, not relying on Lyapunov techniques, have been proposed in [62], [16].

Other interesting issues concerning the stability of switched systems arise when, instead of considering arbitrary switching, one restricts the class of admissible signals, by imposing, for instance, a dwell time constraint [50].

Another rich area of research concerns discrete-time switched systems, where new intriguing phenomena appear, preventing the algebraic characterization of stability even for small dimensions of the state space [59]. It is known that, in this context, stability cannot be tested on periodic signals alone [32].

Finally, let us mention that little is known about infinite-dimensional switched system, with the exception of some results on uniform asymptotic stability ([68], [82], [83]) and some recent papers on optimal control ([49], [90]).
4. Application Domains

4.1. Geometric Modeling and Shape Reconstruction

Modeling 3D shapes is required for all visualization applications where interactivity is a key feature since the observer can change the viewpoint and get an immediate feedback. This interactivity enhances the descriptive power of the medium significantly. For example, visualization of complex molecules helps drug designers to understand their structure. Multimedia applications also involve interactive visualization and include e-commerce (companies can present their products realistically), 3D games, animation and special effects in motion pictures. The uses of geometric modeling also cover the spectrum of engineering, computer-aided design and manufacture applications (CAD/CAM). More and more stages of the industrial development and production pipeline are now performed by simulation, due to the increased performance of numerical simulation packages. Geometric modeling therefore plays an increasingly important role in this area. Another emerging application of geometric modeling with high impact is medical visualization and simulation.

In a broad sense, shape reconstruction consists of creating digital models of real objects from points. Example application areas where such a process is involved are Computer Aided Geometric Design (making a car model from a clay mockup), medical imaging (reconstructing an organ from medical data), geology (modeling underground strata from seismic data), or cultural heritage projects (making models of ancient and or fragile models or places). The availability of accurate and fast scanning devices has also made the reproduction of real objects more effective such that additional fields of applications are coming into reach. The members of GEOMETRICA have a long experience in shape reconstruction and contributed several original methods based upon the Delaunay and Voronoi diagrams.

4.2. Scientific Computing

Meshes are the basic tools for scientific computing using finite element methods. Unstructured meshes are used to discretize domains bounded by complex shapes while allowing local refinements. GEOMETRICA contributes to mesh generation of 2D and 3D possibly curved domains. Most of our methods are based upon Delaunay triangulations, Voronoi diagrams and their variants. Anisotropic meshes are also investigated. We investigate in parallel both greedy and variational mesh generation techniques. The greedy algorithms consist of inserting vertices in an initial coarse mesh using the Delaunay refinement paradigm, while the variational algorithms consists of minimizing an energy related to the shape and size of the elements. Our goal is to show the complementarity of these two paradigms. Quadrangle surface meshes are also of interest for reverse engineering and geometry processing applications. Our goal is to control the final edge alignment, the mesh sizing and the regularity of the quadrangle tiling.
4. Application Domains

4.1. Building a Large Scale Distributed System for Computing

The main application domain of the Large Scale Distributed System developed in Grand-Large is high performance computing. The two main programming models associated with our platform (RPC and MPI) allow to program a large variety of distributed/parallel algorithms following computational paradigms like bag of tasks, parameter sweep, workflow, dataflow, master worker, recursive exploration with RPC, and SPMD with MPI. The RPC programming model can be used to execute concurrently different applications codes, the same application code with different parameters and library function codes. In all these cases, there is no need to change the code. The code must only be compiled for the target execution environment. LSDS are particularly useful for users having large computational needs. They could typically be used in Research and Development departments of Pharmacology, Aerospace, Automotive, Electronics, Petroleum, Energy, Meteorology industries. LSDS can also be used for other purposes than CPU intensive applications. Other resources of the connected PCs can be used like their memory, disc space and networking capacities. A Large Scale Distributed System like XtremWeb can typically be used to harness and coordinated the usage of these resources. In that case XtremWeb deploys on Workers services dedicated to provide and manage a disc space and the network connection. The storage service can be used for large scale distributed fault tolerant storage and distributed storage of very large files. The networking service can be used for server tests in real life conditions (workers deployed on Internet are coordinated to stress a web server) and for networking infrastructure tests in real like conditions (workers of known characteristics are coordinated to stress the network infrastructure between them).

4.2. Security and Reliability of Network Control Protocols

The main application domain for self-stabilizing and secure algorithms is LSDS where correct behaviours must be recovered within finite time. Typically, in a LSDS (such as a high performance computing system), a protocol is used to control the system, submit requests, retrieve results, and ensure that calculus is carried out accordingly to its specification. Yet, since the scale of the system is large, it is likely that nodes fail while the application is executing. While nodes that actually perform the calculus can fail unpredictably, a self-stabilizing and secure control protocol ensures that a user submitting a request will obtain the corresponding result within (presumably small) finite time. Examples of LSDS where self-stabilizing and secure algorithms are used, include global computing platforms, or peer to peer file sharing systems. Another application domain is routing protocols, which are used to carry out information between nodes that are not directly connected. Routing should be understood here in its most general acceptance, e.g. at the network level (Internet routing) or at the application level (on virtual topologies that are built on top of regular topologies in peer to peer systems). Since the topology (actual or virtual) evolves quickly through time, self-stabilization ensures that the routing protocol eventually provides accurate information. However, for the protocol to be useful, it is necessary that it provides extra guarantees either on the stabilization time (to recover quickly from failures) or on the routing time of messages sent when many faults occur. Finally, additional applications can be found in distributed systems that are composed of many autonomous agents that are able to communicate only to a limited set of nodes (due to geographical or power consumption constraints), and whose environment is evolving rapidly. Examples of such systems are wireless sensor networks (that are typically large of 10000+ nodes), mobile autonomous robots, etc. It is completely unrealistic to use centralized control on such networks because they are intrinsically distributed; still strong coordination is required to provide efficient use of resources (bandwidth, battery, etc).
4.3. End-User Tools for Computational Science and Engineering

Another Grand Large application domain is Linear Algebra, which is often required to solve Large Scale Computational Science and Engineering applications. Two main approaches are proposed. First, we have to experiment and evaluate several classical stable numerical methods. Second, we have to propose tools to help end-users to develop such methods.

In addition to the classical supercomputing and the GRID computing, the large scale P2P approach proposes new computing facilities for computational scientists and engineers. Thus, it exists many applications which would use such computing facilities for long period of time. During a first period, many applications will be based on large simulations rather than classical implicit numerical methods, which are more difficult to adapt for such large problems and new programming paradigm as they generated linear algebra problems. Then, implicit method would be developed to have more accurate solutions.

Simulations and large implicit methods always have to compute linear algebra routines. So, they were our first targeted numerical methods (we also remark that the powerful worldwide computing facilities are still rated using a linear algebra benchmark http://www.top500.org). We especially focused on divide-and-conquer and block-based matrix methods to solve dense problems. We have also studied Krylov subspace methods (Lanczos, Arnoldi) and hybrid methods to solve sparse matrix problems. As these applications are utilized for many applications, it is possible to extrapolate the results to different scientific domains.

Many smart tools have to be developed to help the end-user to program such environments, using up-to-date component technologies and languages. At the actual present stage of maturity of this programming paradigm for scientific applications, the main goal is to experiment on large platforms, to evaluate and extrapolate performance, and to propose tools for the end-users; with respect to many parameters and under some specific hypothesis concerning scheduling strategies and multicast speeds [78]. We have to always replace the end-user at the center of this scientific programming. Then, we have to propose a framework to program P2P architectures which completely virtualizes the P2P middleware and the heterogeneous hardware. Our approach is based, on the one hand, on component and coordination languages, and on the other hand, to the development of an ASP, which may be dedicated to a targeted scientific domain. The YML framework provides a solution to the first point since it offers the YvetteML workflow language in order to orchestrate components. This is a very intuitive programming approach and it favors the re-usability of optimized and bug-free components. The abstraction of the underlying P2P middleware is also ensured by YML by means of its back-end mechanism. The end-user of YML can submit a computing task to any kind of peer connected to Internet as long as YML has a back-end in charge of the middleware which is running on this peer. Currently, YML has two back-ends for the XtremWeb and OmniRPC middleware. Another one for Condor will be soon available. The second point concerns the integration of SPIN to YML in order to get a complete programming tool which covers all the needs of the client in order to run applications (based on linear algebra methods) over the Internet. Finally, the conclusion of our work would be a P2P scientific programming methodology based on experimentations and evaluation on an actual P2P development environment.

4.4. Numerical simulations and other intensive applications

For the research on numerical linear algebra, the final goal of the research of Marc Baboulin and Laura Grigori is to make the algorithms and software as generic and widely usable as possible. However, through our academic and industrial collaborators, our research focuses on several important application domains, that we briefly describe in the following.

- Simulation of compositional multiphase Darcy flow in heterogeneous porous media with different type of applications: simulation of reservoir models, simulation of basin models, simulation of geological CO2 underground storage, simulation of underground nuclear waste disposal. This research is performed in the context of the ANR Petal and Petalh projects, in collaboration with partners from IFP and CEA.
- Data analysis in astrophysics: we focus on computationally intensive numerical algorithms arising in the data analysis of current and forthcoming Cosmic Microwave Background (CMB) experiments in
astrophysics. While this application does not involve a PDE, its most complex and time consuming step is solving a generalized least squares problem, which is at the core of our research. This research is performed in the context of the ANR Midas project in collaboration with Paris 7.

- Numerical simulations of incompressible fluid flows: We address the solution of large sparse linear systems coming from the discretization of Helmholtz and Poisson equations that represent the major part of the computational time for solving the Navier-Stokes equations describing a large class of fluid flows. For this application, the objective is to develop algorithms that take advantage of current heterogeneous multicore/GPU architectures. These algorithms must minimize the amount of communication in the linear algebra kernels involved in the computation (e.g. sparse matrix-vector product) and in the choice of the preconditioners. This research is a multidisciplinary collaboration between researchers from University Paris-Sud, INRIA-Saclay and LIMSI/CNRS in the framework of the CALIFHA project (funded by DIM/Digitéo).
4. Application Domains

4.1. Wireless mobile ad hoc networks

Abstract. Mobile wireless networks have numerous applications in rescue and emergency operation, military tactical networking and in wireless high speed access to the internet.

A mobile ad hoc network is a network made of a collection of mobile nodes that gather spontaneously and communicate without requiring a pre-existing infrastructure. Of course a mobile ad hoc network use a wireless communication medium. They can be applied in various contexts:

- military;
- rescue and emergency;
- high speed access to internet.

The military context is the most obvious application of mobile ad hoc networks. Soldiers invading a country won’t subscribe in advance to the local operator. On the reverse side, home units won’t use their local operators firstly because they will likely be disrupted in the first hours of the conflict, and secondly because a wireless communication via an operator is not stealth enough to protect the data and the units. In Checheny, a general has been killed by a missile tracking the uplink signal of his portable phone.

The rescue context is halfway between military and civilian applications. In the september 11 disaster, most of the phone base station of the area have knocked out in less than twenty minutes. The remaining base stations were unable to operate because they could not work in ad hoc mode. The Wireless Emergency Rescue Team recommended afterward that telecom operators should provide ad hoc mode for their infrastructure in order to operate in emergency situation in plain cooperation with police, firemen and hospital networks.

Mobile ad hoc network provide an enhanced coverage for high speed wireless access to the internet. The now very popular WLAN standard, WiFi, provides much larger capacity than mobile operator networks. Using a mobile ad hoc network around hot spots will offer high speed access to much larger community, including cars, busses, trains and pedestrians.

4.2. Services over mobile networks

Abstract. New wireless network calls for new services that fulfill the requirement in terms of mobility and capacity.

The generalization of a new generation of mobile networks calls for a new set of services and applications. For example:

- Indoor and outdoor positioning
- Service discovery and localisation
- Multicast and quality of services

Quality of service has become the central requirement that users expect from a network. High throughput, service continuity are critical issue for multimedia application over the wireless internet where the bandwidth is more scarce than in the wired world. A significant issue in the ad-hoc domain is that of the integrity of the network itself. Routing protocols allow, according to their specifications, any node to participate in the network - the assumption being that all nodes are behaving well and welcome. If that assumption fails - then the network may be subject to malicious nodes, and the integrity of the network fails. An important security service over mobile networks is to ensure that the integrity of the network is preserved even when attacks are launched against the integrity of the network.
4.3. Community Network

Abstract. There is an increasing demand to deploy network within a community, rural or urban, with cabled or wireless access.

Community networks or citizen network are now frequent in big cities. In America most of the main cities have a community network. A community network is using the communication resource of each member (ADSL, Cable and wireless) to provide a general coverage of a city. Pedestrian in the street or in city mails can communicate via a high speed mobile mesh network. This new trend now appears in Europe with many experiments of the OLSR routing protocol in Paris, Lille, Toulouse, Berlin, Brussels, Seattle. The management of such networks is completely distributed and makes them very robust to faults. There is room for smart operators in this business.

4.4. Vehicular Networks

Abstract. Intelligent transport systems require efficient wireless telecommunications.

Vehicular ad hoc networks (VANET) are based on short- to medium-range transmission systems that support both vehicle-to-vehicle and vehicle-to-roadside communications. Vehicular networks will enable vehicular safety applications (safety warnings) as well as non-safety applications (real-time traffic information, routing support, mobile entertainment, and many others). We are interested in developing an efficient routing protocol that takes advantage of the fixed network infrastructure deployed along the roads. We are also studying MAC layer issues in order to provide more priority for security messages which have stringent delivery constraints.

4.5. Large ad hoc networks with sensor nodes

Abstract. Large autonomous wireless sensors in the internet of the things need very well tuned algorithms.

Self-organization is considered as a key element in tomorrow’s Internet architecture. A major challenge concerning the integration of self-organized networks in the Internet is the accomplishment of light weight network protocols in large ad hoc environments.

In this domain, Hipercom’s activity with wireless sensor nodes in collaboration with the Freie Universitaet in Berlin explores various solutions, including extensions of OLSR (for example DHT-OLSR) using programmable sensor nodes co-designed by the Freie Universitaet, and provides one of the largest testbeds of this kind, to date.

4.6. Energy efficient wireless sensor networks

Abstract. Energy efficiency is a key property in wireless sensor networks.

Various techniques are used to contribute to energy efficiency, see for instance the survey we published at the WMNC 2011 conference. In the OCARI2 project, we have designed and implemented an energy efficient routing protocol and a node activity scheduling algorithm allowing router nodes to sleep. We have applied a cross-layering approach allowing the optimization of MAC and network protocols taking into account the application requirements and the environment in which the network operates. This activity has been done in collaboration with our partners EDF, LIMOS and TELIT.
IN-SITU Project-Team

4. Application Domains

4.1. Application Domains

INSITU works actively with users from various application domains in order to understand their specific needs. By studying similar problems in different domains, we can generalize our results and develop more general principles. Our current application domains include:

- Scientific discovery, i.e. the use of advanced interactive technologies by scientists of other disciplines, in particular:
  - Biological research, in cooperation with the Institut Pasteur (Paris), INRA (Institut National de la Recherche Agronomique, Evry), INRA Metarisk\(^1\) (Paris), and other laboratories of the University Paris-Sud;
  - Astronomy, in cooperation with the European Southern Observatory on the ALMA project\(^2\) (Atacama Large Millimiter/submillimeter Array), for array operations monitoring and control of radiotelescopes; and with Institut d’Astrophysique Spatiale\(^3\) on the visualization of large astronomy imagery using ultra-high-resolution wall-sized displays;

- Creative industries (music composition), in cooperation with IRCAM (Institut de Recherche et Coordination Acoustique-Musique, Paris);

- Domestic technologies, in cooperation with ENSCI (Ecole Nationale Supérieure de Création Industrielle, Paris).

We have selected these domains to ensure that we explore and address diverse validation criteria, e.g. enhancing productivity versus increasing communication access, diverse user characteristics, e.g. professionals versus non-professionals, and diverse user environments, e.g. desktops at work versus home versus mobile settings.

\(^1\) http://www.paris.inra.fr/metarisk
\(^2\) http://www.almaobservatory.org/
\(^3\) http://www.ias.u-psud.fr
4. Application Domains

4.1. Application Domains

The typical IT projects to which our technologies contribute aim at efficient and flexible management of complex digital information. The form and nature of the data often varies: Web pages, Office or PDF documents, XML structured data (sometimes obtained through Web service gateways), thesauri, ontologies etc. From such heterogeneous, complex resources, interested parties aim at building storage and processing tools, enabling the efficient storage, classification, annotation, enrichment, and fine-grained search on such data. Sample real-life applications that we have already worked on in this setting are:

- Archiving filtered content from online information sources (journals, blogs, ...) with the purpose of recording their perspective on facts involving specific countries, key political actors etc. (EADS data gathering for intelligence purposes, also an application from the WebContent project)
- Building an XML data warehouse out of public e-mails exchanged in a technical standardization body (in our particular case, the W3C) in order to enable a fine-grained social network analysis to determine key players, opinion leaders etc.
- Building a complete processing chain for digital documents from the medical domain. The process may start with the digitization and text extraction from scanned documents (we does not work in this area), then continues with extraction of named entities, document annotation based on existing domain ontologies, mapping of documents to a central domain ontology, reasoning across scattered data sources for query answering, storing, indexing, and distributing the data (and query results) across distributed players.
- Data produced and made public by numerous public administration offices (in France, Europe, and the world) opens many perspectives for integrating, analyzing, and combining data sources into added-value information sources. Time is also an essential dimension here; so is data matching and reconciliation, since the same entity may be referenced from many different viewpoints and reconciliation is needed when joining data sources. Users of such applications could be public administration analyzing the impact of its policies, social scientists and journalists which already work on the data (but gather it with much difficulty) etc. This applications is gathered from our collaboration with the DataPublica start-up.

Interesting areas of content management, which we do not address, are: audio and video data, natural language processing, data mining, access control and privacy. We collaborate up with other groups specialized in these topics.
4. Application Domains

4.1. Systèmes à événements discrets (productique, réseaux)/Discrete event systems (manufacturing systems, networks)

One important part of applications of max-plus algebra comes from discrete event dynamical systems [6]. Max-plus linear systems, and more generally, monotone nonexpansive dynamical systems, provide natural models for which many analytical results can be applied to performance evaluation problems. For instance, problems like computing the cycle time of asynchronous digital circuits [94], or computing the throughput of a workshop [138] or of a transportation network, and performance evaluation problems for communication networks, are often amenable to max-plus algebra, at least in some simplified form, see in particular [93] and [85]. The max-plus approach has been applied to the analysis of the time behaviour of concurrent systems, and in particular, to the analysis of high level sequence message charts [89], [146]. The Maxplus team collaborates with the Metalau team, working particularly on the applications of max-plus models to the microscopic modelling of road traffic [152], [150], [119].

4.2. Commande optimale et jeux/Optimal control and games

Optimal control and game theory have numerous well established applications fiels: mathematical economy and finance, stock optimization, optimization of networks, decision making, etc. In particular, the Mathfi team works on applications in mathematical finance. There is a tradition of collaboration between researchers of the Maxplus team and of the Mathfi team on these questions, see as an illustration [5] where ideas from the spectral theory of monotone homogeneous maps [3] are applied.

4.3. Recherche opérationnelle/Operations research

L’algèbre max-plus intervient de plusieurs manières en Recherche opérationnelle. Premièrement, il existe des liens profonds entre l’algèbre max-plus et les problèmes d’optimisation discrète, voir [95]. Ces liens conduisent parfois à de nouveaux algorithmes pour les problèmes de recherche opérationnelle classiques,
comme le problème de circuit de poids moyen maximum [104]. Certains problèmes combinatoires, comme des problèmes de programmation disjonctive, peuvent être décomposés par des méthodes de type max-plus [184]. Ensuite, le rôle de l’algèbre max-plus dans les problèmes d’ordonnancement est bien connu depuis les années 60, les dates de complétion pouvant souvent être calculées à partir d’équations linéaires max-plus. Plus récemment, des représentations de problèmes d’ordonnancement ont pu être obtenues à partir de semi-groupes de matrices max-plus : une première représentation a été obtenue dans [126] pour le cas du “jobshop”, une représentation plus simple a été obtenue dans [147] dans le cas du “flowshop”. Ce point de vue algébrique a été très utile dans le cas du “flowshop” : il permet de retrouver des résultats anciens de dominance et d’obtenir ainsi de nouvelles bornes [147]. Finalement, en regardant l’algèbre max-plus comme une limite de l’algèbre classique, on peut utiliser des outils algébriques en optimisation combinatoire [144].

**English version**

Max-plus algebra arise in several ways in Operations Research. First, there are intimate relations between max-plus algebra and discrete optimisation problems, see [95]. Sometimes, these relations lead to new algorithms for classical Operations Research problems, like the maximal circuit mean [104]. There are also special combinatorial problems, like certain problems of disjunctive programming, which can be decomposed by max-plus type methods [184]. Next, the role of max-plus algebra in scheduling problems has been known since the sixties: completion dates can often be computed by max-plus linear equations. Recently, representations of certain scheduling problems using max-plus matrix semigroups have appeared, a first representation was given in [126] for the jobshop case, a simpler representation was given in [147] in the flowshop case. This algebraic point of view turned out to be particularly fruitful in the flowshop case: it allows one to recover old dominance results and to obtain new bounds [147]. Finally, viewing max-plus algebra as a limit of classical algebra allows to use algebraic tools in combinatorial optimisation [144].

4.4. Analyse statique de programmes/Static analysis of computer programs

L’interprétation abstraite est une technique, introduite par P. et R. Cousot [108], qui permet de déterminer des invariants de programmes en calculant des points fixes minimaux d’applications monotones définies sur certains treillis. On associe en effet à chaque point de contrôle du programme un élément du treillis, qui représente une sur-approximation valide de l’ensemble des valeurs pouvant être prises par les variables du programme en ce point. Le treillis le plus simple exprimant des propriétés numériques est celui des produits Cartésiens d’intervalles. Des treillis plus riches permettent de mieux tenir compte de relations entre variables, en particulier, des classes particulières de polyèdres sont souvent employées.

Voici, en guise d’illustration, un petit exemple de programme, avec le système de point fixe associé, pour le treillis des intervalles:

```c
void main() {
    int x=0; // 1
    while (x<100) { // 2
        x=x+1; // 3
        }
        // 4
    }
```

Si l’on s’intéresse par exemple aux valeurs maximales prise par la variable $x$ au point de contrôle 2, soit $x_2^+ := \max x_2$, après une élimination, on parvient au problème de point fixe:

$$x_2^+ = \min (99, \max (0, x_2^+ + 1)),$$  \hfill (1)

qui a pour plus petite solution $x_2^+ = 99$, ce qui prouve que $x$ est majoré par 99 au point 2.
On reconnaît ici un opérateur de point fixe associé à un problème de jeux à deux joueurs et somme nulle. Cette analogie est en fait générale, dans le cadre d’un collaboration que l’équipe entretient depuis plusieurs années avec l’équipe MeASI d’Eric Goubault (CEA et LIX), spécialiste d’analyse statique, nous avons en effet mis progressivement en évidence une correspondance [107], [123], entre les problèmes de jeux à somme nulle et les problèmes d’analyse statique, qui peut se résumer par le dictionnaire suivant:

<table>
<thead>
<tr>
<th>Jeux</th>
<th>Interprétation abstraite</th>
</tr>
</thead>
<tbody>
<tr>
<td>système dynamique</td>
<td>programme</td>
</tr>
<tr>
<td>opérateur de Shapley</td>
<td>fonctionnelle</td>
</tr>
<tr>
<td>espace d’état</td>
<td>(# points de contrôle) × (# degrés de liberté du treillis)</td>
</tr>
<tr>
<td>problème en horizon $n$</td>
<td>exécution de $n$ pas</td>
</tr>
<tr>
<td>limite du problème en horizon fini</td>
<td>invariant optimal (borne)</td>
</tr>
<tr>
<td>itération sur les valeurs</td>
<td>itération de Kleene</td>
</tr>
</tbody>
</table>

Pour que le nombre d’états du jeu soit fini, il est nécessaire de se limiter à des treillis d’ensembles ayant un nombre fini de degrés de liberté, ce qui est le cas de domaines communément utilisés (intervalles, ensembles définis par des contraintes de potentiel de type $x_i - x_j \leq \text{cst}$, mais aussi, les “templates” qui sont des sous-classes de polyèdres introduits récemment par Sankaranarayanan, Sipma et Manna [175]). L’ensemble des actions est alors fini si on se limite à une arithmétique affine. Signalons cependant qu’en toute généralité, on aboutit à des jeux avec un taux d’escompte négatif, ce qui pose des difficultés inédites. Cette correspondance entre jeux et analyse statique est non intuitive, au sens où les actions du minimiseur consistent à sélectionner des points extrêmes de certains polyèdres obtenus par un mécanisme de dualité.

Une pathologie bien répertoriée en analyse statique est la lenteur des algorithmes de point fixe, qui peuvent effectuer un nombre d’itérations considérable (99 itérations pour obtenir le plus petit point fixe de (8)). Celle-ci est usuellement traitée par des méthodes d’accélération de convergence dites d’élargissement et rétrécissement [109], qui ont cependant l’inconvénient de conduire à une perte de précision des invariants obtenus. Nous avons exploité la correspondance entre analyse statique et jeux pour développer des algorithmes d’une nature très différente, s’inspirant de nos travaux antérieurs sur l’itération sur les politiques pour les jeux répétés [124], [102], [103], [7]. Une version assez générale de cet algorithme, adaptée au domaine des templates, est décrite dans [123] et a fait l’objet d’une implémentation prototype. Chaque itération combine de la programmation linéaire et des algorithmes de graphes. Des résultats expérimentaux ont montré le caractère effectif de la méthode, avec souvent un gain en précision par rapport aux approches classiques, par exemple pour des programmes comprenant des boucles imbriquées.

Ce domaine se trouve être en pleine évolution, un enjeu actuel étant de traiter d’une manière qui passe à l’échelle des invariants plus précis, y compris dans des situations où l’arithmétique n’est plus affine.

**English version**

The abstract interpretation method introduced by P. and R. Cousot [108], allows one to determine automatically invariants of programs by computing the minimal fixed point of an order preserving map defined on a complete lattice. To every breakpoint of the program is associated an element of the lattice, which yields a valid overapproximation of the set of reachable values of the vectors of variables of the program, at this breakpoint. The simplest lattice expressing numerical invariants consists of Cartesian products of intervals. More sophisticated lattices, taking into account relations between variables, consisting in particular of subclasses of polyhedra, are often used.

As an illustration, we gave before Eqn (8) a simple example of program, together with the associated fixed-point equation. In this example, the value of the variable $x$ at the breakpoint 2 is bounded by the smallest solution $x^2_j$ of the fixed point problem (8), which is equal to 99.

The fixed point equation (8) is similar to the one arising in the theory of zero-sum repeated games. This analogy turns out to be general. Un a series of joint works of our team with the MeASI team of Eric Goubault (CEA and LIX), we brought progressively to light a correspondence [107], [123], between the zero-sum game problems and the static analysis problems, which can be summarized by the following dictionary:
Games
- dynamical system
- Shapley operator
- state space
- horizon $n$ problem
- limit of the value in horizon $n$

Abstract interpretation
- program
- functional
- (# breakpoints) $\times$ (# degrees of freedom)
- execution of $n$ logical steps
- optimal invariant (bound)
- Kleene iteration

For the game to have a finite state space, we must restrict our attention to lattices of sets with a finite number of degrees of freedom, which is the case of the domains commonly used in static analysis (intervals, sets defined by potentials constraints of the form $x_i - x_j \leq \text{cst}$, and also the subclasses of polyhedra called “templates”, introduced recently by Sankaranarayanan, Sipma and Manna [175]). Then, the action space is finite if the arithmetics of the program is affine. However, in full generality, the games we end up with have a negative discount rate, which raises difficulties which are unfamiliar from the game theory point of view.

This correspondence between games and static analysis turns out to be non intuitive, in that the action of the minimizer consist of selecting an extreme point of a polyhedron arising from a certain duality construction.

A well known pathology in static analysis is the fact that the standard Kleene fixed point algorithm may have a very slow behavior (99 iterations are needed to get the smallest fixed point of (8)). This is usually solved by using some accelerations of convergence, called widening and narrowing [109], which however lead to a loss of precision. We exploited the correspondence between static analysis and games to develop algorithms of a very different nature, inspired by our earlier work on policy iteration for games [124], [102], [103], [7]. A rather general version of this policy iteration algorithm, adapted to the domain of templates, is described in [123], together with a prototype implementation. Every iteration combines linear programming and combinatorial algorithms. Some experimental results indicate that the method often leads to invariants which are more accurate than the ones obtained by alternative methods, in particular for some programs with nested loops.

This topic of research is currently evolving, a question of current interest being to find accurate invariants, in a scalable way, in situations in which the arithmetics is not affine.

4.5. Autres applications/Other applications

L’algèbre max-plus apparaît de manière naturelle dans le calcul de scores de similitudes dans la comparaison de séquences génétiques. Voir par exemple [106].

*English version*

Max-plus algebra arises naturally in the computation of similarity scores, in biological sequence comparison. See for instance [106].
4. Application Domains

4.1. Panorama

MExICo’s research is motivated by problems on system management in several domains:

- In the domain of service oriented computing, it is often necessary to insert some Web service into an existing orchestrated business process, e.g. to replace another component after failures. This requires to ensure, often actively, conformance to the interaction protocol. One therefore needs to synthesize *adaptators* for every component in order to steer its interaction with the surrounding processes.

- Still in the domain of telecommunications, the supervision of a network tends to move from out-of-band technology, with a fixed dedicated supervision infrastructure, to in-band supervision where the supervision process uses the supervised network itself. This new setting requires to revisit the existing supervision techniques using control and diagnosis tools.

- Several recent Intelligent Transport Systems projects aim at providing assistance to drivers, in the way of (partially) automated motorways. We will focus on the modeling and analysis of the collision avoidance problems in critical short sections of motorways.

This list is likely to grow over the next years as we continue our research.

4.2. Autonomous Telecommunications Systems: In-Band Supervision

Participants: Stefan Haar, Serge Haddad.

In the context of traditional hard-wired communication networks, supervision structures for managing faults, configuration, provisioning etc could be developed with a fixed infrastructure, and perform the communication between sensors, supervisors, policy enforcement points etc over a separate network using separate hardware. This rigid, out-of-band technology does not survive passing to today’s and tomorrow’s services and networks. In fact, the dynamic mobility of services combined across sites and domains cannot be captured unless the network used for supervision evolves in the same way and simultaneously, which rules out static solutions; but providing out-of-band infrastructure that grows with the networks to be supervised would be prohibitively expensive, if at all technically feasible. Heterogeneity is the other feature of modern networks that forces a change, since different domains are not likely to agree on a pervasive third-party supervision. Rather, the providers will keep control over the internal state and evolution of their domain, and accept only exchange through standardized outward interfaces.

Supervision has thus to be re-invented on an in-band, autonomous base: monitoring probes deployed on the web, dysfunctions on one peer node diagnosed by another peer in a network with changing configuration, enhanced supervisor and actor capacities of services, etc. MExICo will work on improving the interoperability of service components through continued application of e.g. distributed techniques for control and diagnosis.

4.3. Traffic Safety Control

Participant: Serge Haddad.

The Intelligent Transport Systems (ITS) community tries to deal with the numerous challenges that arise when designing secure and reliable software dedicated to automatic transport systems. Several recent ITS projects aim at providing assistance to drivers and deal with partially automated motorways. The community investigated first a fully automated infrastructure and vehicles approach (as in the PATH project [93]) in the 1990's. That approach was then abandoned in favor of a new line of research and development activities, more centered on safety strategies to ensure properties such as Collision Avoidance or Safety Margin for Assistance Vehicles [94].
This vision relies on cooperative systems where “road operators, infrastructure, vehicles, their drivers and other road users will cooperate to deliver the most efficient, safe, secure and comfortable journeys” [73]. Implementing such a system then follows a peer-to-peer organization where each vehicle must fully cooperate in a time-constrained and safety-critical environment.

In that context, many projects are dealing with safety-oriented applications based on sensors, communication devices and protocols as well as distributed traffic management systems involving cooperation between the infrastructure and vehicles [72], [71], [105]. Thus, reliability, flexibility in the design as well as safety are primary issues. Such systems are even more complex to analyze than previous distributed systems. Consequently, there is a need for a specific methodology and tools to design and analyze them.

We will focus on an approach for the modeling and analysis of the collision avoidance problems in critical short sections of motorways with the aim to check whether a control strategy exists depending on the parameters (speed, safety distances, etc.). We intend to cope with the undecidability of such problems by appropriate discretizations and with high complexity of the obtained systems by using elaborated data structures based on decision diagrams.

4.4. Web Services

Participants: Stefan Haar, Serge Haddad.

Specific applications targeted by MExICo include the problem of adaptation in Service-Oriented Computing (SOC). The challenge is here twofold, stemming both from the distributed nature of services (scattered over the entire web) and their heterogeneous origins.

4.4.1. Context

Web services have become the most frequently used model of design and programming based on components for business applications. Web service languages like BPEL have useful constructors that manage for instance exceptions, (timed guarded) waiting of messages, parallel execution of processes, distant service invocations, etc. Interoperability of components is based on interaction protocols associated with them and often published on public or private registers. In the framework of Web services, these protocols are called abstract processes by contrast with business processes (i.e. services). Composition of components must be analyzed for several reasons and at least to avoid deadlocks during execution. This has led to numerous works that focus on compositional verification, substitution of a component by another one, synthesis of adaptators, etc., and triggered a push towards a unifying theoretical framework (see e.g. [102], [107]).

4.4.2. Problems

Interoperability requires that when a user or a program wants to interact with the component, the knowledge of the interaction protocol is enough. Our previous works have shown that the interaction protocols can be inherently ambiguous: no client can conduct a correct interaction with the component in every scenario. This problem is even more complex when the protocol can evolve during execution due to adaptation requirements. The composition of components also raises interesting problems. When composing optimal components (w.r.t. the number of states for instance) the global component can be non optimal. So one aims at reducing a posteriori or better on the fly the global component. At last, the dynamical insertion of a component in a business process requires to check whether this insertion is behaviorally consistent [109], [96].

We do not intend to check global properties based on a modular verification technique. Rather, given an interaction protocol per component and a global property to ensure, we want to synthesize an adaptator per component such that this property is fulfilled or to detect that there cannot exist such adaptators [92]. In another research direction, one can introduce the concept of utility of a service and then optimize a system i.e. keeping the same utility value while reducing the resources (states, transitions, clocks, etc.).
PARIETAL Project-Team

4. Application Domains

4.1. Application Domains

- Multi-modal brain image registration for the estimation of brain templates.
- Segmentation and dictionary learning techniques for the creation of functional brain atlases.
- Detection of statistical association between the genetic variability and brain characteristics.
- Detection of abnormal data in neuroimaging datasets and robust statistics.
- Evaluation of neuro-computational of vision based on functional neuroimaging experiments.
- Inference of brain states or cognitive variables based on activation patterns.
- Extraction of biomarkers from functional connectivity data for neurodegenerative diseases.
4. Application Domains

4.1. Automated theorem proving

Automated theorem proving has traditionally focused on classical first-order logic, but non-classical logics are increasingly becoming important in the specification and analysis of software. Most type systems are based on (possibly second-order) propositional intuitionistic logic, for example, while resource-sensitive and concurrent systems are most naturally expressed in linear logic.

The members of the Parsifal team have a strong expertise in the design and implementation of performant automated reasoning systems for such non-classical logics. In particular, the Linprover suite of provers \[34\] continue to be the fastest automated theorem provers for propositional and first-order linear logic.

Any non-trivial specification, of course, will involve theorems that are simply too complicated to prove automatically. It is therefore important to design semi-automated systems that allow the user to give high level guidance, while at the same time not having to write every detail of the formal proofs. High level proof languages in fact serve a dual function – they are more readily comprehended by human readers, and they tend to be more robust with respect to maintenance and continued evolution of the systems. Members of the Parsifal team, in association with other INRIA teams and Microsoft Research, have been building a heterogeneous semi-automatic proof system for verifying distributed algorithms \[36\].

On a more foundational level, the team has been developing many new insights into the structure of proofs and the proof search spaces. Two directions, in particular, present tantalizing possibilities:

- The concept of multi-focusing \[37\] can be used to expose concurrency in computational behavior, which can in turn be exploited to prune areas of the proof search space that explore irrelevant interleavings of concurrent actions.
- The use of bounded search, where the bounds can be shown to be complete by meta-theoretic analysis, can be used to circumvent much of the non-determinism inherent in resource-sensitive logics such as linear logic. The lack of proofs of a certain bound can then be used to justify the presence or absence of properties of the encoded computations.

Much of the theoretical work on automated reasoning has been motivated by examples and implementations, and the Parsifal team intends to continue to devote significant effort in these directions.

4.2. Mechanized metatheory

There has been increasing interest in the use of formal methods to provide proofs of properties of programs and programming languages. Tony Hoare’s Grand Challenge titled “Verified Software: Theories, Tools, Experiments” has as a goal the construction of “verifying compilers” for a world where programs would only be produced with machine-verified guarantees of adherence to specified behavior. Guarantees could be given in a number of ways: proof certificates being one possibility.

The POPLMark challenge \[29\] envisions “a world in which mechanically verified software is commonplace: a world in which theorem proving technology is used routinely by both software developers and programming language researchers alike.” The proposers of this challenge go on to say that a “crucial step towards achieving these goals is mechanized reasoning about language metatheory.”

The Parsifal team has developed several tools and techniques for reasoning about the meta-theory of programming languages. One of the most important requirements for programming languages is the ability to reason about data structures with binding constructs up to \(\alpha\)-equivalence. The use of higher-order syntax and nominal techniques for such data structures was pioneered by Miller, Nadathur and Tiu. The Abella system (see Section 3.2) implements a refinement of a number of these ideas and has been used to give full solutions to sections of the POPLMark challenge in addition to fully formal proofs of a number of other theorems in the meta-theory of the \(\lambda\)-calculus.
Now that the Abella system has been in circulation among colleagues during the past couple of years, there are many aspects of the methodology that now need to be addressed. During the summer of 2011, the team employed three interns Carnegie Mellon University and McGill University to work on different aspects of Abella. Particular focus was given to better ways to manipulate specification-logic contexts in the reasoning-logic and with finding ways to have Abella output a proper proof object (different from the scripts that are used to find a proof).

Our colleague Alwen Tiu from the Australian National University has also been building on our Bedwyr model checking tool so that we can build on top of it his SPEC system for doing model checking of spi-calculus expressions. We have adopted his enhancements to Bedwyr and are developing further improvements within the context of the BATT project (see Section 5.3).

4.3. Proof certificates

Members of the Parsifal team have shown how to specify a large variety of proof systems—including natural deduction, the sequent calculus, and various tableau and free deduction systems—uniformly using either focused linear logic \[51\], \[49\] or focused intuitionistic logic \[43\] as the meta-language. In the presence of induction and co-induction, arbitrary finite computations can be embedded into single synthetic steps \[30\]. Additional work \[20\] shows that this same framework can also capture resolution refutations as well as Pratt primality certificates.

An important application then of this work in designing synthetic inference systems based on classical and intuitionistic logic is that of designing a broad spectrum proof certificate. The definition of proof certificates can be remarkably flexible within the simple setting of focused proofs.

The most important implications of such a certificate format would be that most of the worlds theorem provers should be able to print out their proofs and communication them to other provers: these other provers could then check such certificates by expanding the synthetic connectives they contain down into a small and fixed set of “micro” inference rules.
4. Application Domains

4.1. Panorama

Many systems in telecommunication, banking or transportation involve sophisticated software for controlling critical operations. One major problem is to get a high-level of confidence in the algorithms or protocols that have been developed inside the companies or by partners.

Many smartcards in mobile phones are based on a (small) Java virtual machine. The card is supposed to execute applets that are loaded dynamically. The operating system itself is written in C, it implements security functions in order to preserve the integrity of data on the card or to offer authentication mechanisms. Applets are developed in Java, compiled, and then the byte-code is loaded and executed on the card. Applets or the operating systems are relatively small programs but they need to behave correctly and to be certified by an independent entity.

If the user expresses the expected behavior of the program as a formal specification, it is possible for a tool to check whether the program actually behaves according to the requirements.

Avionics or more generally transportation systems are another area were there are critical algorithms involved, for instance in Air Traffic control. We have collaborations in this domain with Dassault-Aviation and National Institute of Aerospace (NIA, Hampton, USA). Since 2011, we started a new collaboration with Mitsubishi Electric R&D Centre Europe (Rennes), on the construction of certified software for railroad transportation. We also recently started a collaboration with Adacore for a new environment for proving Ada source code, which has applications in transportation systems including aerospace.
4. Application Domains

4.1. Application: uncertainties management

Our theoretical works are motivated by and find natural applications to real-world problems in a general frame generally referred to as uncertainty management, that we describe now.

Since a few decades, modeling has gained an increasing part in complex systems design in various fields of industry such as automobile, aeronautics, energy, etc. Industrial design involves several levels of modeling: from behavioural models in preliminary design to finite-elements models aiming at representing sharply physical phenomena. Nowadays, the fundamental challenge of numerical simulation is in designing physical systems while saving the experimentation steps.

As an example, at the early stage of conception in aeronautics, numerical simulation aims at exploring the design parameters space and setting the global variables such that target performances are satisfied. This iterative procedure needs fast multiphysical models. These simplified models are usually calibrated using high-fidelity models or experiments. At each of these levels, modeling requires control of uncertainties due to simplifications of models, numerical errors, data imprecisions, variability of surrounding conditions, etc.

One dilemma in the design by numerical simulation is that many crucial choices are made very early, and thus when uncertainties are maximum, and that these choices have a fundamental impact on the final performances.

Classically, coping with this variability is achieved through model registration by experimenting and adding fixed margins to the model response. In view of technical and economical performance, it appears judicious to replace these fixed margins by a rigorous analysis and control of risk. This may be achieved through a probabilistic approach to uncertainties, that provides decision criteria adapted to the management of unpredictability inherent to design issues.

From the particular case of aircraft design emerge several general aspects of management of uncertainties in simulation. Probabilistic decision criteria, that translate decision making into mathematical/probabilistic terms, require the following three steps to be considered [50]:

1. build a probabilistic description of the fluctuations of the model’s parameters (Quantification of uncertainty sources),
2. deduce the implication of these distribution laws on the model’s response (Propagation of uncertainties),
3. and determine the specific influence of each uncertainty source on the model’s response variability (Sensitivity Analysis).

The previous analysis now constitutes the framework of a general study of uncertainties. It is used in industrial contexts where uncertainties can be represented by random variables (unknown temperature of an external surface, physical quantities of a given material, ... at a given fixed time). However, in order for the numerical models to describe with high fidelity a phenomenon, the relevant uncertainties must generally depend on time or space variables. Consequently, one has to tackle the following issues:

- How to capture the distribution law of time (or space) dependent parameters, without directly accessible data? The distribution of probability of the continuous time (or space) uncertainty sources must describe the links between variations at neighbor times (or points). The local and global regularity are important parameters of these laws, since it describes how the fluctuations at some time (or point) induce fluctuations at close times (or points). The continuous equations representing the studied phenomena should help to propose models for the law of the random fields. Let us notice that interactions between various levels of modeling might also be used to derive distributions of probability at the lowest one.
The navigation between the various natures of models needs a kind of metric which could mathematically describe the notion of granularity or fineness of the models. Of course, the local regularity will not be totally absent of this mathematical definition.

All the various levels of conception, preliminary design or high-fidelity modelling, require registrations by experimentation to reduce model errors. This calibration issue has been present in this frame since a long time, especially in a deterministic optimization context. The random modeling of uncertainty requires the definition of a systematic approach. The difficulty in this specific context is: statistical estimation with few data and estimation of a function with continuous variables using only discrete setting of values.

Moreover, a multi-physical context must be added to these questions. The complex system design is most often located at the interface between several disciplines. In that case, modeling relies on a coupling between several models for the various phenomena and design becomes a multidisciplinary optimization problem. In this uncertainty context, the real challenge turns robust optimization to manage technical and economical risks (risk for non-satisfaction of technical specifications, cost control).

We participate in the uncertainties community through several collaborative research projects (ANR and Pôle SYSTEM@TIC), and also through our involvement in the MASCOT-NUM research group (GDR of CNRS). In addition, we are considering probabilistic models as phenomenological models to cope with uncertainties in the DIGITEO ANIFRAC project. As explained above, we focus on essentially irregular phenomena, for which irregularity is a relevant quantity to capture the variability (e.g. certain biomedical signals, terrain modeling, financial data, etc.). These will be modeled through stochastic processes with prescribed regularity.

4.2. Design of complex systems

Figure 3. Coupling uncertainty between heterogeneous models
The design of a complex (mechanical) system such as aircraft, automobile or nuclear plant involves numerical simulation of several interacting physical phenomena: CFD and structural dynamics, thermal evolution of a fluid circulation, ... For instance, they can represent the resolution of coupled partial differential equations using finite element method. In the framework of uncertainty treatment, the studied “phenomenological model” is a chaining of different models representing the various involved physical phenomena. As an example, the pressure field on an aircraft wing is the result of both aerodynamic and structural mechanical phenomena. Let us consider the particular case of two models of partial differential equations coupled by limit conditions. The direct propagation of uncertainties is impossible since it requires an exploration and then, many calls to costly models. As a solution, engineers use to build reduced-order models: the complex high-fidelity model is substituted with a CPU less costly model. The uncertainty propagation is then realized through the simplified model, taking into account the approximation error (see [ 46 ]).

Interactions between the various models are usually explicited at the finest level (cf. Fig. 3 ). How may this coupling be formulated when the fine structures of exchange have disappeared during model reduction? How can be expressed the interactions between models at different levels (in a multi-level modeling)? The ultimate question would be: how to choose the right level of modeling with respect to performance requirements?

In the multi-physical numerical simulation, two kinds of uncertainties then coexist: the uncertainty due to substitution of high-fidelity models with approximated reduced-order models, and the uncertainty due to the new coupling structure between reduced-order models.

According to the previous discussion, the uncertainty treatment in a multi-physical and multi-level modeling implies a large range of issues, for instance numerical resolutions of PDE (which do not enter into the research topics of Regularity ). Our goal is to contribute to the theoretical arsenal that allows to fly among the different levels of modeling (and then, among the existing numerical simulations). We will focus on the following three axes:

- In the case of a phenomenon represented by two coupled partial differential equations whose resolution is represented by reduced-order models, how to define a probabilistic model of the coupling errors? In connection with our theoretical development, we plan to characterize the regularity of this error in order to quantify its distribution. This research axis is supported by an ANR grant (OPUS project).
- The multi-level modeling assumes the ability to choose the right level of details for the models in adequacy to the goals of the study. In order to do that, a rigorous mathematical definition of the notion of model fineness/granularity would be very helpful. Again, a precise analysis of the fine regularity of stochastic models is expected to give elements toward a precise definition of granularity. This research axis is supported by a a PÅ’le SYSTEM@TIC grant (EHPOC project), and also by a collaboration with EADS.
- Some fine characteristics of the phenomenological model may be used to define the probabilistic behaviour of its variability. The action of modeling a phenomena can be seen as an interpolation issue between given observations. This interpolation can be driven by physical evolution equations or fine analytical description of the physical quantities. We are convinced that Hölder regularity is an essential parameter in that context, since it captures how variations at a given point induce variations at its neighbors. Stochastic processes with prescribed regularity (see section 3.3 ) have already been used to represent various fluctuating phenomena: Internet traffic, financial data, ocean floor. We believe that these models should be relevant to describe solutions of PDE perturbed by uncertain (random) coefficients or limit conditions. This research axis is supported by a PÅ’le SYSTEM@TIC grant (CSDL project).

4.3. Biomedical Applications

ECG analysis and modeling
ECG and signals derived from them are an important source of information in the detection of various pathologies, including e.g. congestive heart failure, arrhythmia and sleep apnea. The fact that the irregularity of ECG bears some information on the condition of the heart is well documented (see e.g. the web resource http://www.physionet.org). The regularity parameters that have been studied so far are mainly the box and regularization dimensions, the local Hölder exponent and the multifractal spectrum [53], [55]. These have been found to correlate well with certain pathologies in some situations. From a general point of view, we participate in this research area in two ways.

- First, we use refined regularity characterizations, such as the regularization dimension, 2-microlocal analysis and advanced multifractal spectra for a more precise analysis of ECG data. This requires in particular to test current estimation procedures and to develop new ones.
- Second, we build stochastic processes that mimic in a faithful way some features of the dynamics of ECG. For instance, the local regularity of RR intervals, estimated in a parametric way based on a modeling by an mBm, displays correlations with the amplitude of the signal, a feature that seems to have remained unobserved so far [3]. In other words, RR intervals behave as SRP. We believe that modeling in a simplified way some aspects of the interplay between the sympathetic and parasympathetic systems might lead to an SRP, and to explain both this self-regulating property and the reasons behind the observed multifractality of records. This will open the way to understanding how these properties evolve under abnormal behaviour.

**Pharmacodynamics and patient drug compliance**

Poor adherence to treatment is a worldwide problem that threatens efficacy of therapy, particularly in the case of chronic diseases. Compliance to pharmacotherapy can range from 5% to 90%. This fact renders clinical tested therapies less effective in ambulatory settings. Increasing the effectiveness of adherence interventions has been placed by the World Health Organization at the top list of the most urgent needs for the health system. A large number of studies have appeared on this new topic in recent years [67], [66]. In collaboration with the pharmacy faculty of Montréal university, we consider the problem of compliance within the context of multiple dosing. Analysis of multiple dosing drug concentrations, with common deterministic models, is usually based on patient full compliance assumption, i.e., drugs are administered at a fixed dosage. However, the drug concentration-time curve is often influenced by the random drug input generated by patient poor adherence behaviour, inducing erratic therapeutic outcomes. Following work already started in Montréal [60], [61], we consider stochastic processes induced by taking into account the random drug intake induced by various compliance patterns. Such studies have been made possible by technological progress, such as the “medication event monitoring system”, which allows to obtain data describing the behaviour of patients.

We use different approaches to study this problem: statistical methods where enough data are available, model-based ones in presence of qualitative description of the patient behaviour. In this latter case, piecewise deterministic Markov processes (PDP) seem a promising path. PDP are non-diffusion processes whose evolution follows a deterministic trajectory governed by a flow between random time instants, where it undergoes a jump according to some probability measure [49]. There is a well-developed theory for PDP, which studies stochastic properties such as extended generator, Dynkin formula, long time behaviour. It is easy to cast a simplified model of non-compliance in terms of PDP. This has allowed us already to obtain certain properties of interest of the random concentration of drug [40]. In the simplest case of a Poisson distribution, we have obtained rather precise results that also point to a surprising connection with infinite Bernouilli convolutions [29], [13], [12]. Statistical aspects remain to be investigated in the general case.
4. Application Domains

4.1. Application Domains

Here are a few examples of applications of research done in SECSI:

- Security of electronic voting schemes: the case of the Helios protocol, used in particular at University of Louvain-la-Neuve (2010) and at the International Association for Cryptologic Research (IACR).
- Security of the protocols involved in the TPM (Trusted Platform Module) chip, a chip present in most PC laptops today, and which is meant to act as a trusted base.
- Security of the European electronic passport—and the discovery of an attack on the French implementation of it.
- The Tookan tool allows one to assess the security of security tokens. These tokens are meant as safes holding secret keys, which should never be permitted to get out unencrypted. Several vulnerabilities discovered. Several interesting customers in banking (HSBC, Barclays), in aeronautics (Boeing), notably.
- Intrusion detection with the Orchids tool: several interested partners, among which EADS Cassidian, Thales, Galois Inc. (USA), the French Direction Générale de l’Armement (DGA).
4. Application Domains

4.1. Introduction

A key goal of SELECT is to produce methodological contributions in statistics. For this reason, the SELECT team works with applications that serve as an important source of interesting practical problems and require innovative methodologies to address them. Most of our applications involve contracts with industrial partners, e.g. in reliability, although we also have several more academic collaborations, e.g. genomics, genetics and neuroimaging.

4.2. Curves classification

The field of classification for complex data as curves, functions, spectra and time series is important. Standard data analysis questions are being revisited to define new strategies that take the functional nature of the data into account. Functional data analysis addresses a variety of applied problems, including longitudinal studies, analysis of fMRI data and spectral calibration.

We are focusing on unsupervised classification. In addition to standard questions as the choice of the number of clusters, the norm for measuring the distance between two observations, and the vectors for representing clusters, we must also address a major computational problem. The functional nature of the data needs to be design efficient anytime algorithms.

4.3. Computer Experiments and Reliability

Since several years, SELECT has collaborations with EDF-DER Maintenance des Risques Industriels group. An important theme concerns the resolution of inverse problems using simulation tools to analyze uncertainty in highly complex physical systems. A collaboration on an analogous topic is developed with Dassault Aviation.

The other major theme concerns probabilistic modeling in fatigue analysis in the context of a research collaboration with SAFRAN an high-technology group (Aerospace propulsion, Aircraft equipment, Defense Security, Communications).

4.4. Neuroimaging

Since 2007 SELECT participates to a working group with team Neurospin (CEA-INSERM-INRIA) on Classification, Statistics and fMRI (functional Magnetic Resonance Imaging) analysis. In this framework two theses have been co-supervised by SELECT and Neurospin researchers (Merlin Keller 2006-2009 and Vincent Michel 2007-2010). The aim of this research is to determine which parts of the brain are activated by different types of stimuli. A model selection approach is useful to avoid “false-positive” detections.

4.5. Analysis of genomic data

For the past few years SELECT has collaborated with Marie-Laure Martin-Magniette (URGV) for the analysis of genomic data. An important theme of this collaboration is using statistically sound model-based clustering methods to discover groups of co-expressed genes from microarray and high-throughput sequencing data. In particular, identifying biological entities that share similar profiles across several treatment conditions, such as co-expressed genes, may help identify groups of genes that are involved in the same biological processes.
4.6. Environment

A study has been achieved by Jean-Michel Poggi, François-Xavier Jollois (Université Paris-Descartes) and Bruno Portier (INSA de Rouen), in the context of a collaboration between AirNormand, Paris Descartes University and INSA of Rouen. They analyzed and forecasted PM10 pollution in Rouen area on six different monitoring sites to quantify the effects of variables of different types, mainly meteorological versus other pollutant measurements. Some recent non parametric statistical methods (random forests, mixture of linear models and nonlinear additive models) have been used and beyond the application, this study shed light on those methods.
4. Application Domains

4.1. Communications

Clearly, our main field of applications is telecommunications. We participate in the protection of information. We are proficient on a theoretical level, and ready to develop applications using modern cryptographic techniques, with a main focus on elliptic curve cryptography and codes based on algebraic curves. One potential application is cryptosystems in environments with limited resources as smart cards, mobile phones, and \textit{ad hoc} networks. For coding, we envisage developing algebraic codes for the erasure channel or distributed storage.
4. Application Domains

4.1. Application Domains

Since its creation, TAO mainstream applications regard Numerical Engineering, Autonomous Robotics, and Control and Games. Two new fields of applications, due to the arrival of Cécile Germain (Pr UPS, 2005), Philippe Caillou (MdC, 2005), Balázs Kégl (CR CNRS LAL, 2006) and Cyril Furtlehner (CR INRIA, 2007) have been considered: Autonomic Computing and Complex Systems.

**Numerical Engineering** still is a major source of applications. The successful OMD (Optimization Multi-Disciplinaire) RNTL/ANR project is being resumed by OMD2, started in July 2009. Collaborations with IFP and PSA automobile industry respectively led to Zyed Bouzarkouna’s and Mouadh Yagoubi’s PhD CIFRE. TAO leads the Work Package “Optimization” in the System@atic CSDL project, responsible for both fundamental research on surrogate models in multi-objective optimization and the setup of a software platform, that lead to Ilya Loshchilov’s PhD work. A collaboration with CEA DM2S was conducted as a Digiteo project and lead to Philippe Rolet’s PhD on simplified models.

**Autonomous Software Robotics** is rooted in our participation to the SYMBRION European IP and SyDiN-MaLaS (ANR-JST, coll. University of Kyushu). On this topic, Jean-Marc Montanier started his PhD in Sept. 2009; Vladimir Skortsov did his Post-doc from Sept. 2009 to Sept. 2010; Weijia Wang and Riad Akrour started their PhDs in Sept. 2010. See Section 6.2.

Our activity in **Control and Games** is chiefly visible through Mogo, already mentioned in the Highlights. Another application regards Brain Computer Interfaces: the Digiteo project *Digibrain* (coll. with CEA List and Neurospin), with Cedric Gouy-Pailler’s postdoc from October 2009 to October 2010.

Applications related to **Autonomic Computing** became an important part of TAO activities, led by Cécile Germain and Balázs Kégl in tight collaboration with the Laboratoire de l’Accélérateur Linéaire (section 6.1). Applications related to **Social Systems** are led by Philippe Caillou and Cyril Furtlehner, respectively investigating multi-agent models for labor market, and road traffic models (ANR project TRAVesti, coordinated by C. Furtlehner, started in 2009). Last but not least, the arrival in TAO of Jamal Atif brought a new application field in image analysis and understanding.
4. Application Domains

4.1. Certified decision procedures

Roughly in the last ten years, proof systems have enjoyed a wider and wider audience, having been used by a growing number of researchers and teams for a growing number of applications. We can list a dozen of INRIA teams who have used Coq in an important way. We can also list the various application fields. It comes as no surprise that these fields are often parts of the genuine activities of other related INRIA teams and appear in more detail in their own reports; among others:

- Computer security: from the formalization of security properties of protocols, to the analysis of cryptographic primitives, through questions of privacy.
- Embedded software, with a growing emphasis on real-time, reactive, software.
- Computer arithmetic: certifying the correctness of the implementation of numerical functions, possibly with explicit rounding errors.
- Formally certified automatic demonstration techniques (like SAT/SMT solvers) either for more trustworthy automatic tools, or to use the latter as formal proof techniques.

4.2. Formalized mathematics

The use of computing power has dramatically increased for the past decades, in all fields of human activity, including most branches of sciences, causing a general need for reliable computing. It also often lies the base for new interdisciplinary interactions. This is also true for so called pure mathematics. One can remark that Thomas Hales’ proof of Kepler’s conjecture, which is an undoubted result of pure mathematics, relies on computations in order to establish thousands of semi-numerical, semi-symbolic inequalities. This is done using techniques of optimization which are typically coming from applied mathematics and have been developed for very concrete applications, often engineering problems. On the other hand, the complete classification of the finite simple groups, also known as the “enormous theorem”, does not rely on any machine computation, but is a huge compound piece of published mathematics. Such level of intricacy also raised a controversy on the level of confidence one should have in the correctness of the whole. We thus see that the computer here contributes to blur the lines between what was traditionally considered “fundamental” or “applied”. In such situations, by providing a common mathematical language, formal proof systems may be the only way to provide a safe join between these various tools, through the formalization of proofs whose correctness is difficult to assess through purely human means.