RESEARCH CENTER
Nancy - Grand Est

FIELD

Activity Report 2014

Section Application Domains

Edition: 2015-03-24

ALGORILLE Project-Team

4. Application Domains

4.1. Promoting parallelism in applications

In addition to direct contributions within our own scientific domain, numerous collaborations have permitted us to test our algorithmic ideas in connection with academics of different application domains and through our association with SUPELEC with some industrial partners: physics, geology, biology, medicine, machine learning or finance.

4.2. Experimental methodologies for the evaluation of distributed systems

Our experimental research axis has a *meta* positioning, targeting all large-scale distributed systems. This versatility allows us to factorize the efforts and maximize our efficiency. The resulting findings are typically used by researchers and developers of systems in the following domains:

- High Performance Computing systems (in particular MPI applications on high-end platforms)
- Cloud environments (in particular virtualized environments)
- Grids (in particular high throughput computing systems)
- Peer-to-peer systems

ALICE Project-Team

4. Application Domains

4.1. Numerical simulation

flow simulation for oil exploration: we co-advised three Ph.D. theses with the Gocad Consortium, that develops modeling algorithms for oil and gas exploration. We developed specialized meshing algorithms, well suited to represent geological layers at various resolutions [27], [19].

optimal transport: this is an active research topics in the mathematics community. Given two measures μ and ν, optimal transport defines a distance between μ and ν, as the minimum cost of "morphing" μ into ν. This distance (called the *Wasserstein distance*) structures the space of measures and offers new ways of solving some highly non-linear PDEs (Monge-Ampere, Fokker-Plank ...). This requires a numerical way of computing the Wasserstein distance and its gradients. We studied a semi-discrete technique [21] (conditionally accepted to ESAIM J. M2AN) that optimizes power diagrams. This is to our knowledge the first numerical implementation of optimal transport for volumetric densities (computes the Wasserstein distance between a sum of Dirac masses and a piece-wise linear density supported on a tetrahedral mesh).

Bose-Einstein condensates: Xavier Antoine (prof. in mathematics at the Université de Lorraine) joined the team on a "delegation" position (Sept. 2013 - Sept. 2014) to explore some common research topics. We are members of the BECASIM project, funded by the ANR ("French NSF"). In a certain sense, a Bose-Einstein condensate is a "Schroedinger cat" made of a few hundred atoms. By special physical means (low temperature and lasers), the probability waves of these atoms are intermixed, thus forming an alternative state of matter. The BECASIM project aims at developing numerical simulation methods for these complicated phenomena (that intermix fluid dynamics, electromagnetics and quantum physics).

4.2. Fabrication

Our work around fabrication and additive manufacturing finds applications in different fields. Our algorithms for fast geometric computations on solids (boolean operations, morphological operations) are useful to model a variety of shapes, from mechanical engineering parts to prosthetics for medical applications. Our techniques allow for simpler modelling and processing of very intricate geometries and therefore also find applications in art and design, for unusual shapes that would be very difficult to obtain otherwise.

BIGS Project-Team (section vide)

CAMUS Team

4. Application Domains

4.1. Application domains

Performance being our main objective, our developments' target applications are characterized by intensive computation phases. Such applications are numerous in the domains of scientific computations, optimization, data mining and multimedia.

Applications involving intensive computations are necessarily high energy consumers. However this consumption can be significantly reduced thanks to optimization and parallelization. Although this issue is not our prior objective, we can expect some positive effects for the following reasons:

- Program parallelization tries to distribute the workload equally among the cores. Thus an equivalent performance, or even a better performance, to a sequential higher frequency execution on one single core, can be obtained.
- Memory and memory accesses are high energy consumers. Lowering the memory consumption, lowering the number of memory accesses and maximizing the number of accesses in the low levels of the memory hierarchy (registers, cache memories) have a positive consequence on execution speed, but also on energy consumption.

CARAMEL Project-Team

4. Application Domains

4.1. Cryptology

The first application domain for our research is cryptology. This includes cryptography (constructive side) and cryptanalysis (breaking systems). For the cryptanalysis part, although it has practical implications, we do not expect any transfer in the classical sense of the term: it is more directed to governmental agencies and the end-users who build their trust, based on the cryptanalysis effort. It is noteworthy that analysis documents from governmental agencies (see e.g., [31]) use cryptanalysis results as their key material.

4.1.1. Cryptography

Our cryptographic contributions are related to multiple facets of the large realm of curve-based cryptology. While it is quite clear that enough algorithms exist in order to provide cryptographers with elliptic curves having a suitably hard discrete logarithm (as found in cryptographic standards for instance), one must bear in mind that refinements of the requirements and extensions to curves of higher genus raise several interesting problems. Our work contributes to expanding the cryptographer's capabilities in these areas.

In the context of genus-2 curves, our work aims at two goals. First, improvements on the group law on selected curves yield better speed for the associated cryptosystems. The cryptographic primitives, and then the whole suite of cryptographic protocols built upon such curves would be accelerated. The second goal is the expansion of the set of curves that can be built given a set of desired properties. Using point counting algorithms for arbitrary curves, a curve offering a 128-bit security level, together with nice properties for fast arithmetic, has been computed by CARAMEL [4]. Another natural target for construction of curves for cryptography is also the suitability of curves for pairings. We expect to be able to compute such curves.

Important objects related to the structure of genus-2 curves are the isogenies between their Jacobians. Computing such isogenies is a key point in understanding important underlying objects such as the endomorphism ring, and can be useful in various situations, including for cryptographic or cryptanalytic applications. The team has produced important results in this context [6], [2].

4.1.2. Cryptanalysis

Our research on cryptanalysis is important for the cryptographic industry: by detecting weak instances, and setting new records we contribute to the definition of recommended families of systems together with their key sizes. The user's confidence in a cryptographic primitive is also related to how well the underlying problem is studied by researchers.

In particular, our involvement in computations with "NFS-like" algorithms encompasses of course the task of assessing the computational limits for integer factorization (as was done by the team by factoring RSA-768 [5]) and discrete-logarithm computations (as was done by the team in 2013 for the field $\mathrm{GF}(2^{809})$ [15]). The impact of the former is quite clear as it concerns the RSA algorithm; record-sized computations attract broad interest and determine updates on key-length recommendations. The latter are particularly important for pairing-based cryptography, since, in this context, one naturally encounters discrete-logarithm problems in extension fields of large degree. To this regard the breakthrough provided by the new quasi-polynomial discrete logarithm [17] is of course of utmost importance.

4.2. Computer Algebra Systems

Some of our software libraries are being used by computer algebra systems. Most of those libraries are free software, with a license that allows proprietary systems to link them. This gives us a maximal visibility, with a large number of users.

4.2.1. Magma

Magma is a very large computational algebra package. It provides a mathematically rigorous environment for computing with algebraic, number-theoretic, combinatoric, and geometric objects. It is developed in Sydney, by the team around John Cannon. It is non-commercial (in the sense that its goal is not to make profit), but is not freely distributed and is not open-source.

Several members of the team have visited Sydney — several years ago — to contribute to the development of Magma, by implementing their algorithms or helping in integrating their software. Our link to Magma exists also via the libraries it uses: it currently links GNU MPFR and GNU MPC for its floating-point calculations, and links GMP-ECM as part of its factorization suite.

4.2.2. Pari/GP

Pari/GP is a computational number theory system that is composed of a C library and an interpreter on top of it. It is developed in Bordeaux, where Karim Belabas from the LFANT project-team is the main maintainer. Its license is GPL. Although we do not directly contribute to this package, we have good contact with the developers.

4.2.3. Sage

Sage is a fairly large scale and open-source computer algebra system written in Python. Sage aggregates a large amount of existing free software, aiming at selecting the fastest free software package for each given task. The motto of Sage is that instead of "reinventing the wheel" all the time, Sage is "building the car". To date, Sage links GNU MPFR, GMP-ECM, and GNU MPC as standard packages.

4.3. Standardization

4.3.1. Floating-point arithmetic

The IEEE 754 standard for floating-point arithmetic was revised in 2008. The main new features are some new formats for decimal computations, and the recommendation of correctly rounded transcendental functions. The new decimal formats should not have an impact on our work, since we either use integer-only arithmetic, or arbitrary-precision binary floating-point arithmetic through the GNU MPFR library.

A new standard (P1788) is currently under construction for interval arithmetic. We are not officially involved in this standard, but we follow the discussions, to check in particular that the proposed standard will also cover arbitrary precision (interval) arithmetic.

CARTE Project-Team

4. Application Domains

4.1. Computer Virology

4.1.1. The theoretical track.

It is rightful to wonder why there are only a few fundamental studies on computer viruses while it is one of the important flaws in software engineering. The lack of theoretical studies explains maybe the weakness in the anticipation of computer diseases and the difficulty to improve defenses. For these reasons, we do think that it is worth exploring fundamental aspects, and in particular self-reproducing behaviors.

4.1.2. The virus detection track

The crucial question is how to detect viruses or self-replicating malwares. Cohen demonstrated that this question is undecidable. The anti-virus heuristics are based on two methods. The first one consists in searching for virus signatures. A signature is a regular expression, which identifies a family of viruses. There are obvious defects. For example, an unknown virus will not be detected, like ones related to a 0-day exploit. We strongly suggest to have a look at the independent audit [45] in order to understand the limits of this method. The second one consists in analysing the behavior of a program by monitoring it. Following [47], this kind of methods is not yet really implemented. Moreover, the large number of false-positive implies this is barely usable. To end this short survey, intrusion detection encompasses virus detection. However, unlike computer virology, which has a solid scientific foundation as we have seen, the IDS notion of "malwares" with respect to some security policy is not well defined. The interested reader may consult [67].

4.1.3. The virus protection track

The aim is to define security policies in order to prevent malware propagation. For this, we need (i) to define what is a computer in different programming languages and setting, (ii) to take into consideration resources like time and space. We think that formal methods like rewriting, type theory, logic, or formal languages, should help to define the notion of a formal immune system, which defines a certified protection.

4.1.4. The experimentation track

This study on computer virology leads us to propose and construct a "high security lab" in which experiments can be done in respect with the French law.

4.2. Computations and Dynamical Systems

4.2.1. Continuous computation theories

Understanding computation theories for continuous systems leads to studying hardness of verification and control of these systems. This has been used to discuss problems in fields as diverse as verification (see e.g., [29]), control theory (see e.g., [37]), neural networks (see e.g., [68]), and so on. We are interested in the formal decidability of properties of dynamical systems, such as reachability [58], the Skolem-Pisot problem [33], the computability of the ω-limit set [57]. Those problems are analogous to verification of safety properties.

Contrary to computability theory, complexity theory over continuous spaces is underdeveloped and not well understood. A central issue is the choice of the representation of objects by discrete data and its effects on the induced complexity notions. As for computability, it is well known that a representation is gauged by the topology it induces. However more structure is needed to capture the complexity notions: topologically equivalent representations may induce different classes of polynomial-time computable objects, e.g., developing a sound complexity theory over continuous structures would enable us to make abstract computability results more applicable by analysing the corresponding complexity issues. We think that the preliminary step towards such a theory is the development of higher-order complexity, which we are currently carrying out.

In contrast with the discrete setting, it is of utmost importance to compare the various models of computation over the reals, as well as their associated complexity theories. In particular, we focus on the General Purpose Analog Computer of Claude Shannon [69], on recursive analysis [74], on the algebraic approach [65] and on Markov computability [60]. A crucial point for future investigations is to fill the gap between continuous and discrete computational models. This is one deep motivation of our work on computation theories for continuous systems.

4.2.2. Analysis and verification of adversary systems

The other research direction on dynamical systems we are interested in is the study of properties of adversary systems or programs, i.e., of systems whose behavior is unknown or indistinct, or which do not have classical expected properties. We would like to offer proof and verification tools, to guarantee the correctness of such systems. On one hand, we are interested in continuous and hybrid systems. In a mathematical sense, a hybrid system can be seen as a dynamical system, whose transition function does not satisfy the classical regularity hypotheses, like continuity, or continuity of its derivative. The properties to be verified are often expressed as reachability properties. For example, a safety property is often equivalent to (non-)reachability of a subset of unsure states from an initial configuration, or to stability (with its numerous variants like asymptotic stability, local stability, mortality, etc ...). Thus we will essentially focus on verification of these properties in various classes of dynamical systems.

We are also interested in rewriting techniques, used to describe dynamic systems, in particular in the adversary context. As they were initially developed in the context of automated deduction, the rewriting proof techniques, although now numerous, are not yet adapted to the complex framework of modelization and programming. An important stake in the domain is then to enrich them to provide realistic validation tools, both in providing finer rewriting formalisms and their associated proof techniques, and in developing new validation concepts in the adversary case, i.e., when usual properties of the systems like, for example, termination are not verified. For several years, we have been developing specific procedures for property proofs of rewriting, for the sake of programming, in particular with an inductive technique, already applied with success to termination under strategies [48], [49], [50], to weak termination [51], sufficient completeness [53] and probabilistic termination [55]. The last three results take place in the context of adversary computations, since they allow for proving that even a divergent program, in the sense where it does not terminate, can give the expected results. A common mechanism has been extracted from the above works, providing a generic inductive proof framework for properties of reduction relations, which can be parametrized by the property to be proved [54], [56]. Provided program code can be translated into rule-based specifications, this approach can be applied to correctness proof of software in a larger context. A crucial element of safety and security of software systems is the problem of resources. We are working in the field of Implicit Computational Complexity. Interpretation based methods like Quasi-interpretations (QI) or sup-interpretations, are the approach we have been developing these last years [62], [63], [64]. Implicit complexity is an approach to the analysis of the resources that are used by a program. Its tools come essentially from proof theory. The aim is to compile a program while certifying its complexity.

CASSIS Project-Team

4. Application Domains

4.1. Verification of Security Protocols

Security protocols such as SET, TLS and Kerberos, are designed for establishing the confidence of electronic transactions. They rely on cryptographic primitives, the purpose of which is to ensure integrity of data, authentication or anonymity of participants, confidentiality of transactions, etc.

Experience has shown that the design of those protocols is often erroneous, even when assuming that cryptographic primitives are perfect, i.e., that an encoded message cannot be decrypted without the appropriate key. An intruder can intercept, analyze and modify the exchanged messages with very few computations and therefore, for example, generate important economic damage.

Analyzing cryptographic protocols is complex because the set of configurations to consider is very large, and can even be *infinite*: one has to consider any number of sessions, any size of messages, sessions interleaving, some algebraic properties of encryption or data structures.

Our objective is to automatize as much as possible the analysis of protocols starting from their specification. This consists in designing a tool that is easy to use, enables the specification of a large number of protocols thanks to a standard high-level language, and can either look for flaws in a given protocol or check whether it satisfies a given property. Such a tool is essential for verifying existing protocols, but also for helping in designing new ones. For our tool to be easy to use, it has to provide a graphical interface allowing a user to easily perform push-button verification.

Our tools for verifying security protocols are available as components of the AVISPA and AVANTSSAR platforms.

4.2. Automated Boundary Testing from Formal Specifications

We have introduced a new approach for test generation from set-oriented formal specifications: the BZ-TT method. This method is based on Constraint Logic Programming (CLP) techniques. The goal is to test every operation of the system at every boundary state using all input boundary values of that operation. It has been validated in several industry case studies for smart card OS and application validation (GSM 11-11 standard [75] and Java Card Virtual Machine Transaction mechanism [77]), information system and for embedded software [85].

This test generation method can be summed up as follows: from the formal model, the system computes boundary values to create boundary states; test cases are generated by traversal of the state space with a preamble part (sequences of operations from the initial state to a boundary state), a body part (critical invocations), an identification part (observation and oracle state computation) and a post-amble part (return path to initial or boundary state). Then, an executable test script file is generated using a test pattern and a table of correspondence between abstract operations (from the model) and concrete ones. This approach differs in several main points from previous works (e.g., [81]): first, using boundary goals as test objectives avoids the complete construction of the reachability graph; second, this process is fully automated and the test engineer could just drive it at the boundary value computation level or for the path computation.

The BZ-TT method is fully supported by the BZ-Testing-Tools tool-set. This environment is a set of tools dedicated to animation and test cases generation from B, Z or State-Chart formal specifications. It is based on the CLPS constraint solver, able to simulate the execution of the specification. By execution, we mean that the solver computes a so-called constrained state by applying the pre- and post-condition of operations. A constrained state is a constraint store where state variables and also input and output variables support constraints.

One orientation of the current work is to go beyond the finiteness assumption limitations by using symbolic constraint propagation during the test generation process. Second orientation is to extend the result to object oriented specifications as UML/OCL.Third orientation is to adapt the method to security aspect.

4.3. Program Debugging and Verification

Catching bugs in programs is difficult and time-consuming. The effort of debugging and proving correct even small units of code can surpass the effort of programming. Bugs inserted while "programming in the small" can have dramatic consequences for the consistency of a whole software system as shown, e.g., by viruses which can spread by exploiting buffer overflows, a bug which typically arises while coding a small portion of code. To detect this kind of errors, many verification techniques have been put forward such as static analysis and software model checking.

Recently, in the program verification community, there seems to be a growing demand for more declarative approaches in order to make the results of the analysis readily available to the end user. To meet this requirement, a growing number of program verification tools integrate some form of theorem proving.

The goals of our research are twofold. First, we perform theoretical investigations of various combinations of propositional and first-order satisfiability checking in order to automate the theorem proving activity required to solve a large class of program analysis problems which can be encoded as first-order formulae. Second, we experimentally investigate how our techniques behave on real problems in order to make program analysis more precise and scalable. Building tools capable of providing a good balance between precision and scalability is one of the crucial challenges to transfer theorem proving technology to the industrial domains.

4.4. Verification of Web Services

Driven by rapidly changing requirements and business needs, IT systems and applications are undergoing a paradigm shift: components are replaced by services, distributed over the network, and composed and reconfigured dynamically in a demand-driven way into service-oriented architectures. Exposing services in future network infrastructures means a wide range of trust and security issues need to be adressed. Solving them is extremely hard since making the service components trustworthy is not sufficient: composing services leads to new subtle and dangerous vulnerabilities due to interference between component services and policies, the shared communication layer, and application functionality. Thus, one needs validation of both the service components and their composition into secure service architectures. In this context, there is an obvious need of applying formal methods. Our project aims at applying our proof and constraint solving techniques to reason on web services. More precisely, we focus on the composition problem in the presence of security policies.

4.5. Model-Checking of Collaborative Systems

Collaborative systems constitute a class of distributed systems where real human interactions are predominant. In these systems, users at geographically distributed sites interact by simultaneously manipulating shared objects like text documents, XML trees, filesystems, etc. To improve data availablity, the shared objects are replicated so that the users update their local replicas and exchange their updates between them. One of the main challenges here is how to ensure the data consistency when the updates are executed in arbitrary orders at different replicas. Operational Transformation (OT) is an optimistic technique which has been proposed to overcome the consistency problem. This technique consists of an application-dependent protocol to enforce the out-of-order execution of updates even though these updates do not naturally commute. The data consistency relies crucially on the correctness of OT protocols whose proof is extremely hard. Indeed, possibly infinitely many cases should be tested. Our research work aims at applying symbolic model-checking techniques to automatically verify OT protocols. Most importantly, we are interested in finding under which conditions the model-checking problem can be reduced to a finite-state model.

COAST Team (section vide)

CORIDA Team

4. Application Domains

4.1. Biology and Medicine

4.1.1. Medicine

We began this year to study a new class of applications of observability theory. The investigated issues concern inverse problems in Magnetic Resonance Imaging (MRI) of moving bodies with emphasis on cardiac MRI. The main difficulty we tackle is due to the fact that MRI is, comparatively to other cardiac imaging modalities, a slow acquisition technique, implying that the object to be imaged has to be still. This is not the case for the heart where physiological motions, such as heart beat or breathing, are of the same order of magnitude as the acquisition time of an MRI image. Therefore, the assumption of sample stability, commonly used in MRI acquisition, is not respected. The violation of this assumption generally results in flow or motion artifacts. Motion remains a limiting factor in many MRI applications, despite different approaches suggested to reduce or compensate for its effects Welch et al. [63]. Mathematically, the problem can be stated as follows: can we reconstruct a moving image by measuring at each time step a line of its Fourier transform? From a control theoretic point of view this means that we want to identify the state of a dynamical system by using an output which is a small part of its Fourier transform (this part may change during the measurement).

There are several strategies to overcome these difficulties but most of them are based on respiratory motion suppression with breath-hold. Usually MRI uses ECG information to acquire an image over multiple cardiac cycles by collecting segments of Fourier space data at the same delay in the cycle Lanzer et al. [53], assuming that cardiac position over several ECG cycles is reproducible. Unfortunately, in clinical situations many subjects are unable to hold their breath or maintain stable apnea. Therefore breath-holding acquisition techniques are limited in some clinical situations. Another approach, so called real-time, uses fast, but low resolution sequences to be faster than heart motion. But these sequences are limited in resolution and improper for diagnostic situations, which require small structure depiction as for coronary arteries.

4.2. Simulation of viscous fluid-structure interactions

Participants: Bruno Pinçon, Jean-François Scheid [correspondant], Takéo Takahashi.

A number of numerical codes for the simulation for fluids and fluid-structure problems has been developed by the team. These codes are mainly written in MATLAB Software with the use of C++ functions in order to improve the sparse array process of MATLAB. We have focused our attention on 3D simulations which require large CPU time resources as well as large memory storage. An efficient 3D Stokes sparse solver for MATLAB is now available. An important work has been performed for the study and the development of a class of preconditioners for iterative solver of 3D Stokes problem. Efficient preconditioner of block preconditioned conjugate gradient type (BPCG) is now implemented. The use of this preconditioner significantly reduces the CPU time for the solution of linear system coming from the Stokes equations. This work has been developed in collaboration with Marc Fuentes, research engineer at Inria Nancy Grand Est. M. Fuentes has also written a PYTHON version of the 3D Stokes solver. A 3D characteristics method for the nonlinear Navier-Stokes equations is now in progress

4.3. Biohydrodynamics MATLAB Toolbox (BHT)

Participants: Alexandre Munnier [correspondant], Bruno Pinçon.

Understanding the locomotion of aquatic animals fascinated the scientific community for a long time. This constant interest has grown from the observation that aquatic mammals and fishes evolved swimming capabilities superior to what has been achieved by naval technology. A better understanding of the biomechanics of swimming may allow one to improve the efficiency, manoeuvrability and stealth of underwater vehicles. During the last fifty years, several mathematical models have been developed. These models make possible the qualitative analysis of swimming propulsion as a continuation of the previously developed quantitative theories. Based on recent mathematical advances, Biohydrodynamics MATLAB Toolbox (BHT) is a collection of M-Files for design, simulation and analysis of articulated bodies' motions in fluid. More widely, BHT allows also to perform easily any kind of numeric experiments addressing the motion of solids in ideal fluids (simulations of so-called fluid-structure interaction systems).

This software is available at http://bht.gforge.inria.fr/ .

MADYNES Project-Team

4. Application Domains

4.1. Mobile, ad-hoc and constrained networks

The results coming out from MADYNES can be applied to any dynamic infrastructure that contributes to the delivery of value added services. While this is a potentially huge application domain, we focus on the following environments at the network level:

1. multicast services,
2. ad-hoc networks,
3. mobile devices and IPv6 networks,
4. voice over IP infrastructure.

All these selected application areas exhibit different dynamicity features. In the context of multicast services, we focus on distribution, monitoring and accounting of key distribution protocols. On *ad-hoc* and dynamic networks we are investigating the provisioning, monitoring, configuration and performance management issues.

Concerning mobile devices, we are interested in their configuration, provisioning and monitoring. IPv6 work goes on in Information Models and on self-configuration of the agents.

4.2. Dynamic services infrastructures

At the service level, dynamics is also increasing very fast. We apply the results of our work on autonomous management on infrastructures which support dynamic composition and for which self-instrumentation and management automation is required.

The target service environments are:

- sensor networks,
- peer-to-peer infrastructures,
- information centric networks,
- ambiant environments.

MAGRIT Project-Team

4. Application Domains

4.1. Augmented reality

We have a significant experience in AR that allowed good progress in building usable, reliable and robust AR systems. Our contributions cover the entire process of AR: matching, pose initialization, 3D tracking, in-situ modeling, handling interaction between real and virtual objects....

4.2. Medical Imaging

For 15 years, we have been working in close collaboration with University Hospital of Nancy and GE Healthcare in interventional neuroradiology. Our common aim is to develop a multimodality framework to help therapeutic decisions and interventional gestures. Contributions of the team focus about the developments of AR tools for neuro-navigation as well as the development of simulation tools of the interventional act for training or planning. Laparoscopic surgery is another field of interest with the development of methods for tracking deformable organs based on bio-mechanical models. Some of these projects are developed in collaboration with the EPC SHACRA.

4.3. Applied mechanics

In experimental solid mechanics, an important problem is to characterize properties of specimen subject to mechanical constraints, which makes it necessary to measure tiny strains. Contactless measurement techniques have emerged in the last few years and are spreading quickly. They are mainly based on images of the surface of the specimen on which a regular grid or a random speckle has been deposited. We are engaged since June 2012 in a transdisciplinary collaboration with Institut Pascal (Clermont-Ferrand Université). The aim is to characterize the metrological performances of these techniques limited by, e.g., the sensor noise, and to improve them by several dedicated image processing tools.

MAIA Project-Team

4. Application Domains

4.1. Decision Making

Our group is involved in several applications of its more fundamental work on autonomous decision making and complex systems. Applications addressed include:

- Robotics, where the decision maker or agent is supported by a physical entity moving in the real world;
- Medicine or Personally Assisted Living, where the agent can be an analytic device recommending tests and/or treatments, or able to gather different sources of information (sensors for example) in order to help a final user, detecting for example anormal situation needing the rescue of a person (fall detection of elderly people, risk of hospitalization of a person suffering from chronic disease;
- Active Sensing, where decisions have to be taken in order to gather information on a system. This can be applied to many fields, like for example monitoring the integrity of airplanes wings or the behavior of people in public areas.

4.2. Ambient intelligence

As the Nancy – Grand Est Research Center scientific strategy pushes the development of plateforms on Robotics and Smart Living Apartments, some members of the team have recentered their research toward “ambient intelligence and AI” . This choice is backed up by the Inria Large-scale initiative project termed PAL (Personal assistant Living) in which we are strongly involved. The regional council of Lorraine also supports this new research line through the CPER, (project "situated computing" or "INFOSITU" http://infositu.loria.fr) whose coordinator is a member of MAIA Team. Within this new domain of research in MAIA, we explore how intelligent decentralized complex systems can help designing intelligent environments dedicated to elderly people with loss of autonomy. This domain of research is currently very active, taking up a societal challenge that developed countries have to address.

MASAIE Project-Team

4. Application Domains

4.1. Metapopulation models

Heterogeneity plays an important role in many infectious disease processes. For instance, spatial heterogeneity is a strong determinant of host-parasite relationships. In modeling spatial or geographic effects on the spread of a disease, a distinction is usually made between diffusion and dispersal models. In diffusion models, spread is to immediately adjacent zones, hence the phenomenon of traveling waves can appear. These models traditionally use partial differential equations. However, there are some important situations that cannot be modeled by PDE. This is the case when the space considered is discrete. For example, when we have to consider sparsely populated regions, the human population is located in patches. The organization of human-hosts into well-defined social units such as families, villages or cities, are good examples of patches. Another example arises in the study of the human African Trypanosomiasis. The vector is the tse-tse fly, and it is known that flies take fewer blood meals in villages than in coffee plantations where the villagers work during the day. For such situations where human or vectors can travel a long distance in a short period of time, dispersal models are more appropriate. These models consider migration of individuals between patches. The infection does not take place during the migration process. The situation is that of a directed graph, where the vertices represent the patches and the arcs represent the links between patches. During the last decade, there has been increased interest in these deterministic metapopulation disease models. We have generalized to n patches the Ross-Macdonald model which describes the dynamics of malaria. We incorporate in our model the fact that some patches can be vector free. We assume that the hosts can migrate between patches, but not the vectors. The susceptible and infectious individuals have the same dispersal rate. We compute the basic reproduction ratio $\mathcal{R}_0$. We prove that if $\mathcal{R}_0 \leq 1$, then the disease-free equilibrium is globally asymptotically stable. When $\mathcal{R}_0 > 1$, we prove that there exists a unique endemic equilibrium, which is globally asymptotically stable on the biological domain minus the disease-free equilibrium.

MASAIE is developing, in the framework of the CAPES-COFECUB project (see international program), a metapopulation model for dengue. This model is for the state of Rio and is using the data of foundation FIOCRUZ.

MULTISPEECH Team

4. Application Domains

4.1. Introduction

Approaches and models developed in the MULTISPEECH project are intended to be used for facilitating oral-based communication in various situations through enhancements of the communication channels, either directly via automatic speech recognition or speech production technologies, or indirectly, thanks to computer assisted language learning. Applications also include the usage of speech technologies for helping people in handicapped situations or for improving their autonomy. Foreseen application domains are related to computer assisted learning, health and autonomy (more precisely aided communication and monitoring), annotation and processing of spoken documents, and multimodal computer interaction.

4.2. Computer assisted learning

Although speaking seems quite natural, learning foreign languages, or learning the mother tongue for people with language deficiencies, represent critical cognitive stages. Hence, many scientific activities have been devoted to these issues either from a production or a perception point of view.

The general guiding principle with respect to computer assisted mother or foreign language learning is to combine modalities or to augment speech to make learning easier. Also, the system should provide indications on what should be corrected, a guidance which is considered as necessary by specialists in the oral aspects of language learning. Consequently, based upon a comparison of the learner's production to a reference, automatic diagnoses of the learner's production can be considered, as well as perceptual feedback relying on an automatic transformation of the learner's voice. For example, with respect to prosody, the diagnosis provided through both a text and a visual display, comes from an evaluation of the melodic curve and of the phoneme durations of the learner's realization; and the perceptual feedback consists in a replacement of the learner's prosodic cues by those of the reference; i.e., the signal of the learner's utterance is modified in order to reflect the prosodic cues (duration and F0) of the reference in order to make the learner aware of the expected prosodic cues. The diagnosis step strongly relies on the studies on categorization of sounds and prosody in the mother tongue and in the second language, and also depends on the influence between them. Furthermore, reliable diagnosis on individual utterances is still a challenge, and elaboration of advanced automatic feedback requires a temporally accurate segmentation of speech utterances into phones and this explains why accurate segmentation of native and non-native speech is also an important topic in the field of acoustic speech modeling.

4.3. Aided communication and monitoring

Speech technologies provide ways of helping people in handicapped situations or improving their autonomy. The following applications are considered in the project.

The first one is related to the tuning of speech recognition technology for providing a means of communication between a speaking person and a hard-of-hearing or a deaf person, through an adequate display of the recognized words and/or syllables, which takes also into account the reliability of the recognized items.

The second application aims at improving pathological voices. In this context, the goal is typically to transform the pathological voice signal in order to make it more intelligible. Ongoing work deals with esophageal voices, i.e., substituted voice learned by a laryngectomized patient who has lost his/her vocal cords after surgery. Voice conversion techniques will be studied further to enhance such voice signals, in order to produce clean and intelligible speech signals in replacement of the pathological voice.

The third application aims at improving the autonomy of elderly or disabled people, and fit with smartrooms. In a first step, source separation techniques could be tuned and should help for locating and monitoring people through the detection of sound events inside apartments. In a longer perspective, adapting speech recognition technologies to the voice of elder people should also be useful for such applications, but this requires the recording of adequate databases. Sound monitoring in other application fields (security, environmental monitoring) could also be envisaged.

4.4. Annotation and processing of spoken documents

The first type of annotation consists in transcribing a spoken document in order to get the corresponding sequences of words, with possibly some complementary information, such as the structure (punctuation) or the modality (affirmation/question) of the utterances to make the reading and understanding easier. Typical applications of the automatic transcription of radio or TV shows, or of any other spoken document, include making possible their access by deaf people, as well as by text-based indexing tools.

The second type of annotation is related to speech-text alignment, which aims at determining the starting and ending times of the words, and possibly of the sounds (phonemes). This is of interest in several cases as for example, for annotating speech corpora for linguistic studies, and for synchronizing lip movements with speech sounds, for example for avatar-based communications. Although good results are currently achieved on clean data, automatic speech-text alignment needs to be improved for properly processing noisy spontaneous speech data and needs to be extended to handle overlapping speech.

Finally, there is also a need for speech signal processing techniques in the field of multimedia content creation and rendering. Relevant techniques include speech and music separation, speech equalization, prosody modification, and speaker conversion.

4.5. Multimodal computer interactions

Speech synthesis has tremendous application in facilitating communication in a human-machine interaction context to make machines more accessible. For example, it started to be widely common to use acoustic speech synthesis in smartphones to make possible the uttering of all the information. This is valuable in particular in the case of handicap, as for blind people. Audiovisual speech synthesis, when used in an application such as a talking head, i.e., virtual 3D animated face synchronized with acoustic speech, is beneficial in particular for hard-of-hearing individuals. This requires an audiovisual synthesis that is intelligible, both acoustically and visually. A talking head could be an intermediate between two persons communicating remotely when their video information is not available, and can also be used in language learning applications as vocabulary tutoring or pronunciation training tool. Expressive acoustic synthesis is of interest for the reading of story, such as audiobook, to facilitate the access to literature (for instance for blind people or illiterate people).

NEUROSYS Team

4. Application Domains

4.1. General remarks

The research directions of the team are motivated by general anaesthesia (GA) that has attracted our attention in the last years. The following paragraphs explain in some detail the motivation of our work on the four major phenomena of GA: loss of consciousness, immobility, amnesia and analgesia.
During general anaesthesia, the electroencephalogram (EEG) on the scalp changes characteristically: increasing the anaesthetic drug concentration the amplitudes of oscillations in the α-band ($\sim 8-12$Hz) and in the $\delta-$band ($2-8$Hz) increase amplitudes in frontal electrodes at low drug concentrations whereas the spectral power decreases in the $\gamma-$band ($\sim 20-60$Hz). This characteristic change in the power is the basis of today's EEG-monitors that assist the anaesthesist in the control of the anaesthesia depths of patients during surgery. However, the conventional monitors exhibit a large variability between the patients detected anaesthetic depth and their real depth. Moreover, a certain number of patients re-gain consciousness during surgery (about $1-2$ out of 1000) and a large percentage of patients suffer from diverse after-effects, such as nausea or long-lasting cognitive impairments such as partial amnesia (from days to weeks). Since surgery under general anaesthesia is part of a hospital's everyday practice, a large number of patients suffer from these events everyday. One reason for the lacking control of such disadvantagous effects is the dramatic lack of knowledge on what is going on in the brain during general anaesthesia and a weak EEG-online monitoring system during anaesthesia. Consequently, to improve the situation of patients during and after surgery and to develop improved anaesthetic procedures or even drugs, research is necessary to learn more about the neural processes in the brain and develop new monitoring machines.

4.2. Level of consciousness

The EEG originates from coherent neural activity of populations in the cortex. Hence to understand better the characteristic power changes in EEG during anaesthesia, it is necessary to study neural population dynamics subject to the concentration of anaesthetic drugs and their action on receptors on the single neuron level. We study mathematical models which will be constrained by the signal features extracted from experimental data, such as EEG (data provided by Jamie Sleigh, University of Auckland and Christoph Destrieux, University of Tours), Local Field Potentials (data provided by Flavio Frohlich, University of North Carolina - Chapel Hill) and behavior. The combination of model and analysis of experimental data provides the optimal framework to reveal new knowledge on the neural origin of behavioral features, such as the loss of consciousness or the un-controlled gain of consciousness during surgery. For instance, modelling studies show that the characteristic changes of spectral power (second-order statistics) are not sufficient to deduce all underlying neural mechanisms. Consequently, additional higher-order statistical measures may provide additional insight into underlying neural mechanisms and may provide a novel marker for the loss of consciousness.

Moreover, the constant supervision of anaesthesized patients in intensive care is a demanding task for the personnel in hospital practice. It is almost not possible to take care of a patient constantly and hence the todays' medicine demands monitoring devices that control automatically the level of anaesthetic drugs based on the patients' neural activity (e.g., EEG). Brain-Computer-Interfaces (BCI) have already demonstrated their potential for the detection of consciousness in non-responsive patients. We will apply the data analysis techniques known in BCI to extract new markers for the depth of anaesthesia. More specifically, for deeper anaesthesia, auditory-evoked and Event-Related Desynchronization/Event-Related Synchronization (ERD/ERS) BCI could be used to better identify the state of consciousness in patients under anaesthesia. In this context, we have established a first contact to the University of Wuerzburg. Another research direction will link intracranial EEG and scalp EEG by characterising micro-awake episodes during sleep.

4.3. Immobility

A research direction will be to take benefit of the relationship between the motor activity and anesthesia. Indeed, even if no movement is visually perceptible, a study by electroencephalographic recordings of brain activity in motor areas, quantifying the characteristics of amplitude and phase synchronization observed in the alpha and beta frequency bands, may reveal an intention movement. This feature is important because it demonstrates that the patient is aware. Thus, we will develop an experimental protocol in collaboration with an anesthesiologist of the regional hospital on stimulating the median nerve at forearm level to track the evolution of the shape of the beta rebound in the motor cortex for various doses of the anesthetic agent.

4.4. Amnesia

Patients sometimes develop post-traumatic disorders associated with the surgery they underwent because they either woke up during the surgery or because the amnesiant effect of the general anaesthesia was only partial, declarative memory being maintained in some unexplained cases. It is still unknown how memory can be maintained under general anaesthesia and it needs to be investigated to improve the recovery from anaesthesia and to avoid as much as possible post-traumatic disorders. To learn more about memory under anaesthesia, we will focus our theoretical studies on the oscillation regimes observed in the hippocampus, mainly in the theta and gamma ranges, which are correlated with memory formation and retrieval.

4.5. Analgesia

One of the most important aspect in general anaesthesia is the loss of pain. During surgery, it is very difficult to find out wether the anesthesized patient feels pain and hence will develop cognitive impairement after surgery. Today, the anesthesiologist knows and detects physiological signs of pain, such as sweat, colour of skin or spontaneous unvoluntary movements. However, more objective criteria based on EEG may assist the pain detection and hence improves the patients' situation. To this end, we analyze large sets of patient EEG-data observed during surgery and aim to extract EEG signal features of pain.

ORPAILLEUR Project-Team

4. Application Domains

4.1. Biology and Chemistry

Participants: Mehwish Alam, Aleksey Buzmakov, Adrien Coulet, Marie-Dominique Devignes, Elias Egho, Nicolas Jay, Bernard Maigret, Amedeo Napoli, Nicolas Pépin-Hermann, Gabin Personeni, David Ritchie, Mohsen Sayed, Malika Smaïl-Tabbone, Yannick Toussaint.

Keywords: knowledge discovery in life sciences, bioinformatics, biology, chemistry, genomics

One major application domain which is currently investigated by the Orpailleur team is related to life sciences, with particular emphasis on biology, medicine, and chemistry. The understanding of biological systems provides complex problems for computer scientists, and the developed solutions bring new research ideas or possibilities for biologists and for computer scientists as well. Accordingly, the Orpailleur team includes biologists, chemists, and a physician, making Orpailleur a very original EPI at Inria. Indeed, the interactions between researchers in biology and researchers in computer science improve not only knowledge about systems in biology, chemistry, and medicine, but knowledge about computer science as well.

Knowledge discovery is gaining more and more interest and importance in life sciences for mining either homogeneous databases such as protein sequences and structures, or heterogeneous databases for discovering interactions between genes and environment, or between genetic and phenotypic data, especially for public health and pharmacogenomics domains. The latter case appears to be one main challenge in knowledge discovery in biology and involves knowledge discovery from complex data depending on domain knowledge.

On the same line as biological data, chemical data are presenting important challenges w.r.t. knowledge discovery, for example for mining collections of molecular structures and collections of chemical reactions in organic chemistry. The mining of such collections is an important task for various reasons among which the challenge of graph mining and the industrial needs (especially in drug design, pharmacology and toxicology). Molecules and chemical reactions are complex data that can be modeled as undirected labeled graphs. One objective for guiding computer-based synthesis in organic chemistry is to discover general synthesis methods (i.e. kinds of "meta-reactions") from currently available chemical reaction databases for designing generic and reusable synthesis plans.

Graph mining methods may play an important role in this framework as illustrated in [125], but Formal Concept Analysis (FCA) can also be used in an efficient and well-founded way [101]. Combining supervised methods –with a training sets where objects are tagged– and unsupervised methods, "jumping emerging patterns" can be detected that characterize classes of interest, e.g. toxic molecules or inhibitors. Then, a hybrid classification method based on FCA can be used for building a concept lattice where some of the concepts can be used as reference classes for classifying unknown objects, for recognition and prediction tasks. Graph mining in the framework of FCA is a very important task on which we are actively working, whose results can be transferred to text mining as well.

4.2. Medicine

Participants: Aleksey Buzmakov, Adrien Coulet, Elias Egho, Nicolas Jay, Jean Lieber, Amedeo Napoli, Matthieu Osmuk, Chedy Raïssi, Yannick Toussaint, Mickaël Zehren.

Keywords: knowledge representation, description logics, classification-based reasoning, case-based reasoning, semantic web, formal concept analysis, sequence mining, text mining

We are working on several applications in medicine, mainly in knowledge management and analysis of patient trajectories as sequences. In the first case, the Kasimir research project is about decision support and knowledge management for the treatment of cancer. This is a multidisciplinary research project in which participate researchers in computer science (Orpailleur), experts in oncology ("Institut de Cancérologie de Lorraine Alexis Vautrin" in Vandœuvre-lès-Nancy), Oncolor (a healthcare network in Lorraine involved in oncology), and A2Zi (a company working in Web technologies and involved in several projects in the medical informatics domain, http://www.a2zi.fr/). For a given cancer localization, a treatment is based on a protocol, which is applied in 70% of the cases and provides a treatment. The 30% remaining cases are "out of the protocol", e.g. contraindication, treatment impossibility, etc. and the protocol should be adapted, based on discussions among specialists. This adaptation process is modeled in Kasimir thanks to CBR, where the semantic Web technologies are used and adapted in the Kasimir project for several years.

Another work is in concern with the analysis of patient trajectories, i.e. the "path" of a patient during illness (chronic illnesses and cancer), considered as sequences. It is important to understand these sequence data and temporal data mining methods are good candidate tools for that. However, these methods should be adapted for addressing the complex nature of medical events. Thus, there is an ongoing work on the analysis of trajectories with different levels of granularity and w.r.t. external domain ontologies. In addition, it is also important to be able to compare and classify trajectories according to their content. This is why there is also a work on the definition of a similarity measure able to take into account the complex nature of trajectories and that can be efficiently implemented for allowing quick and reliable classifications.

4.3. Cooking

Participants: Valmi Dufour-Lussier, Emmanuelle Gaillard, Florence Le Ber, Jean Lieber, Amedeo Napoli, Emmanuel Nauer.

Keywords: cooking, knowledge representation, knowledge discovery, case-based reasoning, semantic wiki

The origin of the Taaable project is the Computer Cooking Contest (CCC). A contestant to CCC is a system that answers queries about recipes, using a recipe base; if no recipe exactly matches the query, then the system adapts another recipe. Taaable is a case-based reasoning system based on various technologies from semantic web, knowledge discovery, knowledge representation and reasoning. From a research viewpoint the system enables to test scientific results and to study the complementarity of various research trends in an application domain which is simple to understand and which raises complex issues at the same time. Taaable has been at the origin of the ANR CONTINT project Kolflow, whose application domain is WikiTaaable, the semantic wiki of Taaable.

4.4. Agronomy

Participants: Sébastien Da Silva, Florence Le Ber [contact person], Jean-François Mari.

Keywords: simulation, Markov model, Formal Concept Analysis, graph

In September, Sébastien da Silva has defended his PhD thesis [13]. His research was conducted in the framework of an Inria-INRA collaboration, which takes place in the INRA research network PAYOTE about landscape modeling. The thesis, supervised both by Claire Lavigne (DR in ecology, INRA Avignon) and Florence Le Ber, was concerned with the characterization and the simulation of hedgerows structures in agricultural landscapes, based on Hilbert-Peano curves and Markov models [6] [13], [66], [98].

An on-going research work about the representation of peasant knowledge is involved within a collaboration with IRD in Madagascar [40]. Sketches drawn by peasants were transformed into graphs and compared thanks to Formal Concept Analysis.

PAREO Project-Team

4. Application Domains

4.1. Application Domains

Beside the theoretical transfer that can be performed via the cooperations or the scientific publications, an important part of the research done in the *Pareo* project-team is published within software. *Tom* is our flagship implementation. It is available via the Inria Gforge (http://gforge.inria.fr) and is one of the most visited and downloaded projects. The integration of high-level constructs in a widely used programming language such as Java may have an impact in the following areas:

- Teaching: when (for good or bad reasons) functional programming is not taught nor used, *Tom* is an interesting alternative to exemplify the notions of abstract data type and pattern-matching in a Java object oriented course.
- Software quality: it is now well established that functional languages such as Caml are very successful to produce high-assurance software as well as tools used for software certification. In the same vein, *Tom* is very well suited to develop, in Java, tools such as provers, model checkers, or static analyzers.
- Symbolic transformation: the use of formal anchors makes possible the transformation of low-level data structures such as C structures or arrays, using a high-level formalism, namely pattern matching, including associative matching. *Tom* is therefore a natural choice each time a symbolic transformation has to be implemented in C or Java for instance. *Tom* has been successfully used to implement the Rodin simplifier, for the B formal method.
- Prototyping: by providing abstract data types, private types, pattern matching, rules and strategies, *Tom* allows the development of quite complex prototypes in a short time. When using Java as the host-language, the full runtime library can be used. Combined with the constructs provided by *Tom*, such as strategies, this procures a tremendous advantage.

One of the most successful transfer is certainly the use of *Tom* made by Business Objects/SAP. Indeed, after benchmarking several other rule based languages, they decided to choose *Tom* to implement a part of their software. *Tom* is used in Paris, Toulouse and Vancouver. The standard representation provided by *Tom* is used as an exchange format by the teams of these sites.

SEMAGRAMME Project-Team

4. Application Domains

4.1. Introduction

Our applicative domains concern natural language processing applications that rely on a deep semantic analysis. For instance, one may cite the following ones:

- textual entailment and inference,
- dialogue systems,
- semantic-oriented query systems,
- content analysis of unstructured documents,
- text transformation and automatic summarization,
- (semi) automatic knowledge acquisition.

However, if the need for semantics seems to be ubiquitous, there is a challenge in finding applications for which a deep semantic analysis results in a real improvement over non semantic-based techniques.

4.2. Text Transformation

Text transformation is an application domain featuring two important sub-fields of computational linguistics:

- parsing, from surface form to abstract representation,
- generation, from abstract representation to surface form.

Text simplification or automatic summarization belong to that domain.

We aim at using the framework of Abstract Categorial Grammars we develop to this end. It is indeed a reversible framework that allows both parsing and generation. Its underlying mathematical structure of λ-calculus makes it fit with our type-theoretic approach to discourse dynamics modeling. The ANR project Polymnie (see section 7.2.1.1) is especially dedicated to this aim.

SHACRA Project-Team (section vide)

TONUS Team

4. Application Domains

4.1. Controlled fusion and ITER

The search for alternative energy sources is a major issue for the future. Among others, controlled thermonuclear fusion in a hot hydrogen plasma is a promising possibility. The principle is to confine the plasma in a toroidal chamber, called a tokamak, and to attain the necessary temperatures to sustain nuclear fusion reactions. The International Thermonuclear Experimental Reactor (ITER) is a tokamak being constructed in Cadarache, France. This was the result of a joint decision by an international consortium made of the European Union, Canada, USA, Japan, Russia, South Korea, India and China. ITER is a huge project. As of today, the budget is estimated at 20 billion euros. The first plasma shot is planned for 2020 and the first deuterium-tritium operation for 2027.

Many technical and conceptual difficulties have to be overcome before the actual exploitation of fusion energy. Consequently, much research has been carried out around magnetically confined fusion. Among these studies, it is important to carry out computer simulations of the burning plasma. Thus, mathematicians and computer scientists are also needed in the design of ITER. The reliability and the precision of numerical simulations allow a better understanding of the physical phenomena and thus would lead to better designs. TONUS's main involvement is in such research.

The required temperatures to attain fusion are very high, of the order of a hundred million degrees. Thus it is imperative to prevent the plasma from touching the tokamak inner walls. This confinement is obtained thanks to intense magnetic fields. The magnetic field is created by poloidal coils, which generate the toroidal component of the field. The toroidal plasma current also induces a poloidal component of the magnetic field that twists the magnetic field lines (see Figure 2). The twisting is very important for the stability of the plasma. The idea goes back to research by Tamm and Sakharov, two Russian physicists, in the 50's.

Other devices are essential for the proper operation of the tokamak: divertor for collecting the escaping particles, microwave heating for reaching higher temperatures, fuel injector for sustaining the fusion reactions, toroidal coils for controlling instabilities, *etc*.

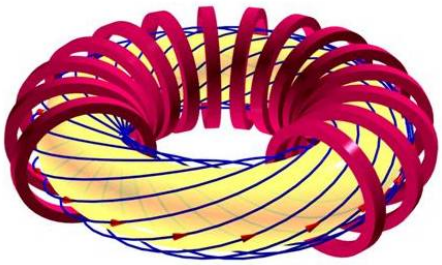

Figure 2. Poloidal coils and magnetic field lines geometry inside a tokamak

4.2. Other applications

The software and numerical methods that we develop can also be applied to other fields of physics or of engineering.

- For instance, we have a collaboration with the company AxesSim in Strasbourg for the development of efficient Discontinuous Galerkin (DG) solvers on hybrid computers. The applications is electromagnetic simulations for the conception of antenna, electronic devices or aircraft electromagnetic compatibility.
- The acoustic conception of large rooms requires huge numerical simulations. It is not always possible to solve the full wave equation and many reduced acoustic models have been developed. A popular model consists in considering "acoustic" particles moving at the speed of sound. The resulting Partial Differential Equation (PDE) is very similar to the Vlasov equation. The same modeling is used in radiation theory. We have started to work on the reduction of the acoustic particles model and realized that our reduction approach perfectly applies to this situation. We plan to supervise a new PhD with CEREMA (Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement) in Strasbourg. The objective is to investigate the model reduction and to implement the resulting acoustic model in our DG solver.

TOSCA Project-Team

4. Application Domains

4.1. Application Domains

TOSCA is interested in developing stochastic models and probabilistic numerical methods. Our present motivations come from models with singular coefficients, with applications in Geophysics, Molecular Dynamics and Neurosciences; Lagrangian modeling in Fluid Dynamics and Meteorology; Population Dynamics, Evolution and Genetics; Neurosciences; and Financial Mathematics.

4.1.1. Stochastic models with singular coefficients: Analysis and simulation

Stochastic differential equations with discontinuous coefficients arise in Geophysics, Chemistry, Molecular Dynamics, Neurosciences, Oceanography, etc. In particular, they model changes of diffusion of fluids, or diffractions of particles, along interfaces.

For practioners in these fields, Monte Carlo methods are popular as they are easy to interpret — one follows particles — and are in general easy to set up. However, dealing with discontinuities presents many numerical and theoretical challenges. Despite its important applications, ranging from brain imaging to reservoir simulation, very few teams in mathematics worldwide are currently working in this area. The Tosca project-team has tackled related problems for several years providing rigorous approach. Based on stochastic analysis as well as interacting with researchers in other fields, we developed new theoretical and numerical approaches for extreme cases such as Markov processes whose generators are of divergence form with discontinuous diffusion coefficient.

The numerical approximation of singular stochastic processes can be combined with backward stochastic differential equations (BSDEs) or branching diffusions to obtain Monte Carlo methods for quasi-linear PDEs with discontinuous coefficients. The theory of BSDEs has been extensively developed since the 1980s, but the general assumptions for their existence can be quite restrictive. Although the probabilistic interpretation of quasi-linear PDEs with branching diffusions has been known for a long time, there have been only a few works on the related numerical methods.

Another motivation to consider stochastic dynamics in a discontinuous setting came to us from time evolution of fragmentation and coagulation phenomena, with the objective to elaborate stochastic models for the avalanche formation of soils, snow, granular materials or other geomaterials. Most of the models and numerical methods for avalanches are deterministic and involve a wide variety of physical parameters such as the density of the snow, the yield, the friction coefficient, the pressure, the basal topography, etc. One of these methods consists in studying the safety factor (or limit load) problem, related to the shallow flow of a visco-plastic fluid/solid with heterogeneous thickness over complex basal topography. The resulting nonlinear partial differential equation of this last theory involves many singularities, which motivates us to develop an alternative stochastic approach based on our past works on coagulation and fragmentation. Our approach consists in studying the evolution of the size of a typical particle in a particle system which fragments in time.

4.1.2. Stochastic Lagrangian modeling in Computational Fluid Dynamics

Stochastic Lagrangian models were introduced in the eighties to simulate complex turbulent flows, particularly two-phase flows. In Computational Fluid Dynamics (CFD), they are intensively used in the so-called Probability Density Functions (PDF) methods in order to model and compute the reaction-phase terms in the fundamental equations of fluid motions. The PDF methods are currently developed in various laboratories by specialists in scientific computation and physicists. However, to our knowledge, we are innovating in two ways:

- our theoretical studies are the pioneering mathematical analysis of Lagrangian stochastic models in CFD;
- our work on the Stochastic Downscaling Method (SDM) for wind simulation is the first attempt to solve the fundamental equations themselves by a fully 3D stochastic particle method.

We emphasize that our numerical analysis is essential to the SDM development which takes benefits from our deep expertise on numerical schemes for McKean-Vlasov-non-linear SDEs.

4.1.3. Population Dynamics, Evolution and Genetics

The activity of the team on stochastic modeling in population dynamics and genetics mainly concerns application in adaptive dynamics, a branch of evolutionary biology studying the interplay between ecology and evolution, ecological modeling, population genetics in growing populations, and stochastic control of population dynamics, with applications to cancer growth modeling. Stochastic modeling in these areas mainly considers individual- based models, where the birth and death of each individual is described. This class of model is well-developed in Biology, but their mathematical analysis is still fragmentary. Another important topic topic in population dynamics is the study of populations conditioned to non-extinction, and of the corresponding stationary distributions, called quasi-stationary distributions (QSD). This domain has been the object of a lot of studies since the 1960's, but we made recently significant progresses on the questions of existence, convergence and numerical approximation of QSDs using probabilistic tools rather than the usual spectral tools.

Our activity in population dynamics also involves a fully new research project on cancer modeling at the cellular level by means of branching processes. In 2010 the International Society for Protons Dynamics in Cancer was launched in order to create a critical mass of scientists engaged in research activities on Proton Dynamics in Cancer, leading to the facilitation of international collaboration and translation of research to clinical development. Actually, a new branch of research on cancer evolution is developing intensively; it aims in particular to understand the role of proteins acting on cancerous cells' acidity, their effects on glycolysis and hypoxia, and the benefits one can expect from controlling pH regulators in view of proposing new therapies.

4.1.4. Stochastic modeling in Neuroscience

It is generally accepted that many different neural processes that take place in the brain do so in the presence of noise. Indeed, one typically observes experimentally underlying variability in the spiking times of an individual neuron in response to an unchanging stimulus, while a predictable overall picture emerges if one instead looks at the average spiking time over a whole group of neurons. Sources of noise that are of interest include ionic currents crossing the neural membrane, synaptic noise, and the global effect of the external environment (such as other parts of the brain).

It is likely that these stochastic components play an important role in the function of both the neurons and the networks they form. The characterization of the noise in the brain, its consequences at a functional level and its role at both a microscopic (individual neuron) level and macroscopic level (network of thousands of neurons) is therefore an important step towards understanding the nervous system.

To this end, a large amount of current research in the neuroscientific literature has involved the addition of noise to classical purely deterministic equations resulting in new phenomena being observed. The aim of the project is thus to rigorously study these new equations in order to be able to shed more light on the systems they describe.

4.1.5. Stochastic modeling in Financial Mathematics

4.1.5.1. Technical Analysis

In the financial industry, there are three main approaches to investment: the fundamental approach, where strategies are based on fundamental economic principles; the technical analysis approach, where strategies are based on past price behaviour; and the mathematical approach where strategies are based on mathematical models and studies. The main advantage of technical analysis is that it avoids model specification, and thus calibration problems, misspecification risks, etc. On the other hand, technical analysis techniques have limited theoretical justifications, and therefore no one can assert that they are risk-less, or even efficient.

4.1.5.2. Financial Risks Estimation and Hedging

Popular models in financial mathematics usually assume that markets are perfectly liquid. In particular, each trader can buy or sell the amount of assets he/she wants at the same price (the "market price"). They moreover assume that the decision taken by the trader does not affect the price of the asset (the small investor assumption). In practice, the assumption of perfect liquidity is never satisfied but the error due to liquidity is generally negligible with respect to other sources of error such as model error or calibration error, etc.

Derivatives of interest rates are singular for at least two reasons: firstly the underlying (interest rate) is not directly exchangeable, and secondly the liquidity costs usually used to hedge interest rate derivatives have large variation in times.

Due to recurrent crises, the problem of risk estimation is now a crucial issue in finance. Regulations have been enforced (Basel Committee II). Most asset management software products on the markets merely provide basic measures (VaR, Tracking error, volatility) and basic risk explanation features (e.g., "top contributors" to risk, sector analysis, etc).

4.1.5.3. Energy and Carbon Markets

With the rise of renewable energy generation (from wind, waves...), engineers face new challenges which heavily rely on stochastic and statistical problems.

Besides, in the context of the beginning of the second phase (the Kyoto phase) in 2008 of the European carbon market, together with the fact that French carbon tax was scheduled to come into law on Jan. 1, 2010, the year 2009 was a key year for the carbon price modeling. Our research approach adopts the point of view of the legislator and energy producers. We used both financial mathematical tools and a game theory approach. Today, with the third phase of the EU-ETS, that didn't yet start, and the report form the Cour des Comptes (October 2013) that pointed out (among many others point) the lack of mathematical modeling on such carbon market design, we continue our research in this direction.

4.1.5.4. Optimal Stopping Problems

The theory of optimal stopping is concerned with the problem of taking a decision at the best time, in order to maximise an expected reward (or minimise an expected cost). We work on the general problem of optimal stopping with random discounting and additional cost of observation.

4.1.5.5. First hitting times distributions

Diffusion hitting times are of great interest in finance (a typical example is the study of barrier options) and also in Geophysics and Neurosciences. On the one hand, analytic expressions for hitting time densities are well known and studied only in some very particular situations (essentially in Brownian contexts). On the other hand, the study of the approximation of the hitting times for stochastic differential eqution is an active area of research since very few results still are available in the literature.

VEGAS Project-Team

3. Application Domains

3.1. Computer graphics

We are interested in the application of our work to virtual prototyping, which refers to the many steps required for the creation of a realistic virtual representation from a CAD/CAM model.

When designing an automobile, detailed physical mockups of the interior are built to study the design and evaluate human factors and ergonomic issues. These hand-made prototypes are costly, time consuming, and difficult to modify. To shorten the design cycle and improve interactivity and reliability, realistic rendering and immersive virtual reality provide an effective alternative. A virtual prototype can replace a physical mockup for the analysis of such design aspects as visibility of instruments and mirrors, reachability and accessibility, and aesthetics and appeal.

Virtual prototyping encompasses most of our work on effective geometric computing. In particular, our work on 3D visibility should have fruitful applications in this domain. As already explained, meshing objects of the scene along the main discontinuities of the visibility function can have a dramatic impact on the realism of the simulations.

3.2. Solid modeling

Solid modeling, i.e., the computer representation and manipulation of 3D shapes, has historically developed somewhat in parallel to computational geometry. Both communities are concerned with geometric algorithms and deal with many of the same issues. But while the computational geometry community has been mathematically inclined and essentially concerned with linear objects, solid modeling has traditionally had closer ties to industry and has been more concerned with curved surfaces.

Clearly, there is considerable potential for interaction between the two fields. Standing somewhere in the middle, our project has a lot to offer. Among the geometric questions related to solid modeling that are of interest to us, let us mention: the description of geometric shapes, the representation of solids, the conversion between different representations, data structures for graphical rendering of models and robustness of geometric computations.

VERIDIS Project-Team

4. Application Domains

4.1. Application Domains

Distributed algorithms and protocols are found at all levels of computing infrastructure, from many-core processors and systems-on-chip to wide-area networks. We are particularly interested in novel paradigms, for example ad-hoc networks that underly mobile and low-power computing or overlay networks and peer-to-peer networking that provide services for telecommunication or cloud computing services. Computing infrastructure must be highly available and is ideally invisible to the end user, therefore correctness is crucial. One should note that standard problems of distributed computing such as consensus, group membership or leader election have to be reformulated for the dynamic context of these modern systems. We are not ourselves experts in the design of distributed algorithms, but work together with domain experts on the modeling and verification of these protocols. These cooperations help us focus on concrete algorithms and ensure that our work is relevant to the distributed algorithm community.

Formal verification techniques that we study can contribute to certify the correctness of systems. In particular, they help assert under which assumptions an algorithm or system functions as required. For example, the highest levels of the Common Criteria for Information Technology Security Evaluation require code analysis, based on mathematically precise foundations. While initially the requirements of certified development have mostly been restricted to safety-critical systems, the cost of unavailable services due to malfunctioning system components and software provides wider incentives for verification. For example, we are working on modeling and verifying medical devices that require closed-loop models of both the system and its environment.