RESEARCH CENTER
Nancy - Grand Est

FIELD

Activity Report 2019

Section Scientific Foundations

Edition: 2020-03-21

ALICE Team

3. Research Program

3.1. Point clouds

Currently, transforming the raw point cloud into a triangular mesh is a long pipeline involving disparate geometry processing algorithms:

- *Point pre-processing:* colorization, filtering to remove unwanted background, first noise reduction along acquisition viewpoint;
- *Registration:* cloud-to-cloud alignment, filtering of remaining noise, registration refinement;
- *Mesh generation:* triangular mesh from the complete point cloud, re-meshing, smoothing.

The output of this pipeline is a locally structured model which is used in downstream mesh analysis methods such as feature extraction, segmentation in meaningful parts or building CAD models.

It is well known that point cloud data contains measurement errors due to factors related to the external environment and to the measurement system itself [37], [32], [20]. These errors propagate through all processing steps: pre-processing, registration and mesh generation. Even worse, the heterogeneous nature of different processing steps makes it extremely difficult to know *how* these errors propagate through the pipeline. To give an example, for cloud-to-cloud alignment it is necessary to estimate normals. However, the normals are forgotten in the point cloud produced by the registration stage. Later on, when triangulating the cloud, the normals are re-estimated on the modified data, thus introducing uncontrollable errors.

We plan to develop new reconstruction, meshing and re-meshing algorithms, with a specific focus on the accuracy and resistance to all defects present in the input raw data. We think that pervasive treatment of uncertainty is the missing ingredient to achieve this goal. We plan to rethink the pipeline with the position uncertainty maintained during the whole process. Input points can be considered either as error ellipsoids [41] or as probability measures [27]. In a nutshell, our idea is to start by computing an error ellipsoid [43], [29] for each point of the raw data, and then to cumulate the errors (approximations) committed at each step of the processing pipeline while building the mesh. In this way, the final users will be able to take the uncertainty knowledge into account and rely on this confidence measure for further analysis and simulations. Quantifying uncertainty for reconstruction algorithms, and propagating them from input data to high-level geometry processing algorithms has never been considered before, possibly due to the very different methodologies of the approaches involved. At the very beginning we will re-implement the entire pipeline, and then attack the weak links through all three reconstruction stages.

3.2. Parameterizations

One of the favorite tools we use in our team are parameterizations. They provide a very powerful way to reveal structures on objects. The most omnipresent application of parameterizations is texture mapping: texture maps provide a way to represent in 2D (on the map) information related to a surface. Once the surface is equipped with a map, we can do much more than a mere coloring of the surface: we can approximate geodesics, edit the mesh directly in 2D or transfer information from one mesh to another.

Parameterizations constitute a family of methods that involve optimizing an objective function, subject to a set of constraints (equality, inequality, being integer, etc.). Computing the exact solution to such problems is beyond any hope, therefore approximations are the only resort. This raises a number of problems, such as the minimization of highly nonlinear functions and the definition of direction fields topology, without forgetting the robustness of the software that puts all this into practice.

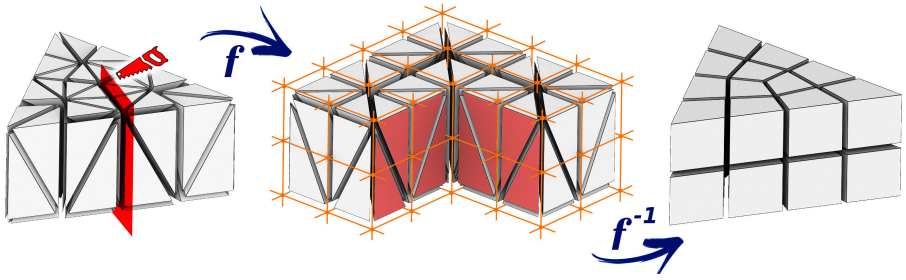


Figure 2. Hex-remeshing via global parameterization. ***Left:*** *Input tetrahedral mesh. To allow for a singular edge in the center, the mesh is cut open along the red plane.* ***Middle:*** *Mesh in parametric space.* ***Right:*** *Output mesh defined by parameterization.*

We are particularly interested in a specific instance of parameterization: hexahedral meshing. The idea [4] is to build a transformation f from the domain to a parametric space, where the deformed domain can be meshed by a regular grid. The inverse transformation f^{-1} applied to this grid produces the hexahedral mesh of the domain, aligned with the boundary of the object. The strength of this approach is that the transformation may admit some discontinuities. Let us show an example: we start from a tetrahedral mesh (Figure 2 , left) and we want deform it in a way that its boundary is aligned with the integer grid. To allow for a singular edge in the output (the valency 3 edge, Figure 2 , right), the input mesh is cut open along the highlighted faces and the central edge is mapped onto an integer grid line (Figure 2 , middle). The regular integer grid then induces the hexahedral mesh with the desired topology.

Current global parameterizations allow grids to be positioned inside geometrically simple objects whose internal structure (the singularity graph) can be relatively basic. We wish to be able to handle more configurations by improving three aspects of current methods:

- Local grid orientation is usually prescribed by minimizing the curvature of a 3D steering field. Unfortunately, this heuristic does not always provide singularity curves that can be integrated by the parameterization. We plan to explore how to embed integrability constraints in the generation of the direction fields. To address the problem, we already identified necessary validity criteria, for example, the permutation of axes along elementary cycles that go around a singularity must preserve one of the axes (the one tangent to the singularity). The first step to enforce this (necessary) condition will be to split the frame field generation into two parts: first we will define a locally stable vector field, followed by the definition of the other two axes by a 2.5D directional field (2D advected by the stable vector field).
- The grid combinatorial information is characterized by a set of integer coefficients whose values are currently determined through numerical optimization of a geometric criterion: the shape of the hexahedra must be as close as possible to the steering direction field. Thus, the number of layers of hexahedra between two surfaces is determined solely by the size of the hexahedra that one wishes to generate. In this setting degenerate configurations arise easily, and we want to avoid them. In practice, mixed integer solvers often choose to allocate a negative or zero number of layers of hexahedra between two constrained sheets (boundaries of the object, internal constraints or singularities). We will study how to inject strict positivity constraints into these cases, which is a very complex problem because of the subtle interplay between different degrees of freedom of the system. Our first results for quad-meshing of surfaces give promising leads, notably thanks to *motorcycle graphs* [21], a notion we wish to extend to volumes.
- Optimization for the geometric criterion makes it possible to control the average size of the hexahedra, but it does not ensure the bijectivity (even locally) of the resulting parameterizations.

Considering other criteria, as we did in 2D [26], would probably improve the robustness of the process. Our idea is to keep the geometry criterion to find the global topology, but try other criteria to improve the geometry.

3.3. Hexahedral-dominant meshing

All global parameterization approaches are decomposed into three steps: frame field generation, field integration to get a global parameterization, and final mesh extraction. Getting a full hexahedral mesh from a global parameterization means that it has positive Jacobian everywhere except on the frame field singularity graph. To our knowledge, there is no solution to ensure this property, but some efforts are done to limit the proportion of failure cases. An alternative is to produce hexahedral dominant meshes. Our position is in between those two points of view:

1. We want to produce full hexahedral meshes;
2. We consider as pragmatic to keep hexahedral dominant meshes as a fallback solution.

The global parameterization approach yields impressive results on some geometric objects, which is encouraging, but not yet sufficient for numerical analysis. Note that while we attack the remeshing with our parameterizations toolset, the wish to improve the tool itself (as described above) is orthogonal to the effort we put into making the results usable by the industry. To go further, our idea (as opposed to [30], [22]) is that the global parameterization should not handle all the remeshing, but merely act as a guide to fill a large proportion of the domain with a simple structure; it must cooperate with other remeshing bricks, especially if we want to take final application constraints into account.

For each application we will take as an input domains, sets of constraints and, eventually, fields (e.g. the magnetic field in a tokamak). Having established the criteria of mesh quality (per application!) we will incorporate this input into the mesh generation process, and then validate the mesh by a numerical simulation software.

BIGS Project-Team

3. Research Program

3.1. Introduction

We give here the main lines of our research that belongs to the domains of probability and statistics. For clarity, we made the choice to structure them in four items. Although this choice was not arbitrary, the outlines between these items are sometimes fuzzy because each of them deals with modeling and inference and they are all interconnected.

3.2. Stochastic modeling

Our aim is to propose relevant stochastic frameworks for the modeling and the understanding of biological systems. The stochastic processes are particularly suitable for this purpose. Among them, Markov chains give a first framework for the modeling of population of cells [80], [57]. Piecewise deterministic processes are non diffusion processes also frequently used in the biological context [47], [56], [49]. Among Markov model, we developed strong expertise about processes derived from Brownian motion and Stochastic Differential Equations [72], [55]. For instance, knowledge about Brownian or random walk excursions [79], [71] helps to analyse genetic sequences and to develop inference about it. However, nature provides us with many examples of systems such that the observed signal has a given Hölder regularity, which does not correspond to the one we might expect from a system driven by ordinary Brownian motion. This situation is commonly handled by noisy equations driven by Gaussian processes such as fractional Brownian motion of fractional fields. The basic aspects of these differential equations are now well understood, mainly thanks to the so-called rough paths tools [63], but also invoking the Russo-Vallois integration techniques [73]. The specific issue of Volterra equations driven by fractional Brownian motion, which is central for the subdiffusion within proteins problem, is addressed in [48]. Many generalizations (Gaussian or not) of this model have been recently proposed for some Gaussian locally self-similar fields, or for some non-Gaussian models [60], or for anisotropic models [44].

3.3. Estimation and control for stochastic processes

We develop inference about stochastic processes that we use for modeling. Control of stochastic processes is also a way to optimise administration (dose, frequency) of therapy.

There are many estimation techniques for diffusion processes or coefficients of fractional or multifractional Brownian motion according to a set of observations [59], [40], [46]. But, the inference problem for diffusions driven by a fractional Brownian motion is still in its infancy. Our team has a good expertise about inference of the jump rate and the kernel of Piecewise Deterministic Markov Processes (PDMP) [37], [38], [36], [39]. However, there are many directions to go further into. For instance, previous works made the assumption of a complete observation of jumps and mode, that is unrealistic in practice. We tackle the problem of inference of "Hidden PDMP". As an example, in pharmacokinetics modeling inference, we want to take into account for presence of timing noise and identification from longitudinal data. We have expertise on this subjects [41], and we also used mixed models to estimate tumor growth [42].

We consider the control of stochastic processes within the framework of Markov Decision Processes [70] and their generalization known as multi-player stochastic games, with a particular focus on infinite-horizon problems. In this context, we are interested in the complexity analysis of standard algorithms, as well as the proposition and analysis of numerical approximate schemes for large problems in the spirit of [43]. Regarding complexity, a central topic of research is the analysis of the Policy Iteration algorithm, which has made significant progress in the last years [82], [69], [54], [78], but is still not fully understood. For large problems, we have a long experience of sensitivity analysis of approximate dynamic programming algorithms for Markov Decision Processes [76], [75], [77], [62], [74], and we currently investigate whether/how similar ideas may be adapted to multi-player stochastic games.

3.4. Algorithms and estimation for graph data

A graph data structure consists of a set of nodes, together with a set of pairs of these nodes called edges. This type of data is frequently used in biology because they provide a mathematical representation of many concepts such as biological structures and networks of relationships in a population. Some attention has recently been focused in the group on modeling and inference for graph data.

Network inference is the process of making inference about the link between two variables taking into account the information about other variables. [81] gives a very good introduction and many references about network inference and mining. Many methods are available to infer and test edges in Gaussian graphical models [81], [64], [52], [53]. However, when dealing with abundance data, because inflated zero data, we are far from gaussian assumption and we want to develop inference in this case.

Among graphs, trees play a special role because they offer a good model for many biological concepts, from RNA to phylogenetic trees through plant structures. Our research deals with several aspects of tree data. In particular, we work on statistical inference for this type of data under a given stochastic model. We also work on lossy compression of trees via directed acyclic graphs. These methods enable us to compute distances between tree data faster than from the original structures and with a high accuracy.

3.5. Regression and machine learning

Regression models and machine learning aim at inferring statistical links between a variable of interest and covariates. In biological study, it is always important to develop adapted learning methods both in the context of *standard* data and also for data of high dimension (with sometimes few observations) and very massive or online data.

Many methods are available to estimate conditional quantiles and test dependencies [68], [58]. Among them we have developed nonparametric estimation by local analysis via kernel methods [50], [51] and we want to study properties of this estimator in order to derive a measure of risk like confidence band and test. We study also many other regression models like survival analysis, spatio temporal models with covariates. Among the multiple regression models, we want to develop omnibus tests that examine several assumptions together.

Concerning the analysis of high dimensional data, our view on the topic relies on the *French data analysis school*, specifically on Factorial Analysis tools. In this context, stochastic approximation is an essential tool [61], which allows one to approximate eigenvectors in a stepwise manner [67], [65], [66]. BIGS aims at performing accurate classification or clustering by taking advantage of the possibility of updating the information "online" using stochastic approximation algorithms [45]. We focus on several incremental procedures for regression and data analysis like linear and logistic regressions and PCA (Principal Component Analysis).

We also focus on the biological context of high-throughput bioassays in which several hundreds or thousands of biological signals are measured for a posterior analysis. We have to account for the inter-individual variability within the modeling procedure. We aim at developing a new solution based on an ARX (Auto Regressive model with eXternal inputs) model structure using the EM (Expectation-Maximisation) algorithm for the estimation of the model parameters.

CAMUS Project-Team

3. Research Program

3.1. Research Directions

The various objectives we are expecting to reach are directly related to the search of adequacy between the sofware and the new multicore processors evolution. They also correspond to the main research directions suggested by Hall, Padua and Pingali in [56]. Performance, correctness and productivity must be the users' perceived effects. They will be the consequences of research works dealing with the following issues:

- Issue 1: Static Parallelization and Optimization
- Issue 2: Profiling and Execution Behavior Modeling
- Issue 3: Dynamic Program Parallelization and Optimization, Virtual Machine
- Issue 4: Proof of Program Transformations for Multicores

The development of efficient and correct applications for multicore processors requires stepping in every application development phase, from the initial conception to the final run.

Upstream, all potential parallelism of the application has to be exhibited. Here static analysis and transformation approaches (issue 1) must be performed, resulting in *multi-parallel* intermediate code advising the running virtual machine about all the parallelism that can be taken advantage of. However the compiler does not have much knowledge about the execution environment. It obviously knows the instruction set, it can be aware of the number of available cores, but it does not know the actual available resources at any time during the execution (memory, number of free cores, etc.).

That is the reason why a "virtual machine" mechanism will have to adapt the application to the resources (issue 3). Moreover the compiler will be able to take advantage only of a part of the parallelism induced by the application. Indeed some program information (variables values, accessed memory adresses, etc.) being available only at runtime, another part of the available parallelism will have to be generated on-the-fly during the execution, here also, thanks to a dynamic mechanism.

This on-the-fly parallelism extraction will be performed using speculative behavior models (issue 2), such models allowing to generate speculative parallel code (issue 3). Between our behavior modeling objectives, we can add the behavior monitoring, or profiling, of a program version. Indeed, the complexity of current and future architectures avoids assuming an optimal behavior regarding a given program version. A monitoring process will make it possible to select on-the-fly the best parallelization.

These different parallelization steps are schematized in figure 1 .

Our project relies on the conception of a production chain for efficient execution of an application on a multicore architecture. Each link of this chain has to be formally verified in order to ensure correctness as well as efficiency. More precisely, it has to be ensured that the compiler produces a correct intermediate code, and that the virtual machine actually performs the parallel execution semantically equivalent to the source code: every transformation applied to the application, either statically by the compiler or dynamically by the virtual machine, must preserve the initial semantics. This must be proved formally (issue 4).

In the following, those different issues are detailed while forming our global, long term vision of what has to be done.

3.2. Static Parallelization and Optimization

Participants: Vincent Loechner, Philippe Clauss, Éric Violard, Cédric Bastoul, Arthur Charguéraud, Bérenger Bramas, Harenome Ranaivoarivony-Razanajato.

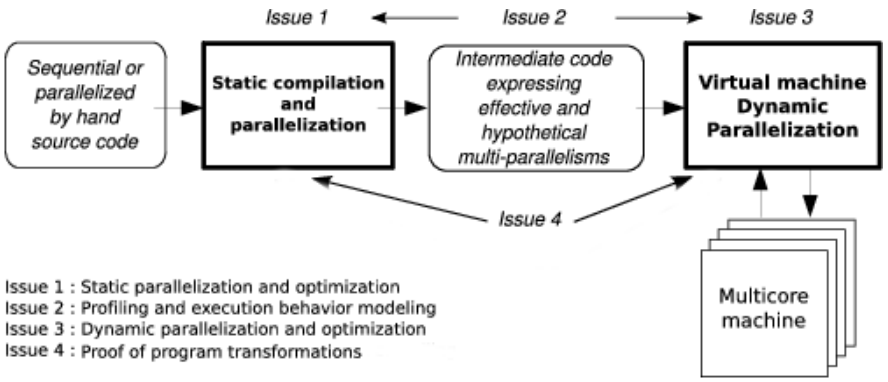


Figure 1. Steps for Automatic parallelization on multicore architectures.

Static optimizations, from source code at compile time, benefit from two decades of research in automatic parallelization: many works address the parallelization of loop nests accessing multi-dimensional arrays, and these works are now mature enough to generate efficient parallel code [53]. Low-level optimizations, in the assembly code generated by the compiler, have also been extensively dealt with for single-core and require few adaptations to support multicore architectures. Concerning multicore specific parallelization, we propose to explore two research directions to take full advantage of these architectures: adapting parallelization to multicore architectures and expressing many potential parallelisms.

3.3. Profiling and Execution Behavior Modeling

Participants: Alain Ketterlin, Philippe Clauss, Salwa Kobeissi.

The increasing complexity of programs and hardware architectures makes it ever harder to characterize beforehand a given program's run time behavior. The sophistication of current compilers and the variety of transformations they are able to apply cannot hide their intrinsic limitations. As new abstractions like transactional memories appear, the dynamic behavior of a program strongly conditions its observed performance. All these reasons explain why empirical studies of sequential and parallel program executions have been considered increasingly relevant. Such studies aim at characterizing various facets of one or several program runs, *e.g.*, memory behavior, execution phases, etc. In some cases, such studies characterize more the compiler than the program itself. These works are of tremendous importance to highlight all aspects that escape static analysis, even though their results may have a narrow scope, due to the possible incompleteness of their input data sets.

3.4. Dynamic Parallelization and Optimization, Virtual Machine

Participants: Philippe Clauss, Salwa Kobeissi, Jens Gustedt, Alain Ketterlin, Muthena Abdul-Wahab, Daniel Salas, Bérenger Bramas.

Dynamic parallelization and optimization has become essential with the advent of the new multicore architectures. When using a dynamic scheme, the performed instructions are not only dedicated to the application functionalities, but also to its control and its transformation, and so in its own interest. Behaving like a computer virus, such a scheme should rather be qualified as a "vitamin". It perfectly knows the current characteristics of the execution environment and owns some qualitative information thanks to a behavior modeling process (issue 2). It provides a significant optimization ability compared to a static compiler, while observing the evolution of the availability of live resources.

3.5. Proof of Program Transformations for Multicores

Participants: Éric Violard, Alain Ketterlin, Julien Narboux, Nicolas Magaud, Arthur Charguéraud.

Our main objective consists in certifying the critical modules of our optimization tools (the compiler and the virtual machine). First we will prove the main loop transformation algorithms which constitute the core of our system.

The optimization process can be separated into two stages: the transformations consisting in optimizing the sequential code and in exhibiting parallelism, and those consisting in optimizing the parallel code itself. The first category of optimizations can be proved within a sequential semantics. For the other optimizations, we need to work within a concurrent semantics. We expect the first stage of optimization to produce data-race free code. For the second stage of optimization we will first assume that the input code is data-race free. We will prove those transformations using Appel's concurrent separation logic [57]. Proving transformations involving programs which are not data-race free will constitute a longer term research goal.

CAPSID Project-Team

3. Research Program

3.1. Classifying and Mining Protein Structures and Protein Interactions

3.1.1. Context

The scientific discovery process is very often based on cycles of measurement, classification, and generalisation. It is easy to argue that this is especially true in the biological sciences. The proteins that exist today represent the molecular product of some three billion years of evolution. Therefore, comparing protein sequences and structures is important for understanding their functional and evolutionary relationships [67], [48]. There is now overwhelming evidence that all living organisms and many biological processes share a common ancestry in the tree of life. Historically, much of bioinformatics research has focused on developing mathematical and statistical algorithms to process, analyse, annotate, and compare protein and DNA sequences because such sequences represent the primary form of information in biological systems. However, there is growing evidence that structure-based methods can help to predict networks of protein-protein interactions (PPIs) with greater accuracy than those which do not use structural evidence [52], [70]. Therefore, developing techniques which can mine knowledge of protein structures and their interactions is an important way to enhance our knowledge of biology [39].

3.1.2. Formalising and Exploiting Domain Knowledge

Concerning protein structure classification, we aim to explore novel classification paradigms to circumvent the problems encountered with existing hierarchical classifications of protein folds and domains. In particular it will be interesting to set up fuzzy clustering methods taking advantage of our previous work on gene functional classification [43], but instead using Kpax domain-domain similarity matrices. A non-trivial issue with fuzzy clustering is how to handle similarity rather than mathematical distance matrices, and how to find the optimal number of clusters, especially when using a non-Euclidean similarity measure. We will adapt the algorithms and the calculation of quality indices to the Kpax similarity measure. More fundamentally, it will be necessary to integrate this classification step in the more general process leading from data to knowledge called Knowledge Discovery in Databases (KDD) [46].

Another example where domain knowledge can be useful is during result interpretation: several sources of knowledge have to be used to explicitly characterise each cluster and to help decide its validity. Thus, it will be useful to be able to express data models, patterns, and rules in a common formalism using a defined vocabulary for concepts and relationships. Existing approaches such as the Molecular Interaction (MI) format [49] developed by the Human Genome Organization (HUGO) mostly address the experimental wet lab aspects leading to data production and curation [58]. A different point of view is represented in the Interaction Network Ontology (INO), a community-driven ontology that aims to standardise and integrate data on interaction networks and to support computer-assisted reasoning [71]. However, this ontology does not integrate basic 3D concepts and structural relationships. Therefore, extending such formalisms and symbolic relationships will be beneficial, if not essential, when classifying the 3D shapes of proteins at the domain family level.

Domain family classification is also relevant for studying domain-domain interactions (DDI). Our previous work on Knowledge-Based Docking (KBDOCK, [3], [5] will be updated and extended using newly published DDIs. Methods for inferring new DDIs from existing protein-protein interactions (PPIs) will be developped. Efforts should be made for validating such inferred DDIs so that they can be used to enrich DDI classification and predict new PPIs.

In parallel, we also intend to design algorithms for leveraging information embedded in biological knowledge graphs (also known as complex networks). Knowledge graphs mostly represent PPIs, integrated with various properties attached to proteins, such as pathways, drug binding or relation with diseases. Setting up similarity measures for proteins in a knowledge graph is a difficult challenge. Our objective is to extract useful knowledge from such graphs in order to better understand and highlight the role of multi-component assemblies in various types of cell or organisms. Ultimately, knowledge graphs can be used to model and simulate the functioning of such molecular machinery in the context of the living cell, under physiological or pathological conditions.

3.1.3. Function Annotation in large protein graphs

Knowledge of the functional properties of proteins can shed considerable light on how they might interact. However, huge numbers of protein sequences in public databases such as UniProt/TrEMBL lack any functional annotation, and the functional annotation of such sequences is a highly challenging problem. We are developing graph-based and machine learning techniques to annotate automatically the available unannotated sequences with functional properties such as EC numbers and Gene Ontology (GO) terms (note that these terms are organized hierarchically allowing generalization/specialization reasoning). The idea is to transfer annotations from expert-reviewed sequences present in the UniProt/SwissProt database (about 560 thousands entries) to unreviewed sequences present in the UniProt/TrEMBL database (about 80% of 180 millions entries). For this, we have to learn from the UniProt/SwissProt database how to compute the similarity of proteins sharing identical or similar functional annotations. Various similarity measures can be tested using cross-validation approches in the UniProt/SwissProt database. For instance, we can use primary sequence or domain signature similarities. More complex similarities can be computed with graph-embedding techniques.

This work is in progress with Bishnu Sarker's PhD project and a first approach called GrAPFI (Graph-based Automatic Protein Function Inference) was presented at conferences in 2018 [11], [12].

3.2. Integrative Multi-Component Assembly and Modeling

3.2.1. Context

At the molecular level, each PPI is embodied by a physical 3D protein-protein interface. Therefore, if the 3D structures of a pair of interacting proteins are known, it should in principle be possible for a docking algorithm to use this knowledge to predict the structure of the complex. However, modeling protein flexibility accurately during docking is very computationally expensive. This is due to the very large number of internal degrees of freedom in each protein, associated with twisting motions around covalent bonds. Therefore, it is highly impractical to use detailed force-field or geometric representations in a brute-force docking search. Instead, most protein docking algorithms use fast heuristic methods to perform an initial rigid-body search in order to locate a relatively small number of candidate binding orientations, and these are then refined using a more expensive interaction potential or force-field model, which might also include flexible refinement using molecular dynamics (MD), for example.

3.2.2. Polar Fourier Docking Correlations

In our *Hex* protein docking program [60], the shape of a protein molecule is represented using polar Fourier series expansions of the form

$$\sigma(\underline{x}) = \sum_{nlm} a_{nlm} R_{nl}(r) y_{lm}(\theta, \phi), \tag{1}$$

where $\sigma(\underline{x})$ is a 3D shape-density function, a_{nlm} are the expansion coefficients, $R_{nl}(r)$ are orthonormal Gauss-Laguerre polynomials and $y_{lm}(\theta, \phi)$ are the real spherical harmonics. The electrostatic potential, $\phi(\underline{x})$, and charge density, $\rho(\underline{x})$, of a protein may be represented using similar expansions. Such representations allow the *in vacuo* electrostatic interaction energy between two proteins, A and B, to be calculated as [51]

$$E = \frac{1}{2} \int \phi_A(\underline{x}) \rho_B(\underline{x}) \mathrm{d}\underline{x} + \frac{1}{2} \int \phi_B(\underline{x}) \rho_A(\underline{x}) \mathrm{d}\underline{x}. \tag{2}$$

This equation demonstrates using the notion of *overlap* between 3D scalar quantities to give a physics-based scoring function. If the aim is to find the configuration that gives the most favourable interaction energy, then it is necessary to perform a six-dimensional search in the space of available rotational and translational degrees of freedom. By re-writing the polar Fourier expansions using complex spherical harmonics, we showed previously that fast Fourier transform (FFT) techniques may be used to accelerate the search in up to five of the six degrees of freedom [61]. Furthermore, we also showed that such calculations may be accelerated dramatically on modern graphics processor units [10], [7]. Consequently, we are continuing to explore new ways to exploit the polar Fourier approach.

3.2.3. Assembling Symmetrical Protein Complexes

Although protein-protein docking algorithms are improving [62], [53], it still remains challenging to produce a high resolution 3D model of a protein complex using *ab initio* techniques. This is mainly due to the problem of structural flexibility described above. However, with the aid of even just one simple constraint on the docking search space, the quality of docking predictions can improve considerably [10], [61]. In particular, many protein complexes involve symmetric arrangements of one or more sub-units, and the presence of symmetry may be exploited to reduce the search space considerably [38], [59], [66]. For example, using our operator notation (in which $\widehat{R}$ and $\widehat{T}$ represent 3D rotation and translation operators, respectively), we have developed an algorithm which can generate and score candidate docking orientations for monomers that assemble into cyclic (C_n) multimers using 3D integrals of the form

$$E_{AB}(y,\alpha,\beta,\gamma) = \int \left[\widehat{T}(0,y,0)\widehat{R}(\alpha,\beta,\gamma)\phi_A(\underline{x})\right] \times \left[\widehat{R}(0,0,\omega_n)\widehat{T}(0,y,0)\widehat{R}(\alpha,\beta,\gamma)\rho_B(\underline{x})\right] \mathrm{d}\underline{x}, \quad (3)$$

where the identical monomers A and B are initially placed at the origin, and $\omega_n = 2\pi/n$ is the rotation about the principal n-fold symmetry axis. This example shows that complexes with cyclic symmetry have just 4 rigid body degrees of freedom (DOFs), compared to $6(n-1)$ DOFs for non-symmetrical n-mers. We have generalised these ideas in order to model protein complexes that crystallise into any of the naturally occurring point group symmetries (C_n, D_n, T, O, I). This approach was published in 2016 [8], and was subsequently applied to several symmetrical complexes from the "CAPRI" blind docking experiment [45]. Although we currently use shape-based FFT correlations, the symmetry operator technique may equally be used to build and refine candidate solutions using a more accurate coarse-grained (CG) force-field scoring function.

3.2.4. Coarse-Grained Models

Many approaches have been proposed in the literature to take into account protein flexibility during docking. The most thorough methods rely on expensive atomistic simulations using MD. However, much of a MD trajectory is unlikely to be relevant to a docking encounter unless it is constrained to explore a putative protein-protein interface. Consequently, MD is normally only used to refine a small number of candidate rigid body docking poses. A much faster, but more approximate method is to use "coarse-grained" (CG) normal mode analysis (NMA) techniques to reduce the number of flexible degrees of freedom to just one or a handful of the most significant vibrational modes [57], [44], [54], [55]. In our experience, docking ensembles of NMA conformations does not give much improvement over basic FFT-based soft docking [68], and it is very computationally expensive to use side-chain repacking to refine candidate soft docking poses [4].

In the last few years, CG force-field models have become increasingly popular in the MD community because they allow very large biomolecular systems to be simulated using conventional MD programs [37]. Typically, a CG force-field representation replaces the atoms in each amino acid with from 2 to 4 "pseudo-atoms", and it assigns each pseudo-atom a small number of parameters to represent its chemo-physical properties. By directly attacking the quadratic nature of pair-wise energy functions, coarse-graining can speed up MD simulations by up to three orders of magnitude. Nonetheless, such CG models can still produce useful models of very large multi-component assemblies [65]. Furthermore, this kind of CG model effectively integrates out many of the internal DOFs to leave a smoother but still physically realistic energy surface [50]. We are currently developing a CG scoring function for fast protein-protein docking and multi-component assembly. This work is part of

the PhD project of Maria-Elisa Ruiz-Echartea [19], [64]. Beyond this PhD project, the CG scoring function will be exploited in all our docing projects, especially for RNA-Protein docking (see below).

3.2.5. Assembling Multi-Component Complexes and Integrative Structure Modeling

We also want to develop related approaches for integrative structure modeling using cryo-electron microscopy (cryo-EM). Thanks to recent developments in cryo-EM instruments and technologies, it is now feasible to capture low resolution images of very large macromolecular machines. However, while such developments offer the intriguing prospect of being able to trap biological systems in unprecedented levels of detail, there will also come with an increasing need to analyse, annotate, and interpret the enormous volumes of data that will soon flow from the latest instruments. In particular, a new challenge that is emerging is how to fit previously solved high resolution protein structures into low resolution cryo-EM density maps. However, the problem here is that large molecular machines will have multiple sub-components, some of which will be unknown, and many of which will fit each part of the map almost equally well. Thus, the general problem of building high resolution 3D models from cryo-EM data is like building a complex 3D jigsaw puzzle in which several pieces may be unknown or missing, and none of which will fit perfectly. We wish to proceed firstly by putting more emphasis on the single-body terms in the scoring function [42], and secondly by using fast CG representations and knowledge-based distance restraints to prune large regions of the search space. This work has made some progress during the PhD project of Maria Elisa Ruiz Echartea but still requires further efforts.

3.2.6. Protein-Nucleic Acids Interactions

As well as playing an essential role in the translation of DNA into proteins, RNA molecules carry out many other essential biological functions in cells, often through their interactions with proteins. A critical challenge in modelling such interactions computationally is that the RNA is often highly flexible, especially in single-stranded (ssRNA) regions of its structure. These flexible regions are often very important because it is through their flexibility that the RNA can adjust its 3D conformation in order to bind to a protein surface. However, conventional protein-protein docking algorithms generally assume that the 3D structures to be docked are rigid, and so are not suitable for modeling protein-RNA interactions. There is therefore much interest in developing protein-RNA docking algorithms which can take RNA flexibility into account. This research topic has been initiated with the recruitement of Isaure Chauvot de Beauchêne in 2016 and is becoming a major activity in the team. A novel flexible docking algorithm is currently under development in the team. It first docks small fragments of ssRNA (typically three nucleotides at a time) onto a protein surface, and then combinatorially reassembles those fragments in order to recover a contiguous ssRNA structure on the protein surface [41], [40].

As the correctness of the initial docking of the fragments settles an upper limit to the correctness of the full model, we are now focusing on improving that step. A key component of our docking tool is the energy function of the protein - fragment interactions, that is used both to drive the sampling (positioning of the fragments) by minimization and to discriminate the correct final positions from decoys (i.e. false positives). We are developing a new knowledge-based energy function that will be learnt by machine-learning methods from public structural data on ssRNA-protein complexes.

In the future, we will improve the combinatorial algorithm used for reassembling the docked fragments using experimental constraints and machine-learning approaches.

CARAMBA Project-Team

3. Research Program

3.1. The Extended Family of the Number Field Sieve

The Number Field Sieve (NFS) has been the leading algorithm for factoring integers for more than 20 years, and its variants have been used to set records for discrete logarithms in finite fields. It is reasonable to understand NFS as a framework that can be used to solve various sorts of problems. Factoring integers and computing discrete logarithms are the most prominent for the cryptographic observer, but the same framework can also be applied to the computation of class groups.

The state of the art with NFS is built from numerous improvements of its inner steps. In terms of algorithmic improvements, the recent research activity on the NFS family has been rather intense. Several new algorithms have been discovered since 2014, notably for non-prime fields, and their practical reach has been demonstrated by actual experiments.

The algorithmic contributions of the CARAMBA members to NFS would hardly be possible without access to a dependable software implementation. To this end, members of the CARAMBA team have been developing the Cado-NFS software suite since 2007. Cado-NFS is now the most widely visible open-source implementation of NFS, and is a crucial platform for developing prototype implementations for new ideas for the many sub-algorithms of NFS. Cado-NFS is free software (LGPL) and follows an open development model, with publicly accessible development repository and regular software releases. Competing free software implementations exist, such as `msieve`, developed by J. Papadopoulos (whose last commit is from August 2018). In Lausanne, T. Kleinjung develops his own code base, which is unfortunately not public.

The work plan of CARAMBA on the topic of the Number Field Sieve algorithm and its cousins includes the following aspects:

- Pursue the work on NFS, which entails in particular making it ready to tackle larger challenges. Several of the important computational steps of NFS that are currently identified as stumbling blocks will require algorithmic advances and implementation improvements. We will illustrate the importance of this work by computational records.
- Work on the specific aspects of the computation of discrete logarithms in finite fields.
- As a side topic, the application of the broad methodology of NFS to the treatment of "ideal lattices" and their use in cryptographic proposals based on Euclidean lattices is also relevant.

3.2. Algebraic Curves for Cryptology

The challenges associated with algebraic curves in cryptology are diverse, because of the variety of mathematical objects to be considered. These challenges are also connected to each other. On the cryptographic side, efficiency matters. With the standardization of TLS 1.3 in 2018 [34], the curves x25519 and x448 have entered the base specification of standard. These curves were designed by academia and offer an excellent compromise between efficiency and security.

On the cryptanalytic side, the discrete logarithm problem on (Jacobians of) curves has resisted all attempts for many years. Among the currently active topics, the decomposition algorithms raise interesting problems related to polynomial system solving, as do attempts to solve the discrete logarithm problem on curves defined over binary fields. In particular, while it is generally accepted that the so-called Koblitz curves (base field extensions of curves defined over $\mathrm{GF}(2)$) are likely to be a weak class among the various curve choices, no concrete attack supports this claim fully.

The research objectives of CARAMBA on the topic of algebraic curves for cryptology are as follows:

- Work on the practical realization of some of the rich mathematical theory behind algebraic curves. In particular, some of the fundamental mathematical objects have potentially important connections to the broad topic of cryptology: Abel-Jacobi map, Theta functions, computation of isogenies, computation of endomorphisms, complex multiplication.
- Improve the point counting algorithms so as to be able to tackle larger problems. This includes significant work connected to polynomial systems.
- Seek improvements on the computation of discrete logarithms on curves, including by identifying weak instances of this problem.

3.3. Symmetric Cryptography

Since the recruiting of Marine Minier in September 2016 as a Professor at the Université de Lorraine, and of Virginie Lallemand as a CNRS researcher in October 2018, a new research domain has emerged in the CARAMBA team: symmetric key cryptology. Accompanied in this adventure by non-permanent team members, we are tackling problems related to both design and analysis. A large part of our recent researches has been motivated by the Lightweight Cryptography Standardization Process of the NIST [0] that embodies a crucial challenge of the last decade: finding ciphers that are suitable for resource-constrained devices.

On a general note, the working program of CARAMBA in symmetric cryptography is defined as follows:

- Develop automatic tools based on constraint programming to help finding optimum attack parameters. The effort will be focused on the AES standard and on recent lightweight cipher proposals.
- Contribute to the security and performance analysis effort required to sort out the candidates for the NIST Lightweight Cryptography Standardization Process.
- Study how to protect services execution on dedicated platforms using white-box cryptography and software obfuscation methods.

3.4. Computer Arithmetic

Computer arithmetic is part of the common background of all team members, and is naturally ubiquitous in our application domains. However involved the mathematical objects considered may be, dealing with them first requires to master more basic objects: integers, finite fields, polynomials, and real and complex floating-point numbers. Libraries such as GNU MP, GNU MPFR, GNU MPC do an excellent job for these, both for small and large sizes (we rarely, if ever, focus on small-precision floating-point data, which explains our lack of mention of libraries relevant to it).

Most of our involvement in subjects related to computer arithmetic is to be understood in connection to our applications to the Number Field Sieve and to abelian varieties. As such, much of the research work we envision will appear as side-effects of developments in these contexts. On the topic of arithmetic work *per se*:

- We will seek algorithmic and practical improvements to the most basic algorithms. That includes for example the study of advanced algorithms for integer multiplication, and their practical reach.
- We will continue to work on the arithmetic libraries in which we have crucial involvement, such as GNU MPFR, GNU MPC, GF2X, MPFQ, and also GMP-ECM.

3.5. Polynomial Systems

Systems of polynomial equations have been part of the cryptographic landscape for quite some time, with applications to the cryptanalysis of block and stream ciphers, as well as multivariate cryptographic primitives.

[0] National Institute of Standard and Technology.

Polynomial systems arising from cryptology are usually not generic, in the sense that they have some distinct structural properties, such as symmetries, or bi-linearity for example. During the last decades, several results have shown that identifying and exploiting these structures can lead to dedicated Gröbner basis algorithms that can achieve large speedups compared to generic implementations [29], [28].

Solving polynomial systems is well done by existing software, and duplicating this effort is not relevant. However we develop test-bed open-source software for ideas relevant to the specific polynomial systems that arise in the context of our applications. The TinyGB software is our platform to test new ideas.

We aim to work on the topic of polynomial system solving in connection with our involvement in the aforementioned topics.

- We have high expertise on Elliptic Curve Cryptography in general. On the narrower topic of the Elliptic Curve Discrete Logarithm Problem on small characteristic finite fields, the highly structured polynomial systems that are involved match well our expertise on the topic of polynomial systems. Once a very hot topic in 2015, activity on this precise problem seems to have slowed down. Yet, the conjunction of skills that we have may lead to results in this direction in the future.
- The hiring of Marine Minier is likely to lead the team to study particular polynomial systems in contexts related to symmetric key cryptography.
- More centered on polynomial systems *per se*, we will mainly pursue the study of the specificities of the polynomial systems that are strongly linked to our targeted applications, and for which we have significant expertise [29], [28]. We also want to see these recent results provide practical benefits compared to existing software, in particular for systems relevant for cryptanalysis.

COAST Project-Team

3. Research Program

3.1. Introduction

Our scientific foundations are grounded on distributed collaborative systems supported by sophisticated data sharing mechanisms and on service oriented computing with an emphasis on orchestration and on non-functional properties. Distributed collaborative systems enable distributed group work supported by computer technologies. Designing such systems requires an expertise in Distributed Systems and in Computer-supported collaborative Work research area. Besides theoretical and technical aspects of distributed systems, the design of distributed collaborative systems must take into account the human factor to offer solutions suitable for users and groups. The Coast team vision is to move away from a centralized authority based collaboration toward a decentralized collaboration. Users will have full control over their data. They can store them locally and decide with whom to share them. The Coast team investigates the issues related to the management of distributed shared data and coordination between users and groups. Service oriented Computing [16] is an established domain on which the ECOO, Score and now the Coast teams have been contributing for a long time. It refers to the general discipline that studies the development of computer applications on the web. A service is an independent software program with a specific functional context and capabilities published as a service contract (or more traditionally an API). A service composition aggregates a set of services and coordinates their interactions. The scale, the autonomy of services, the heterogeneity and some design principles underlying Service Oriented Computing open new research questions that are at the basis of our research. They span the disciplines of **distributed computing**, **software engineering** and **computer supported collaborative work** (CSCW). Our approach to contribute to the general vision of Service Oriented Computing is to focus on the issue of the efficient and flexible construction of reliable and secure high-level services. We aim to achieve it through the coordination/orchestration/composition of other services provided by distributed organizations or people.

3.2. Consistency Models for Distributed Collaborative Systems

Collaborative systems are distributed systems that allow users to share data. One important issue is to manage consistency of shared data according to concurrent access. Traditional consistency criteria such as serializability, linearizability are not adequate for collaborative systems. Causality, Convergence and Intention preservation (CCI) [21] are more suitable for developing middleware for collaborative applications. We develop algorithms for ensuring CCI properties on collaborative distributed systems. Constraints on the algorithms are different according to the kind of distributed system and to the data structure. The distributed system can be centralized, decentralized or peer-to-peer. The type of data can include strings, growable arrays, ordered trees, semantic graphs and multimedia data.

3.3. Optimistic Replication

Replication of data among different nodes of a network promotes reliability, fault tolerance, and availability. When data are mutable, consistency among the different replicas must be ensured. Pessimistic replication is based on the principle of single-copy consistency while optimistic replication allows the replicas to diverge during a short time period. The consistency model for optimistic replication [19] is called eventual consistency, meaning that replicas are guaranteed to converge to the same value when the system is idle. Our research focuses on the two most promising families of optimistic replication algorithms for ensuring CCI:

- operational transformation (OT) algorithms [14]
- algorithms based on commutative replicated data types (CRDT) [18].

Operational transformation algorithms are based on the application of a transformation function when a remote modification is integrated into the local document. Integration algorithms are generic, being parametrised by operational transformation functions which depend on replicated document types. The advantage of these algorithms is their genericity. These algorithms can be applied to any data type and they can merge heterogeneous data in a uniform manner. Commutative replicated data types is a new class of algorithms initiated by WooT [15], the first algorithm designed WithOut Operational Transformations. They ensure consistency of highly dynamic content on peer-to-peer networks. Unlike traditional optimistic replication algorithms, they can ensure consistency without concurrency control. CRDT algorithms rely on natively commutative operations defined on abstract data types such as lists or ordered trees. Thus, they do not require a merge algorithm or an integration procedure.

3.4. Process Orchestration and Management

Process Orchestration and Management is considered as a core discipline behind Service Management and Computing. It includes the analysis, the modelling, the execution, the monitoring and the continuous improvement of enterprise processes and is for us a central domain of studies. Many efforts have been devoted establishing standard business process models founded on well-grounded theories (e.g. Petri Nets) that meet the needs of business analysts, software engineers and software integrator. This led to heated debate in the Business Process Management (BPM) community as the two points of view are very difficult to reconcile. On one side, business people in general require models that are easy to use and understand and that can be quickly adapted to exceptional situations. On the other side, IT people need models with an operational semantic in order to be able transform them into executable artifacts. Part of our work has been an attempt to reconcile these points of view. This resulted in the development of the Bonita BPM system. It resulted also more recently on our work in crisis management where the same people are designing, executing and monitoring the process as it executes. More generally, and at a larger scale, we have been considering the problem of processes spanning the barriers of organizations. This leads to the more general problem of service composition as a way to coordinate inter organizational construction of applications. These applications provide value, based on the composition of lower level services [12].

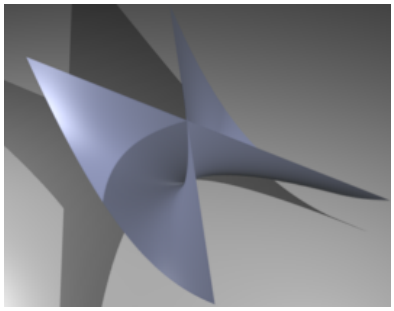
3.5. Service Composition

Recently, we started a study on service composition for software architects where services are coming from different providers with different plans (capacity, degree of resilience...). The objective is to support the architects to select the most accurate services (wrt. to their requirements, both functional and non-functional) and plans for building their software. We also compute the properties that we enforce for the composition of these services.

GAMBLE Project-Team

3. Research Program

3.1. Non-linear computational geometry



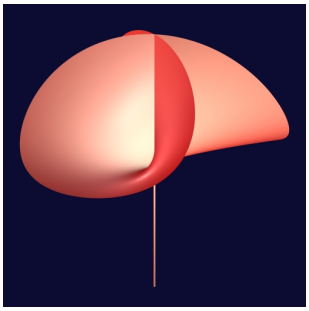


Figure 1. Two views of the Whitney umbrella (on the left, the "stick" of the umbrella, i.e., the negative z-axis, is missing). Right picture from [Wikipedia], left picture from [Lachaud et al.].

As mentioned above, curved objects are ubiquitous in real world problems and in computer science and, despite this fact, there are very few problems on curved objects that admit robust and efficient algorithmic solutions without first discretizing the curved objects into meshes. Meshing curved objects induces a loss of accuracy which is sometimes not an issue but which can also be most problematic depending on the application. In addition, discretization induces a combinatorial explosion which could cause a loss in efficiency compared to a direct solution on the curved objects (as our work on quadrics has demonstrated with flying colors [50], [51], [52], [54], [58]). But it is also crucial to know that even the process of computing meshes that approximate curved objects is far from being resolved. As a matter of fact there is no algorithm capable of computing in practice meshes with certified topology of even rather simple singular 3D surfaces, due to the high constants in the theoretical complexity and the difficulty of handling degenerate cases. Part of the difficulty comes from the unintuitive fact that the structure of an algebraic object can be quite complicated, as depicted in the Whitney umbrella (see Figure 1), surface of equation $x^2 = y^2 z$ on which the origin (the "special" point of the surface) is a vertex of the arrangement induced by the surface while the singular locus is simply the whole z-axis. Even in 2D, meshing an algebraic curve with the correct topology, that is in other words producing a correct drawing of the curve (without knowing where the domain of interest is), is a very difficult problem on which we have recently made important contributions [37], [38], [59].

It is thus to be understood that producing practical robust and efficient algorithmic solutions to geometric problems on curved objects is a challenge on all and even the most basic problems. The basicness and fundamentality of two problems we mentioned above on the intersection of 3D quadrics and on the drawing in a topologically certified way of plane algebraic curves show rather well that the domain is still in its infancy. And it should be stressed that these two sets of results were not anecdotal but flagship results produced during the lifetime of the VEGAS team (the team preceding GAMBLE).

There are many problems in this theme that are expected to have high long-term impacts. Intersecting NURBS (Non-uniform rational basis splines) in a certified way is an important problem in computer-aided design and manufacturing. As hinted above, meshing objects in a certified way is important when topology matters. The 2D case, that is essentially drawing plane curves with the correct topology, is a fundamental problem with

far-reaching applications in research or R&D. Notice that on such elementary problems it is often difficult to predict the reach of the applications; as an example, we were astonished by the scope of the applications of our software on 3D quadric intersection [0] which was used by researchers in, for instance, photochemistry, computer vision, statistics and mathematics.

3.2. Non-Euclidean computational geometry

Figure 2. Left: 3D mesh of a gyroid (triply periodic surface) [61]. Right: Simulation of a periodic Delaunay triangulation of the hyperbolic plane [33].

Triangulations, in particular Delaunay triangulations, in the *Euclidean space* $\mathbb{R}^d$ have been extensively studied throughout the 20th century and they are still a very active research topic. Their mathematical properties are now well understood, many algorithms to construct them have been proposed and analyzed (see the book of Aurenhammer *et al.* [32]). Some members of GAMBLE have been contributing to these algorithmic advances (see, e.g. [36], [68], [47], [35]); they have also contributed robust and efficient triangulation packages through the state-of-the-art Computational Geometry Algorithms Library CGAL whose impact extends far beyond computational geometry. Application fields include particle physics, fluid dynamics, shape matching, image processing, geometry processing, computer graphics, computer vision, shape reconstruction, mesh generation, virtual worlds, geophysics, and medical imaging. [0]

It is fair to say that little has been done on non-Euclidean spaces, in spite of the large number of questions raised by application domains. Needs for simulations or modeling in a variety of domains [0] ranging from the infinitely small (nuclear matter, nano-structures, biological data) to the infinitely large (astrophysics) have led us to consider 3D periodic Delaunay triangulations, which can be seen as Delaunay triangulations in the 3D *flat torus*, quotient of $\mathbb{R}^3$ under the action of some group of translations [42]. This work has already yielded a fruitful collaboration with astrophysicists [55], [69] and new collaborations with physicists are emerging. To the best of our knowledge, our CGAL package [41] is the only publicly available software that computes Delaunay triangulations of a 3D flat torus, in the special case where the domain is cubic. This case, although restrictive, is already useful. [0] We have also generalized this algorithm to the case of general d-dimensional compact flat manifolds [43]. As far as non-compact manifolds are concerned, past approaches, limited to the two-dimensional case, have stayed theoretical [60].

Interestingly, even for the simple case of triangulations on the *sphere*, the software packages that are currently available are far from offering satisfactory solutions in terms of robustness and efficiency [40].

[0] QI: web.
[0] See Projects using CGAL for details.
[0] See CGAL Prospective Workshop on Geometric Computing in Periodic Spaces, Subdivide and Tile: Triangulating spaces for understanding the world, Computational geometry in non-Euclidean spaces, Shape Up 2015 : Exercises in Materials Geometry and Topology
[0] See examples at Projects using CGAL

Moreover, while our solution for computing triangulations in hyperbolic spaces can be considered as ultimate [33], the case of *hyperbolic manifolds* has hardly been explored. Hyperbolic manifolds are quotients of a hyperbolic space by some group of hyperbolic isometries. Their triangulations can be seen as hyperbolic periodic triangulations. Periodic hyperbolic triangulations and meshes appear for instance in geometric modeling [62], neuromathematics [45], or physics [65]. Even the case of the Bolza surface (a surface of genus 2, whose fundamental domain is the regular octagon in the hyperbolic plane) shows mathematical difficulties [34], [57].

3.3. Probability in computational geometry

In most computational geometry papers, algorithms are analyzed in the worst-case setting. This often yields too pessimistic complexities that arise only in pathological situations that are unlikely to occur in practice. On the other hand, probabilistic geometry provides analyses with great precision [63], [64], [39], but using hypotheses with much more randomness than in most realistic situations. We are developing new algorithmic designs improving state-of-the-art performance in random settings that are not overly simplified and that can thus reflect many realistic situations.

Twelve years ago, smooth analysis was introduced by Spielman and Teng analyzing the simplex algorithm by averaging on some noise on the data [67] (and they won the Gödel prize). In essence, this analysis smoothes the complexity around worst-case situations, thus avoiding pathological scenarios but without considering unrealistic randomness. In that sense, this method makes a bridge between full randomness and worst case situations by tuning the noise intensity. The analysis of computational geometry algorithms within this framework is still embryonic. To illustrate the difficulty of the problem, we started working in 2009 on the smooth analysis of the size of the convex hull of a point set, arguably the simplest computational geometry data structure; then, only one very rough result from 2004 existed [46] and we only obtained in 2015 breakthrough results, but still not definitive [49], [48], [53].

Another example of a problem of different flavor concerns Delaunay triangulations, which are rather ubiquitous in computational geometry. When Delaunay triangulations are computed for reconstructing meshes from point clouds coming from 3D scanners, the worst-case scenario is, again, too pessimistic and the full randomness hypothesis is clearly not adapted. Some results exist for "good samplings of generic surfaces" [31] but the big result that everybody wishes for is an analysis for random samples (without the extra assumptions hidden in the "good" sampling) of possibly non-generic surfaces.

Trade-offs between full randomness and worst case may also appear in other forms such as dependent distributions, or random distributions conditioned to be in some special configurations. Simulating these kinds of geometric distributions is currently out of reach for more than a few hundred points [56] although it has practical applications in physics or networks.

3.4. Discrete geometric structures

Our work on discrete geometric structures develops in several directions, each one probing a different type of structure. Although these objects appear unrelated at first sight, they can be tackled by the same set of probabilistic and topological tools.

A first research topic is the study of *Order types*. Order types are combinatorial encodings of finite (planar) point sets, recording for each triple of points the orientation (clockwise or counterclockwise) of the triangle they form. This already determines properties such as convex hulls or half-space depths, and the behaviour of algorithms based on orientation predicates. These properties for all (infinitely many) n-point sets can be studied through the finitely many order types of size n. Yet, this finite space is poorly understood: its estimated size leaves an exponential margin of error, no method is known to sample it without concentrating on a vanishingly small corner, the effect of pattern exclusion or VC dimension-type restrictions are unknown. These are all directions we actively investigate.

A second research topic is the study of *Embedded graphs and simplicial complexes.* Many topological structures can be effectively discretized, for instance combinatorial maps record homotopy classes of embedded graphs and simplicial complexes represent a large class of topological spaces. This raises many structural and algorithmic questions on these discrete structures; for example, given a closed walk in an embedded graph, can we find a cycle of the graph homotopic to that walk? (The complexity status of that problem is unknown.) Going in the other direction, some purely discrete structures can be given an associated topological space that reveals some of their properties (*e.g.* the Nerve theorem for intersection patterns). An open problem is for instance to obtain fractional Helly theorems for set system of bounded topological complexity.

Another research topic is that of *Sparse inclusion-exclusion formulas.* For any family of sets $A_1, A_2, ..., A_n$, by the principle of inclusion-exclusion we have

$$\mathbb{1}_{\bigcup_{i=1}^{n} A_i} = \sum_{I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|+1} \mathbb{1}_{\bigcap_{i \in I} A_i} \tag{4}$$

where $\mathbb{1}_X$ is the indicator function of X. This formula is universal (it applies to any family of sets) but its number of summands grows exponentially with the number n of sets. When the sets are balls, the formula remains true if the summation is restricted to the regular triangulation; we proved that similar simplifications are possible whenever the Venn diagram of the A_i is sparse. There is much room for improvements, both for general set systems and for specific geometric settings. Another interesting problem (the subject of the PhD thesis of Galatée Hemery) is to combine these simplifications with the inclusion-exclusion algorithms developed, for instance, for graph coloring.

LARSEN Project-Team

3. Research Program

3.1. Lifelong Autonomy

3.1.1. Scientific Context

So far, only a few autonomous robots have been deployed for a long time (weeks, months, or years) outside of factories and laboratories. They are mostly mobile robots that simply "move around" (e.g., vacuum cleaners or museum "guides") and data collecting robots (e.g., boats or underwater "gliders" that collect data about the water of the ocean).

A large part of the long-term autonomy community is focused on simultaneous localization and mapping (SLAM), with a recent emphasis on changing and outdoor environments [25], [34]. A more recent theme is life-long learning: during long-term deployment, we cannot hope to equip robots with everything they need to know, therefore some things will have to be learned along the way. Most of the work on this topic leverages machine learning and/or evolutionary algorithms to improve the ability of robots to react to unforeseen changes [25], [32].

3.1.2. Main Challenges

The first major challenge is to endow robots with a stable situation awareness in open and dynamic environments. This covers both the state estimation of the robot itself as well as the perception/representation of the environment. Both problems have been claimed to be solved but it is only the case for static environments [30].

In the LARSEN team, we aim at deployment in environments shared with humans which imply dynamic objects that degrade both the mapping and localization of a robot, especially in cluttered spaces. Moreover, when robots stay longer in the environment than for the acquisition of a snapshot map, they have to face structural changes, such as the displacement of a piece of furniture or the opening or closing of a door. The current approach is to simply update an implicitly static map with all observations with no attempt at distinguishing the suitable changes. For localization in not-too-cluttered or not-too-empty environments, this is generally sufficient as a significant fraction of the environment should remain stable. But for life-long autonomy, and in particular navigation, the quality of the map, and especially the knowledge of the stable parts, is primordial.

A second major obstacle to move robots outside of labs and factories is their fragility: Current robots often break in a few hours, if not a few minutes. This fragility mainly stems from the overall complexity of robotic systems, which involve many actuators, many sensors, and complex decisions, and from the diversity of situations that robots can encounter. Low-cost robots exacerbate this issue because they can be broken in many ways (high-quality material is expensive), because they have low self-sensing abilities (sensors are expensive and increase the overall complexity), and because they are typically targeted towards non-controlled environments (e.g., houses rather than factories, in which robots are protected from most unexpected events). More generally, this fragility is a symptom of the lack of adaptive abilities in current robots.

3.1.3. Angle of Attack

To solve the state estimation problem, our approach is to combine classical estimation filters (Extended Kalman Filters, Unscented Kalman Filters, or particle filters) with a Bayesian reasoning model in order to internally simulate various configurations of the robot in its environment. This should allow for adaptive estimation that can be used as one aspect of long-term adaptation. To handle dynamic and structural changes in an environment, we aim at assessing, for each piece of observation, whether it is static or not.

We also plan to address active sensing to improve the situation awareness of robots. Literally, active sensing is the ability of an interacting agent to act so as to control what it senses from its environment with the typical objective of acquiring information about this environment. A formalism for representing and solving active sensing problems has already been proposed by members of the team [24] and we aim to use this to formalize decision making problems of improving situation awareness.

Situation awareness of robots can also be tackled by cooperation, whether it be between robots or between robots and sensors in the environment (led out intelligent spaces) or between robots and humans. This is in rupture with classical robotics, in which robots are conceived as self-contained. But, in order to cope with as diverse environments as possible, these classical robots use precise, expensive, and specialized sensors, whose cost prohibits their use in large-scale deployments for service or assistance applications. Furthermore, when all sensors are on the robot, they share the same point of view on the environment, which is a limit for perception. Therefore, we propose to complement a cheaper robot with sensors distributed in a target environment. This is an emerging research direction that shares some of the problematics of multi-robot operation and we are therefore collaborating with other teams at Inria that address the issue of communication and interoperability.

To address the fragility problem, the traditional approach is to first diagnose the situation, then use a planning algorithm to create/select a contingency plan. But, again, this calls for both expensive sensors on the robot for the diagnosis and extensive work to predict and plan for all the possible faults that, in an open and dynamic environment, are almost infinite. An alternative approach is then to skip the diagnosis and let the robot discover by trial and error a behavior that works in spite of the damage with a reinforcement learning algorithm [39], [32]. However, current reinforcement learning algorithms require hundreds of trials/episodes to learn a single, often simplified, task [32], which makes them impossible to use for real robots and more ambitious tasks. We therefore need to design new trial-and-error algorithms that will allow robots to learn with a much smaller number of trials (typically, a dozen). We think the key idea is to guide online learning on the physical robot with dynamic simulations. For instance, in our recent work, we successfully mixed evolutionary search in simulation, physical tests on the robot, and machine learning to allow a robot to recover from physical damage [33], [1].

A final approach to address fragility is to deploy several robots or a swarm of robots or to make robots evolve in an active environment. We will consider several paradigms such as (1) those inspired from collective natural phenomena in which the environment plays an active role for coordinating the activity of a huge number of biological entities such as ants and (2) those based on online learning [29]. We envision to transfer our knowledge of such phenomenon to engineer new artificial devices such as an intelligent floor (which is in fact a spatially distributed network in which each node can sense, compute and communicate with contiguous nodes and can interact with moving entities on top of it) in order to assist people and robots (see the principle in [37], [29], [23]).

3.2. Natural Interaction with Robotic Systems

3.2.1. Scientific Context

Interaction with the environment is a primordial requirement for an autonomous robot. When the environment is sensorized, the interaction can include localizing, tracking, and recognizing the behavior of robots and humans. One specific issue lies in the lack of predictive models for human behavior and a critical constraint arises from the incomplete knowledge of the environment and the other agents.

On the other hand, when working in the proximity of or directly with humans, robots must be capable of safely interacting with them, which calls upon a mixture of physical and social skills. Currently, robot operators are usually trained and specialized but potential end-users of robots for service or personal assistance are not skilled robotics experts, which means that the robot needs to be accepted as reliable, trustworthy and efficient [42]. Most Human-Robot Interaction (HRI) studies focus on verbal communication [38] but applications such as assistance robotics require a deeper knowledge of the intertwined exchange of social and physical signals to provide suitable robot controllers.

3.2.2. Main Challenges

We are here interested in building the bricks for a situated Human-Robot Interaction (HRI) addressing both the physical and social dimension of the close interaction, and the cognitive aspects related to the analysis and interpretation of human movement and activity.

The combination of physical and social signals into robot control is a crucial investigation for assistance robots [40] and robotic co-workers [36]. A major obstacle is the control of physical interaction (precisely, the control of contact forces) between the robot and the human while both partners are moving. In mobile robots, this problem is usually addressed by planning the robot movement taking into account the human as an obstacle or as a target, then delegating the execution of this "high-level" motion to whole-body controllers, where a mixture of weighted tasks is used to account for the robot balance, constraints, and desired end-effector trajectories [26].

The first challenge is to make these controllers easier to deploy in real robotics systems, as currently they require a lot of tuning and can become very complex to handle the interaction with unknown dynamical systems such as humans. Here, the key is to combine machine learning techniques with such controllers.

The second challenge is to make the robot react and adapt online to the human feedback, exploiting the whole set of measurable verbal and non-verbal signals that humans naturally produce during a physical or social interaction. Technically, this means finding the optimal policy that adapts the robot controllers online, taking into account feedback from the human. Here, we need to carefully identify the significant feedback signals or some metrics of human feedback. In real-world conditions (i.e., outside the research laboratory environment) the set of signals is technologically limited by the robot's and environmental sensors and the onboard processing capabilities.

The third challenge is for a robot to be able to identify and track people on board. The motivation is to be able to estimate online either the position, the posture, or even moods and intentions of persons surrounding the robot. The main challenge is to be able to do that online, in real-time and in cluttered environments.

3.2.3. Angle of Attack

Our key idea is to exploit the physical and social signals produced by the human during the interaction with the robot and the environment in controlled conditions, to learn simple models of human behavior and consequently to use these models to optimize the robot movements and actions. In a first phase, we will exploit human physical signals (e.g., posture and force measurements) to identify the elementary posture tasks during balance and physical interaction. The identified model will be used to optimize the robot whole-body control as prior knowledge to improve both the robot balance and the control of the interaction forces. Technically, we will combine weighted and prioritized controllers with stochastic optimization techniques. To adapt online the control of physical interaction and make it possible with human partners that are not robotics experts, we will exploit verbal and non-verbal signals (e.g., gaze, touch, prosody). The idea here is to estimate online from these signals the human intent along with some inter-individual factors that the robot can exploit to adapt its behavior, maximizing the engagement and acceptability during the interaction.

Another promising approach already investigated in the LARSEN team is the capability for a robot and/or an intelligent space to localize humans in its surrounding environment and to understand their activities. This is an important issue to handle both for safe and efficient human-robot interaction.

Simultaneous Tracking and Activity Recognition (STAR) [41] is an approach we want to develop. The activity of a person is highly correlated with his position, and this approach aims at combining tracking and activity recognition to benefit one from another. By tracking the individual, the system may help infer its possible activity, while by estimating the activity of the individual, the system may make a better prediction of his/her possible future positions (especially in the case of occlusions). This direction has been tested with simulator and particle filters [28], and one promising direction would be to couple STAR with decision making formalisms like partially observable Markov decision processes (POMDPs). This would allow us to formalize problems such as deciding which action to take given an estimate of the human location and activity. This could also formalize other problems linked to the active sensing direction of the team: how the robotic system

should choose its actions in order to have a better estimate of the human location and activity (for instance by moving in the environment or by changing the orientation of its cameras)?

Another issue we want to address is robotic human body pose estimation. Human body pose estimation consists of tracking body parts by analyzing a sequence of input images from single or multiple cameras.

Human posture analysis is of high value for human robot interaction and activity recognition. However, even if the arrival of new sensors like RGB-D cameras has simplified the problem, it still poses a great challenge, especially if we want to do it online, on a robot and in realistic world conditions (cluttered environment). This is even more difficult for a robot to bring together different capabilities both at the perception and navigation level [27]. This will be tackled through different techniques, going from Bayesian state estimation (particle filtering), to learning, active and distributed sensing.

MAGRIT Team

3. Research Program

3.1. Matching and 3D tracking

One of the most basic problems currently limiting AR applications is the registration problem. The objects in the real and virtual worlds must be properly aligned with respect to each other, or the illusion that the two worlds coexist will be compromised.

As a large number of potential AR applications are interactive, real time pose computation is required. Although the registration problem has received a lot of attention in the computer vision community, the problem of real-time registration is still far from being a solved problem, especially for unstructured environments. Ideally, an AR system should work in all environments, without the need to prepare the scene ahead of time, independently of the variations in experimental conditions (lighting, weather condition,...)

For several years, the MAGRIT project has been aiming at developing on-line and marker-less methods for camera pose computation. The main difficulty with on-line tracking is to ensure robustness of the process over time. For off-line processes, robustness is achieved by using spatial and temporal coherence of the considered sequence through move-matching techniques. To get robust open-loop systems, we have investigated various methods, ranging from statistical methods to the use of hybrid camera/sensor systems. Many of these methods are dedicated to piecewise-planar scenes and combine the advantage of move-matching methods and model-based methods. In order to reduce statistical fluctuations in viewpoint computation, which lead to unpleasant jittering or sliding effects, we have also developed model selection techniques which allow us to noticeably improve the visual impression and to reduce drift over time. Another line of research which has been considered in the team to improve the reliability and the robustness of pose algorithms is to combine the camera with another form of sensor in order to compensate for the shortcomings of each technology.

The success of pose computation over time largely depends on the quality of the matching at the initialization stage. Indeed, the current image may be very different from the appearances described in the model both on the geometrical and the photometric sides. Research is thus conducted in the team on the use of probabilistic methods to establish robust correspondences of features. The use of *a contrario* methods has been investigated to achieve this aim [7]. We especially addressed the complex case of matching in scenes with repeated patterns which are common in urban scenes. We are also investigating the problem of matching images taken from very different viewpoints which is central for the re-localization issue in AR. Within the context of a scene model acquired with structure-from-motion techniques, we are currently investigating the use of viewpoint simulation in order to allow successful pose computation even if the considered image is far from the positions used to build the model [15].

Recently, the issue of tracking deformable objects has gained importance in the team. This topic is mainly addressed in the context of medical applications through the design of bio-mechanical models guided by visual features [2]. We have successfully investigated the use of such models in laparoscopy, with a vascularized model of the liver and with a hyper-elastic model for tongue tracking in ultrasound images. However, these results have been obtained so far in relatively controlled environments, with non-pathological cases. When clinical routine applications are to be considered, many parameters and considerations need to be taken into account. Among the problems that need to be addressed are more realistic model representations, the specification of the range of physical parameters and the need to enforce the robustness of the tracking with respect to outliers, which are common in the interventional context.

3.2. Image-based Modeling

Modeling the scene is a fundamental issue in AR for many reasons. First, pose computation algorithms often use a model of the scene or at least some 3D knowledge on the scene. Second, effective AR systems require a

model of the scene to support interactions between the virtual and the real objects such as occlusions, lighting reflections, contacts... in real-time. Unlike pose computation which has to be performed in a sequential way, scene modeling can be considered as an off-line or an on-line problem depending on the requirements of the targeted application. Interactive in-situ modeling techniques have thus been developed with the aim to enable the user to define what is relevant at the time the model is being built during the application. On the other hand, we also proposed off-line multimodal techniques, mainly dedicated to AR medical applications, with the aim of obtaining realistic and possibly dynamic models of organs suitable for real-time simulation [3].

In-situ modeling
In-situ modeling allows a user to directly build a 3D model of his/her surrounding environment and verify the geometry against the physical world in real-time. This is of particular interest when using AR in unprepared environments or building scenes that either have an ephemeral existence (e.g., a film set) or cannot be accessed frequently (e.g., a nuclear power plant). We have especially investigated two systems, one based on the image content only and the other based on multiple data coming from different sensors (camera, inertial measurement unit, laser rangefinder). Both systems use the camera-mouse principle [34] (i.e., interactions are performed by aiming at the scene through a video camera) and both systems have been designed to acquire polygonal textured models, which are particularly useful for camera tracking and object insertion in AR.

Multimodal modeling for real-time simulation
With respect to classical AR applications, AR in medical context differs in the nature and the size of the data which are available: a large amount of multimodal data is acquired on the patient or possibly on the operating room through sensing technologies or various image acquisitions [32]. The challenge is to analyze these data, to extract interesting features, to fuse and to visualize this information in a proper way. Within the MAGRIT team, we address several key problems related to medical augmented environments. Being able to acquire multimodal data which are temporally synchronized and spatially registered is the first difficulty we face when considering medical AR. Another key requirement of AR medical systems is the availability of 3D (+t) models of the organ/patient built from images, to be overlaid onto the users' view of the environment.

Methods for multimodal modeling are strongly dependent on the imaging modalities and the organ specificities. We thus only address a restricted number of medical applications –interventional neuro-radiology, laparoscopic surgery– for which we have a strong expertise and close relationships with motivated clinicians. In these applications, our aim is to produce realistic models and then realistic simulations of the patient to be used for the training of surgeons or the re-education of patients.

One of our main applications is about neuroradiology. For the last 20 years, we have been working in close collaboration with the neuroradiology laboratory (CHRU-University Hospital of Nancy) and GE Healthcare. As several imaging modalities are now available in an intraoperative context (2D and 3D angiography, MRI, ...), our aim is to develop a multi-modality framework to assist therapeutic decision and treatment.

We have mainly been interested in the effective use of a multimodality framework in the treatment of arteriovenous malformations (AVM) and aneurysms in the context of interventional neuroradiology. The goal of interventional gestures is to guide endoscopic tools towards the pathology with the aim to perform embolization of the AVM or to fill the aneurysmal cavity by placing coils. We have proposed and developed multimodality and augmented reality tools which make various image modalities (2D and 3D angiography, fluoroscopic images, MRI, ...) cooperate in order to assist physicians in clinical routine. One of the successes of this collaboration is the implementation of the concept of *augmented fluoroscopy*, which helps the surgeon to guide endoscopic tools towards the pathology. Lately, in cooperation with the team MIMESIS, we have proposed new methods for implicit modeling of the vasculature with the aim of obtaining near real-time simulation of the coil deployment in the aneurysm [3]. These works open the way towards near real-time patient-based simulations of interventional gestures both for training and for planning.

3.3. Parameter estimation

Many problems in computer vision or image analysis can be formulated in terms of parameter estimation from image-based measurements. This is the case of many problems addressed in the team such as pose

computation or image-guided estimation of 3D deformable models. Often traditional robust techniques which take into account the covariance on the measurements are sufficient to achieve reliable parameter estimation. However, depending on their number, their spatial distribution and the uncertainty on these measurements, some problems are very sensitive to noise and there is a considerable interest in considering how parameter estimation could be improved if additional information on the noise were available. Another common problem in our field of research is the need to estimate constitutive parameters of the models, such as (bio)-mechanical parameters for instance. Direct measurement methods are destructive, and elaborating image-based methods is thus highly desirable. Besides designing appropriate estimation algorithms, a fundamental question is to understand what group of parameters under study can be reliably estimated from a given experimental setup.

This line of research is relatively new in the team. One of the challenges is to improve image-based parameter estimation techniques considering sensor noise and specific image formation models. In a collaboration with the Pascal Institute (Clermont Ferrand), metrological performance enhancement for experimental solid mechanics has been addressed through the development of dedicated signal processing methods [6]. In the medical field, specific methods based on an adaptive evolutionary optimization strategy have been designed for estimating respiratory parameters [8]. In the context of designing realistic simulators for neuroradiology, we are now considering how parameters involved in the simulation could be adapted to fit real images.

MFX Project-Team

3. Research Program

3.1. Research Program

We focus on the computational aspects of shape modeling and processing for digital fabrication: dealing with shape complexity, revisiting design and customization of existing parts in view of the novel possibilities afforded by AM, and providing a stronger integration between modeling and the capabilities of the target processes.
We tackle on the following challenges:

- develop **novel shape synthesis and shape completion algorithms** that can help users model shapes with features in the scale of microns to meters, while following functional, structural, geometric and fabrication requirements;
- propose methodologies to help *expert* designers **describe shapes** and designs that can later be **customized and adapted** to different use cases;
- develop novel algorithms to **adapt and prepare complex designs** for fabrication in a given technology, including the possibility to modify aspects of the design while preserving its functionality;
- develop novel techniques to **unlock the full potential of fabrication processes**, improving their versatility in terms of feasible shapes as well as their capabilities in terms of accuracy and quality of deposition;
- develop **novel shape representations, data-structures, visualization and interaction techniques** to support the integration of our approaches into a single, unified software framework that covers the full chain from modeling to printing instructions;
- **integrate novel capabilities** enabled by advances in additive manufacturing processes and materials **in the modeling and processing chains**, in particular regarding the use of functional materials (*e.g.* piezoelectric, conductive, shrinkable).

Our approach is to cast a holistic view on the aforementioned challenges, by considering modeling and fabrication as a single, unified process. Thus, the modeling techniques we seek to develop will take into account the geometric constraints imposed by the manufacturing processes (minimal thickness, overhang angles, trapped material) as well as the desired object functionality (rigidity, porosity). To allow for the modeling of complex shapes, and to adapt the same initial design to different technologies, we propose to develop techniques that can automatically synthesize functional details within parts. At the same time, we will explore ways to increase the versatility of the manufacturing processes, through algorithms that are capable of exploiting additional degrees of freedom (*e.g.*, curved layering [21]), can introduce new capabilities (*e.g.*, material mixing [22]) and improve part accuracy (*e.g.*, adaptive slicing [20]).

Our research program is organized along three main research directions. The first one focuses on the automatic synthesis of shapes with intricate multi-scale geometries, that conform to the constraints of additive manufacturing technologies. The second direction considers geometric and algorithmic techniques for the actual fabrication of the modeled object. We aim to further improve the capabilities of the manufacturing processes with novel deposition strategies. The third direction focuses on computational design algorithms to help model parts with gradient of properties, as well as to help customizing existing designs for their reuse.

These three research directions interact strongly, and cross-pollinate: *e.g.*, novel possibilities in manufacturing unlock novel possibilities in terms of shapes that can be synthesized. Stronger synthesis methods allow for further customization.

MIMESIS Team

3. Research Program

3.1. Real-time computational models for interactive applications

The principal objective of this challenge is to improve, at the numerical level, the efficiency, robustness, and quality of the simulations (see Fig. 2). An important part of our research is dedicated to the development of computational models that remain compatible with real-time computation, i.e., which allow immediate visual or haptic feedback. This typically requires computation times below $50ms$ and in some cases around $1ms$. Such advanced models can not only increase the realism of future training systems, but also act as a bridge toward the development of patient-specific solutions for computer-aided interventions. Additionally, such simulations should run on (high-end) consumer level computers (i.e. with a single multi-core CPU and a dedicated GPU). To reach these goals, we are investigating novel finite element techniques able to cope with complex, potentially ill-defined input data. After developing Smoothed FEM for real-time simulations, we are developing meshless techniques and immersed boundary methods. The first one is well suited for topological changes, which we sometimes need to account for in our simulations. The second is expected to lead to more stable, and numerically efficient, formulations of the finite element method. We are also developing numerical techniques to compute the complex interactions that can take place between anatomical structures or between medical devices and organs. Boundary conditions are known to also play an important role in the solution of such problems. Therefore we are investigating solutions to both identify and model the interactions that take place between the structure of interest and its anatomical environment.

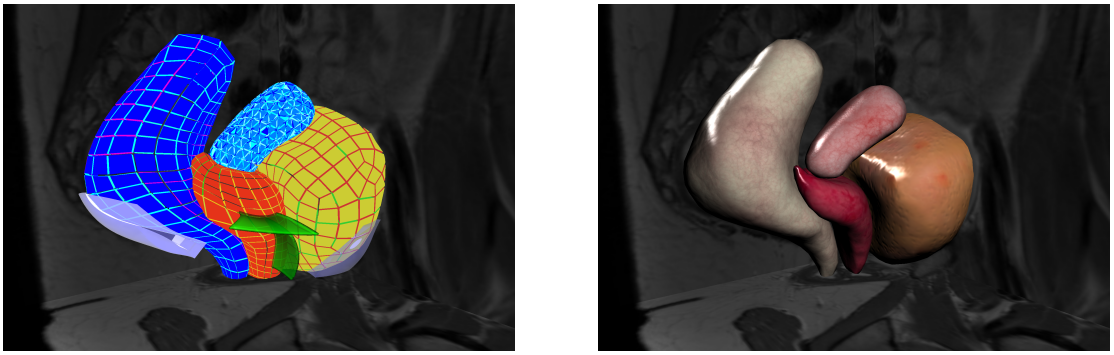


Figure 2. Model of the pelvis with (left) the finite element models of different anatomical structures and (right) their visual representations. Complex interactions take place between these deformable structures. The simulation is computed at interactive rates

3.2. Data-driven simulations

Data-driven simulation has been a recent area of research in our team (see Fig. 3). We have demonstrated that it has the potential to bridge the gap between medical imaging and clinical routine by adapting pre-operative data to the time of the procedure. In the areas of non-rigid registration and augmented reality during surgery, we have demonstrated the benefit of our physics-based approaches with several key publications in major conferences (MICCAI, CVPR, IPCAI, ISMAR).

We have continued this work with an **emphasis on robustness to uncertainty and outliers** in the information extracted in real-time from image data, as well as real-time parameter estimation. This is currently done by **combining Bayesian methods with advanced physics-based methods** to handle uncertainties in image-driven simulations (MICCAI 2017, CVCS 2018).

Finally, Bayesian or similar methods require to perform a large amount of simulations to sample the domain space, even when using efficient methods such as Reduced Order Unscented Kalman Filters. For this reason, we are investigating the use of neural networks to perform predictions instead of using full numerical simulations. Our latest paper [22] at MICCAI 2019 shows it is possible to **teach a neural network from numerical simulations** and **predict**, with good accuracy, **the deformation of an organ**.

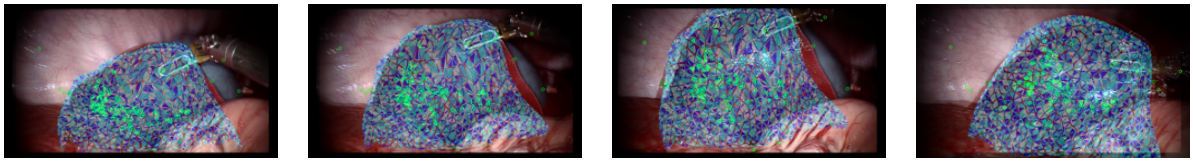


Figure 3. Real-time deformation of a virtual liver according to tissue motion tracked in laparoscopic images.

MOCQUA Team

3. Research Program

3.1. Quantum Computing

While it can be argued that the quantum revolution has already happened in cryptography [39] or in optics [38], quantum computers are far from becoming a common commodity, with only a few teams around the world working on a practical implementation. In fact, one of the most commonly known examples of a quantum computer, the D-Wave 2X System, defies the usual definition of a computer: it is not general-purpose, and can only solve (approximately) a very specific hardwired problem.

Most current prototypes of a quantum computer differ fundamentally on the hardware substrate, and it is quite hard to predict which solution will finally be adopted. The landscape of quantum programming languages is also constantly evolving. Comparably to compiler design, the foundation of quantum software therefore relies on an intermediate representation that is suitable for manipulation, easy to produce from software and easily encodable into hardware. The language of choice for this is the ZX-calculus.

Regardless of the actual model that will be accepted by the industry, it is becoming clear that some of the hurdles into scaling up quantum computers from a few qubits to very large arrays will remain. As an example, current implementations of quantum computers working on hundreds of qubits indeed are not able to form and maintain all possible forms of entanglement between qubits. This raises two questions. First, does this restrict the computational power, and the supposed advantage of the quantum computer over the classical computer? Second, how to ensure that a quantum program that was designed for a theoretical quantum computer will work on the practical implementations? This will be investigated, in particular by providing static analysis methods for evaluating a priori how much entanglement a quantum program needs.

3.2. Higher-Order Computing

While programs often operate on natural numbers or finite structures such as graphs or finite strings, they can also take functions as input. In that case, the program is said to perform higher-order computations, or to compute a higher-order functional. Functional programming or object-oriented programming are important paradigms allowing higher-order computations.

While the theory of computation is well developed for first-order programs, difficulties arise when dealing with higher-order programs. There are many non-equivalent ways of presenting inputs to such programs: an input function can be presented as a black-box, encoded in an infinite binary sequence, or sometimes by a finite description. Comparing those representations is an important problem. A particularly useful application of higher-order computations is to compute with infinite objects that can be represented by functions or symbolic sequences. The theory works well in many cases (to be precise, when these objects live in a topological space with a countable basis [42]), but is not well understood in other interesting cases. For instance, when the inputs are the second-order functionals (of type $(\mathbb{N} \to \mathbb{N}) \to (\mathbb{N} \to \mathbb{N})$), the classical theory does not apply and many problems are still open.

3.3. Dynamical Systems

The most natural example of a computation with infinite precision is the simulation of a dynamical system. The underlying space might be $\mathbb{R}^n$ in the case of the simulation of physical systems, or the Cantor space $\{0,1\}^{\mathbb{Z}}$ in the case of discrete dynamical systems.

From the point of view of computation, the main point of interest is the link between the long-term behavior of a system and its initial configuration. There are two questions here: (a) predict the behavior, (b) design dynamical systems with some prescribed behavior. The first will be mainly examined through the angle of reachability and more generally control theory for hybrid systems.

The model of cellular automata will be of particular interest. This computational model is relevant for simulating complex global phenomena which emerge from simple interactions between simple components. It is widely used in various natural sciences (physics, biology, etc.) and in computer science, as it is an appropriate model to reason about errors that occur in systems with a great number of components.

The simulation of a physical dynamical system on a computer is made difficult by various aspects. First, the parameters of the dynamical systems are seldom exactly known. Secondly, the simulation is usually non exact: real numbers are usually represented by floating-point numbers, and simulations of cellular automata only simulate the behavior of finite or periodic configurations. For some chaotic systems, this means that the simulation can be completely irrelevant.

MULTISPEECH Project-Team

3. Research Program

3.1. Beyond black-box supervised learning

This research axis focuses on fundamental, domain-agnostic challenges relating to deep learning, such as the integration of domain knowledge, data efficiency, or privacy preservation. The results of this axis naturally apply in the domains studied in the two other research axes.

3.1.1. Integrating domain knowledge

State-of-the-art methods in speech and audio are based on neural networks trained for the targeted task. This paradigm faces major limitations: lack of interpretability and of guarantees, large data requirements, and inability to generalize to unseen classes or tasks. We intend to research **deep generative models** as a way to learn task-agnostic probabilistic models of audio signals and design inference methods to combine and reuse them for a variety of tasks. We will pursue our investigation of hybrid methods that combine the representational power of deep learning with **statistical signal processing** expertise by leveraging recent optimization techniques for non-convex, non-linear inverse problems. We will also explore the integration of deep learning and **symbolic reasoning** to increase the generalization ability of deep models and to empower researchers/engineers to improve them.

3.1.2. Learning from little/no labeled data

While fully labeled data are costly, unlabeled data are cheap but provide intrinsically less information. **Weakly supervised learning** based on not-so-expensive incomplete and/or noisy labels is a promising middle ground. This entails modeling label noise and leveraging it for unbiased training. Models may depend on the labeler, the spoken context (voice command), or the temporal structure (ambient sound analysis). We will also keep studying **transfer learning** to adapt an expressive (audiovisual) speech synthesizer trained on a given speaker to another speaker for which only neutral voice data has been collected.

3.1.3. Preserving privacy

Some voice technology companies process users' voices in the cloud and store them for training purposes, which raises privacy concerns. We aim to **hide speaker identity** and (some) speaker states and traits from the speech signal, and evaluate the resulting automatic speech/speaker recognition accuracy and subjective quality/intelligibility/identifiability, possibly after removing private words from the training data. We will also explore **semi-decentralized learning** methods for model personalization, and seek to obtain statistical guarantees.

3.2. Speech production and perception

This research axis covers topics related to the production of speech through articulatory modeling and multimodal expressive speech synthesis, and topics related to the perception of speech through the categorization of sounds and prosody in native and in non-native speech.

3.2.1. Articulatory modeling

Articulatory speech synthesis will rely on further 2D and 3D modeling of the vocal tract as well as of the **dynamics of the vocal tract** from real-time MRI data. The prediction of glottis opening will also be considered so as to produce better quality acoustic events for consonants. The **coarticulation model** developed to handle the animation of the visible articulators will be extended to control the face and the tongue. This will help characterize links between the vocal tract and the face, and illustrate inner mouth articulation to learners. The suspension of articulatory movements in stuttering speech will also be studied.

3.2.2. Multimodal expressive speech

The dynamic realism of the animation of the talking head, which has a direct impact on audiovisual intelligibility, will continue to be our goal. Both the **animation** of the lower part of the face relating to speech and of the upper part relating to the facial expression will be considered, and development will continue towards a multilingual talking head. We will investigate further the modeling of **expressivity** both for audio-only and for audiovisual speech synthesis. We will also evaluate the benefit of the talking head in various use cases, including children with language and learning disabilities or deaf people.

3.2.3. Categorization of sounds and prosody

Reading and speaking are basic skills that need to be mastered. Further analysis of schooling experience will allow a better understanding of reading acquisition, especially for children with some language impairment. With respect to L1/L2 language interference [0] , a special focus will be set on the impact of L2 prosody on segmental realizations. Prosody will also be considered for its implication on the structuration of speech communication, including on discourse particles. Moreover, we will experiment the usage of speech technologies for computer assisted language learning in middle and high schools, and, hopefully, also for helping children learning to read.

3.3. Speech in its environment

The themes covered by this research axis correspond to the acoustic environment analysis, to speech enhancement and noise robustness, and to linguistic and semantic processing.

3.3.1. Acoustic environment analysis

Audio scene analysis is key to characterize the environment in which spoken communication may take place. We will investigate audio event detection methods that exploit both strongly/weakly labeled and unlabeled data, operate in real-world conditions, can discover novel events, and provide a semantic interpretation. We will keep working on source localization in the presence of nearby acoustic reflectors. We will also pursue our effort at the interface of **room acoustics** to blindly estimate room properties and develop acoustics-aware signal processing methods. Beyond spoken communication, this has many applications to surveillance, robot audition, building acoustics, and augmented reality.

3.3.2. Speech enhancement and noise robustness

We will pursue **speech enhancement** methods targeting several distortions (echo, reverberation, noise, overlapping speech) for both speech and speaker recognition applications, and extend them to ad-hoc arrays made of the microphones available in our daily life using multi-view learning. We will also continue to explore statistical signal models **beyond the usual zero-mean complex Gaussian model** in the time-frequency domain, e.g., deep generative models of the signal phase. **Robust acoustic modeling** will be achieved by learning domain-invariant representations or performing unsupervised domain adaptation on the one hand, and by extending our uncertainty-aware approach to more advanced (e.g., nongaussian) uncertainty models and accounting for the additional uncertainty due to short utterances on the other hand, with application to speaker and language recognition “in the wild”.

3.3.3. Linguistic and semantic processing

We will seek to address robust speech recognition by exploiting word/sentence embeddings carrying **semantic information** and combining them with acoustical uncertainty to rescore the recognizer outputs. We will also combine semantic content analysis with text obfuscation models (similar to the label noise models to be investigated for weakly supervised training of speech recognition) for the task of detecting and classifying (hateful, aggressive, insulting, ironic, neutral, etc.) **hate speech** in social media.

[0] L1 refers to the speaker’s native language, and L2 to a speaker’s second language, usually learned later as a foreign language

NEUROSYS Project-Team

3. Research Program

3.1. Main Objectives

The main challenge in computational neuroscience is the high complexity of neural systems. The brain is a complex system and exhibits a hierarchy of interacting subunits. On a specific hierarchical level, such subunits evolve on a certain temporal and spatial scale. The interactions of small units on a low hierarchical level build up larger units on a higher hierarchical level evolving on a slower time scale and larger spatial scale. By virtue of the different dynamics on each hierarchical level, until today the corresponding mathematical models and data analysis techniques on each level are still distinct. Only few analysis and modeling frameworks are known which link successfully at least two hierarchical levels.
After extracting models for different description levels, they are typically applied to obtain simulated activity which is supposed to reconstruct features in experimental data. Although this approach appears straightforward, it presents various difficulties. Usually the models involve a large set of unknown parameters which determine the dynamical properties of the models. To optimally reconstruct experimental features, it is necessary to formulate an inverse problem to extract optimally such model parameters from the experimental data. Typically this is a rather difficult problem due to the low signal-to-noise ratio in experimental brain signals. Moreover, the identification of signal features to be reconstructed by the model is not obvious in most applications. Consequently an extended analysis of the experimental data is necessary to identify the interesting data features. It is important to combine such a data analysis step with the parameter extraction procedure to achieve optimal results. Such a procedure depends on the properties of the experimental data and hence has to be developed for each application separately. Machine learning approaches that attempt to mimic the brain and its cognitive processes have had a lot of success in classification problems during the last decade. These hierarchical and iterative approaches use non-linear functions, which imitate neural cell responses, to communicate messages between neighboring layers. In our team, we work towards developing polysomnography-specific classifiers that might help in linking the features of particular interest for building systems for sleep signal classification with sleep mechanisms, with the accent on memory consolidation during the Rapid Eye Movement (REM) sleep phase.

3.2. Challenges

Techniques for the implementation and analysis of models achieved promises to be able to construct novel data monitors. This construction involves additional challenges and requires contact with realistic environments. By virtue of the specific applications of the research, close contact to hospitals and medical companies shall be established over a longer term in order to (i) gain deeper insight into the specific application of the devices and (ii) build specific devices in accordance with the actual need. Collaborations with local and national hospitals and the pharmaceutical industry already exist.

3.3. Research Directions

- From the microscopic to the mesoscopic scale:
  One research direction focuses on the *relation of single-neuron activity* on the microscopic scale *to the activity of neuronal populations*. To this end, the team investigates the stochastic dynamics of single neurons subject to external random inputs and involving random microscopic properties, such as random synaptic strengths and probability distributions of spatial locations of membrane ion channels. Such an approach yields a stochastic model of single neurons and allows the derivation of a stochastic neural population model.

  This bridge between the microscopic and mesoscopic scale may be performed via two pathways. The analytical and numerical treatment of the microscopic model may be called a *bottom-up approach*,

since it leads to a population activity model based on microscopic activity. This approach allows theoretical neural population activity to be compared to experimentally obtained population activity. The *top-down approach* aims at extracting signal features from experimental data gained from neural populations which give insight into the dynamics of neural populations and the underlying microscopic activity. The work on both approaches represents a well-balanced investigation of the neural system based on the systems properties.

- From the mesoscopic to the macroscopic scale:
  The other research direction aims to link neural population dynamics to macroscopic activity and behavior or, more generally, to phenomenological features. This link is more indirect but a very powerful approach to understand the brain, e.g., in the context of medical applications. Since real neural systems, such as in mammals, exhibit an interconnected network of neural populations, the team studies analytically and numerically the network dynamics of neural populations to gain deeper insight into possible phenomena, such as traveling waves or enhancement and diminution of certain neural rhythms. Electroencephalography (EEG) is a powerful brain imaging technique to study the overall brain activity in real time non-invasively. However it is necessary to develop robust techniques based on stable features by investigating the time and frequency domains of brain signals. Two types of information are typically used in EEG signals: (i) transient events such as evoked potentials, spindles and K-complexes and (ii) the power in specific frequency bands.

ORPAILLEUR Project-Team

3. Research Program

3.1. Hybrid and Exploratory Knowledge Discovery

Keywords: knowledge discovery in databases, knowledge discovery in databases guided by domain knowledge, data mining, data exploration, formal concept analysis, classification, pattern mining, numerical methods in data mining.

Knowledge discovery in databases (KDD) aims at discovering intelligible and reusable patterns in possibly large databases. These patterns can then be interpreted as knowledge units to be reused in knowledge-based systems. From an operational point of view, the KDD process is based on three main steps: (i) selection and preparation of the data, (ii) data mining, (iii) interpretation of the discovered patterns. Moreover, the KDD process is iterative, interactive, and generally controlled by an expert of the data domain, called the analyst. The analyst selects and interprets a subset of the extracted units for obtaining knowledge units having a certain plausibility. In this view, KDD is an exploratory process similar to "exploratory data analysis".

The KDD process –as implemented in the Orpailleur team– is based on data mining methods which are either symbolic or numerical. Symbolic methods are based on pattern mining (e.g. mining frequent itemsets, association rules, sequences...), Formal Concept Analysis (FCA) and extensions such as Pattern Structures and Relational Concept Analysis (RCA), and redescription mining. Numerical methods are based on Random Forests, Support Vector Machines (SVM), Neural Networks, and probabilistic approaches such as second-order Hidden Markov Models (HMM). Moreover, for being able to deal with complex data, numerical data mining methods can be associated with symbolic methods, for improving applicability and efficiency of knowledge discovery. This is particularly true in classification, where supervised and unsupervised approaches may be combined with benefits.

A main operation in the research work of Orpailleur is "classification", which is a polymorphic process involved in modeling, mining, representing, and reasoning tasks. In this way, domain knowledge, when available, can improve and guide the KDD process, materializing the idea of *Knowledge Discovery guided by Domain Knowledge* or KDDK. In KDDK, domain knowledge plays a role at each step of KDD: the discovered patterns can be interpreted as knowledge units and reused for problem-solving activities in knowledge systems, implementing the exploratory process "mining, interpreting, modeling, representing, and reasoning". Then knowledge discovery can be considered as a key task in knowledge engineering (KE), having an impact in various semantic activities, e.g. information retrieval, recommendation, and ontology engineering. In addition, if knowledge discovery can feed knowledge-based systems, in turn, domain knowledge can be used to support the knowledge discovery process.

Finally, life sciences, i.e. agronomy, biology, chemistry, and medicine, are application domains where the Orpailleur team has a very rich experience. The team intends to keep and to extend this experience, paying also more attention to the impact of knowledge discovery in the real world. This should lead to the design of green (sustainable), explainable, and fair data mining systems.

3.2. Text Mining

Keywords: text mining, knowledge discovery from texts, text classification, annotation, ontology engineering from texts.

The objective of a text mining process is to extract useful knowledge units from large collections of texts [71]. The text mining process shows specific characteristics due to the fact that texts are complex objects written in natural language. The information in a text is expressed in an informal way, following linguistic rules, making text mining a difficult task. A text mining process has to take into account –as much as possible– paraphrases, ambiguities, specialized vocabulary and terminology. This is why the preparation of texts for text mining is usually dependent on linguistic resources and methods.

From a knowledge discovery perspective, text mining aims at extracting "interesting units" (nouns and relations) from texts with the help of domain knowledge encoded within a knowledge base. The process is roughly similar for text annotation. Text mining is especially useful in the context of semantic web for ontology engineering. In the Orpailleur team, we work on the mining of real-world texts in application domains such as biology and medicine, using numerical and symbolic data mining methods. Accordingly, the text mining process may be involved in a loop used to enrich and to extend linguistic resources. In turn, linguistic and ontological resources can be exploited to guide a "knowledge-based text mining process".

3.3. Knowledge Systems and Web of Data

Keywords: knowledge engineering, web of data, semantic web, ontology, description logics, classification-based reasoning, case-based reasoning, information retrieval, recommendation.

The web of data constitutes a good platform for experimenting ideas on knowledge engineering (KE) and knowledge discovery. A software agent may be able to read, understand, and manipulate information on the web, if and only if the knowledge necessary for achieving those tasks is available. This is why domain knowledge and ontologies are of main importance. OWL ("Web Ontology Language" https://www.w3.org/OWL/) is based on description logics (DLs [72]) and is the representation language commonly used for designing ontologies. In OWL, knowledge units are represented by classes having properties and instances. Concepts are organized within a partially ordered set based on a subsumption relation, and the inference services are based on subsumption and classification.

Actually, there are many interconnections between concept lattices in FCA and ontologies, e.g. the partial order underlying an ontology can be supported by a concept lattice. Moreover, a pair of implications within a concept lattice can provide a possible materialization of a concept definition in an ontology. In this way, we study how the web of data, considered as a set of knowledge sources, e.g. DBpedia, Wikipedia, Yago, Freebase, can be mined for guiding the design of a knowledge base, and further, how knowledge discovery techniques can be applied for allowing a better usage of the web of data, e.g. Linked Open Data (LOD) classification and completion.

Then, a part of the research work in Knowledge Engineering is oriented towards knowledge discovery in the web of data, as, with the increased interest in machine processable data, more and more data is now published in RDF (Resource Description Framework) format. Particularly, we are interested in the completeness of the data and their potential to provide concept definitions in terms of necessary and sufficient conditions. We have proposed algorithms based on FCA and Redescription Mining which allow data exploration as well as the discovery of definition (bidirectional implication rules).

PESTO Project-Team

3. Research Program

3.1. Modelling

Before being able to analyse and properly design security protocols, it is essential to have a model with a precise semantics of the protocols themselves, the attacker and its capabilities, as well as the properties a protocol must ensure.

Most current languages for protocol specification are quite basic and do not provide support for global state, loops, or complex data structures such as lists, or Merkle trees. As an example we may cite Hardware Security Modules that rely on a notion of *mutable global state* which does not arise in traditional protocols, see e.g. the discussion by Herzog [53].

Similarly, the properties a protocol should satisfy are generally not precisely defined, and stating the "right" definitions is often a challenging task in itself. In the case of authentication, many protocol attacks were due to the lack of a precise meaning, cf. [52]. While the case of authentication has been widely studied, the recent digitalisation of all kinds of transactions and services, introduces a plethora of new properties, including for instance anonymity in e-voting, untraceability of RFID tokens, verifiability of computations that are out-sourced, as well as sanitisation of data in social networks. We expect that many privacy and anonymity properties may be modelled as particular observational equivalences in process calculi [48], or indistinguishability between cryptographic games [3]; sanitisation of data may also rely on information-theoretic measures.

We also need to take into account that the attacker model changes. While historically the attacker was considered to control the communication network, we may nowadays argue that even (part of) the host executing the software may be compromised through, e.g., malware. This situation motivates the use of secure elements and multi-factor authentication with out-of-band channels. A typical example occurs in e-commerce: to validate an online payment a user needs to enter an additional code sent by the bank via SMS to the user's mobile phone. Such protocols require the possession of a physical device in addition to the knowledge of a password which could have been leaked on an untrusted platform. The fact that data needs to be copied by a human requires these data to be *short*, and hence amenable to brute-force attacks by an attacker or guessing.

3.2. Analysis

3.2.1. Generic proof techniques

Most automated tools for verifying security properties rely on techniques stemming from automated deduction. Often existing techniques do however not apply directly, or do not scale up due to state explosion problems. For instance, the use of Horn clause resolution techniques requires dedicated resolution methods [41] [44]. Another example is unification modulo equational theory, which is a key technique in several tools, e.g. [51]. Security protocols however require to consider particular equational theories that are not naturally studied in classical automated reasoning. Sometimes, even new concepts have been introduced. One example is the finite variant property [46], which is used in several tools, e.g., *Akiss* [44], Maude-NPA [51] and Tamarin [54]. Another example is the notion of asymmetric unification [50] which is a variant of unification used in Maude-NPA to perform important *syntactic* pruning techniques of the search space, even when reasoning modulo an equational theory. For each of these topics we need to design efficient decision procedures for a variety of equational theories.

3.2.2. Dedicated procedures and tools

We design dedicated techniques for automated protocol verification. While existing techniques for security protocol verification are efficient and have reached maturity for verification of confidentiality and authentication properties (or more generally safety properties), our goal is to go beyond these properties and the standard attacker models, verifying the properties and attacker models identified in Section 3.1 . This includes techniques that:

- can analyse *indistinguishability* properties, including for instance anonymity and unlinkability properties, but also properties stated in simulation-based (also known as universally composable) frameworks, which express the security of a protocol as an ideal (correct by design) system;
- take into account protocols that rely on a notion of *mutable global state* which does not arise in traditional protocols, but is essential when verifying tamper-resistant hardware devices, e.g., the RSA PKCS#11 standard, IBM's CCA and the trusted platform module (TPM);
- consider attacker models for protocols relying on *weak secrets* that need to be copied or remembered by a human, such as multi-factor authentication.

These goals are beyond the scope of most current analysis tools and require both theoretical advances in the area of verification, as well as the design of new efficient verification tools.

3.3. Design

Given our experience in formal analysis of security protocols, including both protocol proofs and finding of flaws, it is tempting to use our experience to design protocols with security in mind and security proofs. This part includes both provably secure design techniques, as well as the development of new protocols.

3.3.1. General design techniques

Design techniques include *composition results* that allow one to design protocols in a modular way [47], [45]. Composition results come in many flavours: they may allow one to compose protocols with different objectives, e.g. compose a key exchange protocol with a protocol that requires a shared key or rely on a protocol for secure channel establishment, compose different protocols in parallel that may re-use some key material, or compose different sessions of the same protocol.

Another area where composition is of particular importance is Service Oriented Computing, where an "orchestrator" must combine some available component services, while guaranteeing some security properties. In this context, we work on the automated synthesis of the orchestrator or monitors for enforcing the security goals. These problems require the study of new classes of automata that communicate with structured messages.

3.3.2. New protocol design

We also design new protocols. Application areas that seem of particular importance are:

- External hardware devices such as security APIs that allow for flexible key management, including key revocation, and their integration in security protocols. The security *fiasco* of the PKCS#11 standard [43], [49] witnesses the need for new protocols in this area.
- Election systems that provide strong security guarantees. We have been working (in collaboration with the Caramba team) on a prototype implementation of an e-voting system, Belenios (http://belenios.gforge.inria.fr).
- Mechanisms for publishing personal information (e.g. on social networks) in a controlled way.

RESIST Team

3. Research Program

3.1. Overview

The Resist project aims at designing, implementing and validating novel models, algorithms and tools to **make networked systems elastic and resilient so as to enhance their scalability and security**, assuming users, applications and devices whose volume and heterogeneity will continue to increase.

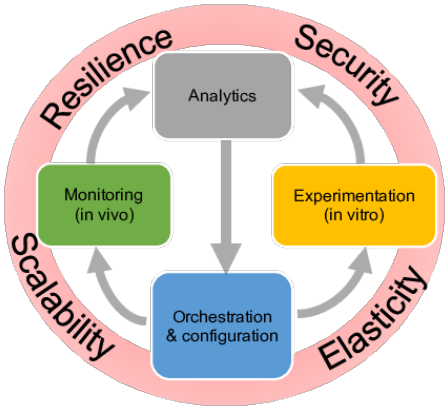


Figure 1. The Resist project

Softwarization of networks and **data analytics** are key enablers to design intelligent methods to orchestrate – *i.e.* configure in a synchronized and distributed manner – both network and system resources. Intelligent **orchestration** leverages relevant data for decision-making using **data analytics**. Input data reflecting the past, current and even future (predicted) states of the system are used to build relevant knowledge. Two approaches are pursued to generate knowledge and to validate orchestration decisions. First, a running system can be **monitored in vivo**. Second, **in vitro experimentation** in a controlled environment (simulators, emulators and experimental platforms) is helpful to reproduce a running system with a high reliability and under different hypotheses. Monitoring and experimentation are steered and configured through orchestration according to the two intertwined loops illustrated in Figure 1 .

Accordingly Resist is thus structured into four main research objectives (activities) namely Monitoring, Experimentation, Analytics and Orchestration.

3.2. Monitoring

The evolving nature of the Internet ecosystem and its continuous growth in size and heterogeneity call for a better understanding of its characteristics, limitations, and dynamics, both locally and globally so as to improve application and protocol design, detect and correct anomalous behaviors, and guarantee performance.

To face these scalability issues, **appropriate monitoring models, methods and algorithms are required for data collection, analysis and sharing** from which knowledge about Internet traffic and usage can be extracted. Measuring and collecting traces necessitate user-centered and data-driven paradigms to cover the wide scope of heterogeneous user activities and perceptions. In this perspective, we propose monitoring algorithms and architectures for large scale environments involving mobile and Internet of Things (IoT) devices.

Resist also assesses **the impact of the Internet infrastructure evolution integrating network softwarization on monitoring**, for example the need for dedicated measurement methodologies. We take into account not only the technological specifics of such paradigms for their monitoring but also the ability to use them for collecting, storing and processing monitoring data in an accurate and cost-effective manner.

Crowd-sourcing and third-party involvement are gaining in popularity, paving the way for massively distributed and collaborative monitoring. We thus investigate opportunistic mobile crowdsensing in order to collect user activity logs along with contextual information (social, demographic, professional) to effectively measure end-users' **Quality of Experience**. However, collaborative monitoring raises serious concerns regarding trust and sensitive data sharing (open data). Data anonymization and sanitization need to be carefully addressed.

3.3. Experimentation

Of paramount importance in our target research context is experimental validation using testbeds, simulators and emulators. In addition to using various existing experimentation methodologies, Resist contributes in **advancing the state of the art in experimentation methods and experimental research practices**, particularly focusing on elasticity and resilience.

We develop and deploy testbeds and emulators for **experimentation with new networking paradigms** such as SDN and NFV, to enable large-scale in-vitro experiments combining all aspects of Software-Defined Infrastructures (server virtualization, SDN/NFV, storage). Such fully controlled environments are particularly suitable for our experiments on resilience, as they ease the management of fault injection features.

We are playing a central role in the development of the Grid'5000 testbed [44] and our objective is to reinforce our collaborations with other testbeds, towards a **testbed federation** in order to enable experiments to scale to multiple testbeds, providing a diverse environment reflecting the Internet itself.

Moreover, our research focuses on extending the infrastructure virtualization capabilities of our Distem [47] emulator, which provides a flexible software-based experimental environment.

Finally, methodological aspects are also important for ensuring **trustworthy and reproducible experiments**, and raises many challenges regarding testbed design, experiment description and orchestration, along with automated or assisted provenance data collection [45].

3.4. Analytics

A large volume of data is processed as part of the operations and management of networked systems. These include traditional monitoring data generated by network components and components' configuration data, but also data generated by dedicated network and system probes.

Understanding and predicting security incidents or system ability to scale requires the elaboration of novel **data analytics techniques** capable to cope with large volumes of data generated from various sources, in various formats, possibly incomplete, non-fully described or even encrypted.

We use machine learning techniques (*e.g.* Topological Data Analysis or multilayer perceptrons) and leverage our domain knowledge to fine-tune them. For instance, machine learning on network data requires the definition of new distance metrics capable to capture the properties of network configurations, packets and flows similarly to edge detection in image processing. Resist contributes to developing and making publicly available an **analytics framework dedicated to networked systems** to support Intelligence-Defined Networked Systems.

Specifically, the goal of the Resist analytics framework is to facilitate the extraction of knowledge useful for **detecting, classifying or predicting security or scalability issues**. The extracted knowledge is then leveraged for orchestration purposes to achieve system elasticity and guarantee its resilience. Indeed, predicting when, where and how issues will occur is very helpful in deciding the provisioning of resources at the right time and place. Resource provisioning can be done either reactively to solve the issues or proactively to prepare the networked system for absorbing the incident (resiliency) in a timely manner thanks to its elasticity.

While the current trend is towards centralization where the collected data is exported to the cloud for processing, we seek to extend this model by also developing and evaluating novel approaches in which **data analytics is seamlessly embedded within the monitored systems**. This combination of big data analytics with network softwarization enablers (SDN, NFV) can enhance the scalability of the monitoring and analytics infrastructure.

3.5. Orchestration

The ongoing transformations in the Internet ecosystem including network softwarization and cloudification bring new management challenges in terms of service and resource orchestration. Indeed, the growing sophistication of Internet applications and the complexity of services deployed to support them require novel models, architectures and algorithms for their automated **configuration** and **provisioning**. Network applications are more and more instantiated through the **composition of services, including virtualized hardware and software resources**, that are offered by **multiple providers** and are subject to changes and updates over time. In this dynamic context, efficient orchestration becomes fundamental for ensuring performance, resilience and security of such applications. We are investigating the chaining of different functions for supporting the security protection of smart devices, based on the networking behavior of their applications.

From a resilience viewpoint, this orchestration at the network level allows the dynamic **reconfiguration of resources** to absorb the effects of congestions, such as link-flooding behaviors. The goal is to drastically reduce the effects of these congestions by imposing dynamic policies on all traffic where the network will adapt itself until it reaches a stable state. We also explore mechanisms for **detecting and remediating potential dysfunctions** within a virtualized network. Corrective operations can be performed through dynamically composed VNFs (Virtualized Network Functions) based on available resources, their dependencies (horizontal and vertical), and target service constraints. We also conduct research on verification methods for automatically assessing and validating the composed chains.

From a security viewpoint, this orchestration provides **prevention mechanisms** that capture adversaries' intentions early and **enforces security policies** in advance through the available resources, to be able to proactively mitigate their attacks. We mainly rely on the results obtained in our research activity on security analytics to build such policies, and the orchestration part focuses on the required algorithms and methods for their automation.

SEMAGRAMME Project-Team

3. Research Program

3.1. Overview

The research program of Sémagramme aims to develop models based on well-established mathematics. We seek two main advantages from this approach. On the one hand, by relying on mature theories, we have at our disposal sets of mathematical tools that we can use to study our models. On the other hand, developing various models on a common mathematical background will make them easier to integrate, and will ease the search for unifying principles.

The main mathematical domains on which we rely are formal language theory, symbolic logic, and type theory.

3.2. Formal Language Theory

Formal language theory studies the purely syntactic and combinatorial aspects of languages, seen as sets of strings (or possibly trees or graphs). Formal language theory has been especially fruitful for the development of parsing algorithms for context-free languages. We use it, in a similar way, to develop parsing algorithms for formalisms that go beyond context-freeness. Language theory also appears to be very useful in formally studying the expressive power and the complexity of the models we develop.

3.3. Symbolic Logic

Symbolic logic (and, more particularly, proof-theory) is concerned with the study of the expressive and deductive power of formal systems. In a rule-based approach to computational linguistics, the use of symbolic logic is ubiquitous. As we previously said, at the level of syntax, several kinds of grammars (generative, categorial...) may be seen as basic deductive systems. At the level of semantics, the meaning of an utterance is captured by computing (intermediate) semantic representations that are expressed as logical forms. Finally, using symbolic logics allows one to formalize notions of inference and entailment that are needed at the level of pragmatics.

3.4. Type Theory and Typed λ-Calculus

Among the various possible logics that may be used, Church's simply typed λ-calculus and simple theory of types (a.k.a. higher-order logic) play a central part. On the one hand, Montague semantics is based on the simply typed λ-calculus, and so is our syntax-semantics interface model. On the other hand, as shown by Gallin, the target logic used by Montague for expressing meanings (i.e., his intensional logic) is essentially a variant of higher-order logic featuring three atomic types (the third atomic type standing for the set of possible worlds).

SPHINX Project-Team

3. Research Program

3.1. Control and stabilization of heterogeneous systems

Fluid-Structure Interaction Systems (FSIS) are present in many physical problems and applications. Their study involves solving several challenging mathematical problems:

- **Nonlinearity:** One has to deal with a system of nonlinear PDE such as the Navier-Stokes or the Euler systems;
- **Coupling:** The corresponding equations couple two systems of different types and the methods associated with each system need to be suitably combined to solve successfully the full problem;
- **Coordinates:** The equations for the structure are classically written with Lagrangian coordinates whereas the equations for the fluid are written with Eulerian coordinates;
- **Free boundary:** The fluid domain is moving and its motion depends on the motion of the structure. The fluid domain is thus an unknown of the problem and one has to solve a free boundary problem.

In order to control such FSIS systems, one has first to analyze the corresponding system of PDE. The oldest works on FSIS go back to the pioneering contributions of Thomson, Tait and Kirchhoff in the 19th century and Lamb in the 20th century, who considered simplified models (potential fluid or Stokes system). The first mathematical studies in the case of a viscous incompressible fluid modeled by the Navier-Stokes system and a rigid body whose dynamics is modeled by Newton's laws appeared much later [119], [114], [94], and almost all mathematical results on such FSIS have been obtained in the last twenty years.

The most studied FSIS is the problem modeling a **rigid body moving in a viscous incompressible fluid** ([77], [73], [112], [83], [88], [116], [118], [102], [86]). Many other FSIS have been studied as well. Let us mention [104], [91], [87], [76], [64], [82], [65], [84] for different fluids. The case of **deformable structures** has also been considered, either for a fluid inside a moving structure (e.g. blood motion in arteries) or for a moving deformable structure immersed in a fluid (e.g. fish locomotion). The obtained coupled FSIS is a complex system and its study raises several difficulties. The main one comes from the fact that we gather two systems of different nature. Some studies have been performed for approximations of this system: [69], [64], [97], [78], [67]). Without approximations, the only known results [74], [75] were obtained with very strong assumptions on the regularity of the initial data. Such assumptions are not satisfactory but seem inherent to this coupling between two systems of different natures. In order to study self-propelled motions of structures in a fluid, like fish locomotion, one can assume that the **deformation of the structure is prescribed and known**, whereas its displacement remains unknown ([110]). This permits to start the mathematical study of a challenging problem: understanding the locomotion mechanism of aquatic animals. This is related to control or stabilization problems for FSIS. Some first results in this direction were obtained in [92], [66], [106].

3.2. Inverse problems for heterogeneous systems

The area of inverse problems covers a large class of theoretical and practical issues which are important in many applications (see for instance the books of Isakov [93] or Kaltenbacher, Neubauer, and Scherzer [95]). Roughly speaking, an inverse problem is a problem where one attempts to recover an unknown property of a given system from its response to an external probing signal. For systems described by evolution PDE, one can be interested in the reconstruction from partial measurements of the state (initial, final or current), the inputs (a source term, for instance) or the parameters of the model (a physical coefficient for example). For stationary or periodic problems (i.e. problems where the time dependence is given), one can be interested in determining from boundary data a local heterogeneity (shape of an obstacle, value of a physical coefficient describing the medium, etc.). Such inverse problems are known to be generally ill-posed and their study leads to investigate the following questions:

- *Uniqueness.* The question here is to know whether the measurements uniquely determine the unknown quantity to be recovered. This theoretical issue is a preliminary step in the study of any inverse problem and can be a hard task.
- *Stability.* When uniqueness is ensured, the question of stability, which is closely related to sensitivity, deserves special attention. Stability estimates provide an upper bound for the parameter error given some uncertainty on data. This issue is closely related to the so-called observability inequality in systems theory.
- *Reconstruction.* Inverse problems being usually ill-posed, one needs to develop specific reconstruction algorithms which are robust to noise, disturbances and discretization. A wide class of methods is based on optimization techniques.

We can split our research in inverse problems into two classes which both appear in FSIS and CWS:

1. **Identification for evolution PDE.**

   Driven by applications, the identification problem for systems of infinite dimension described by evolution PDE has seen in the last three decades a fast and significant growth. The unknown to be recovered can be the (initial/final) state (e.g. state estimation problems [59], [85], [89], [115] for the design of feedback controllers), an input (for instance source inverse problems [56], [68], [79]) or a parameter of the system. These problems are generally ill-posed and many regularization approaches have been developed. Among the different methods used for identification, let us mention optimization techniques ([72]), specific one-dimensional techniques (like in [60]) or observer-based methods as in [100].

   In the last few years, we have developed observers to solve initial data inverse problems for a class of linear systems of infinite dimension. Let us recall that observers, or Luenberger observers [99], have been introduced in automatic control theory to estimate the state of a dynamical system of finite dimension from the knowledge of an output (for more references, see for instance [103] or [117]). Using observers, we have proposed in [105], [90] an iterative algorithm to reconstruct initial data from partial measurements for some evolution equations.We are deepening our activities in this direction by considering more general operators or more general sources and the reconstruction of coefficients for the wave equation. In connection with this problem, we study the stability in the determination of these coefficients. To achieve this, we use geometrical optics, which is a classical albeit powerful tool to obtain quantitative stability estimates on some inverse problems with a geometrical background, see for instance [62], [61].

2. **Geometric inverse problems.**

   We investigate some geometric inverse problems that appear naturally in many applications, like medical imaging and non destructive testing. A typical problem we have in mind is the following: given a domain Ω containing an (unknown) local heterogeneity ω, we consider the boundary value problem of the form

$$\begin{cases} Lu = 0, & (\Omega \setminus \omega) \\ u = f, & (\partial\Omega) \\ Bu = 0, & (\partial\omega) \end{cases}$$

   where L is a given partial differential operator describing the physical phenomenon under consideration (typically a second order differential operator), B the (possibly unknown) operator describing the boundary condition on the boundary of the heterogeneity and f the exterior source used to probe the medium. The question is then to recover the shape of ω and/or the boundary operator B from some measurement Mu on the outer boundary $\partial\Omega$. This setting includes in particular inverse scattering problems in acoustics and electromagnetics (in this case Ω is the whole space and the data are far

field measurements) and the inverse problem of detecting solids moving in a fluid. It also includes, with slight modifications, more general situations of incomplete data (i.e. measurements on part of the outer boundary) or penetrable inhomogeneities. Our approach to tackle this type of problems is based on the derivation of a series expansion of the input-to-output map of the problem (typically the Dirichlet-to-Neumann map of the problem for the Calderón problem) in terms of the size of the obstacle.

3.3. Numerical analysis and simulation of heterogeneous systems

Within the team, we have developed in the last few years numerical codes for the simulation of FSIS and CWS. We plan to continue our efforts in this direction.

- In the case of FSIS, our main objective is to provide computational tools for the scientific community, essentially to solve academic problems.
- In the case of CWS, our main objective is to build tools general enough to handle industrial problems. Our strong collaboration with Christophe Geuzaine's team in Liège (Belgium) makes this objective credible, through the combination of DDM (Domain Decomposition Methods) and parallel computing.

Below, we explain in detail the corresponding scientific program.

- **Simulation of FSIS:** In order to simulate fluid-structure systems, one has to deal with the fact that the fluid domain is moving and that the two systems for the fluid and for the structure are strongly coupled. To overcome this free boundary problem, three main families of methods are usually applied to numerically compute in an efficient way the solutions of the fluid-structure interaction systems. The first method consists in suitably displacing the mesh of the fluid domain in order to follow the displacement and the deformation of the structure. A classical method based on this idea is the A.L.E. (Arbitrary Lagrangian Eulerian) method: with such a procedure, it is possible to keep a good precision at the interface between the fluid and the structure. However, such methods are difficult to apply for large displacements (typically the motion of rigid bodies). The second family of methods consists in using a *fixed mesh* for both the fluid and the structure and to simultaneously compute the velocity field of the fluid with the displacement velocity of the structure. The presence of the structure is taken into account through the numerical scheme. Finally, the third class of methods consists in transforming the set of PDEs governing the flow into a system of integral equations set on the boundary of the immersed structure. The members of SPHINX have already worked on these three families of numerical methods for FSIS systems with rigid bodies (see e.g. [109], [96], [111], [107], [108], [101]).
- **Simulation of CWS:** Solving acoustic or electromagnetic scattering problems can become a tremendously hard task in some specific situations. In the high frequency regime (i.e. for small wavelength), acoustic (Helmholtz's equation) or electromagnetic (Maxwell's equations) scattering problems are known to be difficult to solve while being crucial for industrial applications (e.g. in aeronautics and aerospace engineering). Our particularity is to develop new numerical methods based on the hybridization of standard numerical techniques (like algebraic preconditioners, etc.) with approaches borrowed from asymptotic microlocal analysis. Most particularly, we contribute to building hybrid algebraic/analytical preconditioners and quasi-optimal Domain Decomposition Methods (DDM) [63], [80], [81] for highly indefinite linear systems. Corresponding three-dimensional solvers (like for example GetDDM) will be developed and tested on realistic configurations (e.g. submarines, complete or parts of an aircraft, etc.) provided by industrial partners (Thales, Airbus). Another situation where scattering problems can be hard to solve is the one of dense multiple (acoustic, electromagnetic or elastic) scattering media. Computing waves in such media requires us to take into account not only the interactions between the incident wave and the scatterers, but also the effects of the interactions between the scatterers themselves. When the number of scatterers is very large (and possibly at high frequency [58], [57]), specific deterministic or stochastic numerical methods and algorithms are needed. We introduce new optimized numerical methods for solving such complex

configurations. Many applications are related to this problem *e.g.* for osteoporosis diagnosis where quantitative ultrasound is a recent and promising technique to detect a risk of fracture. Therefore, numerical simulation of wave propagation in multiple scattering elastic media in the high frequency regime is a very useful tool for this purpose.

TONUS Project-Team

3. Research Program

3.1. Kinetic models for plasmas

The fundamental model for plasma physics is the coupled Vlasov-Maxwell kinetic model: the Vlasov equation describes the distribution function of particles (ions and electrons), while the Maxwell equations describe the electromagnetic field. In some applications, it may be necessary to take relativistic particles into account, which leads to consider the relativistic Vlasov equation, even if in general, tokamak plasmas are supposed to be non-relativistic. The distribution function of particles depends on seven variables (three for space, three for the velocity and one for time), which yields a huge amount of computation. To these equations we must add several types of source terms and boundary conditions for representing the walls of the tokamak, the applied electromagnetic field that confines the plasma, fuel injection, collision effects, etc.

Tokamak plasmas possess particular features, which require developing specialized theoretical and numerical tools.

Because the magnetic field is strong, the particle trajectories have a very fast rotation around the magnetic field lines. A full resolution would require a prohibitive amount of computation. It is necessary to develop reduced models for large magnetic fields in order to obtain tractable calculations. The resulting model is called a gyrokinetic model. It allows us to reduce the dimensionality of the problem. Such models are implemented in GYSELA and Selalib.

On the boundary of the plasma, the collisions can no more be neglected. Fluid models, such as MagnetoHydroDynamics (MHD) become again relevant. For the good operation of the tokamak, it is necessary to control MHD instabilities that arise at the plasma boundary. Computing these instabilities requires special implicit numerical discretizations with excellent long time behavior.

In addition to theoretical modelling tools, it is necessary to develop numerical schemes adapted to kinetic, gyrokinetic and fluid models. Three kinds of methods are studied in TONUS: Particle-In-Cell (PIC) methods, semi-Lagrangian and fully Eulerian approaches.

3.1.1. Gyrokinetic models: theory and approximation

In most phenomena where oscillations are present, we can establish a three-model hierarchy: (i) the model parameterized by the oscillation period, (ii) the limit model and (iii) the two-scale model, possibly with its corrector. In a context where one wishes to simulate such a phenomenon where the oscillation period is small and the oscillation amplitude is not small, it is important to have numerical methods based on an approximation of the two-scale model. If the oscillation period varies significantly over the domain of simulation, it is important to have numerical methods that approximate properly and effectively the model parameterized by the oscillation period and the two-scale model. Implementing two-scale numerical methods (for instance by Frénod et al. [27]) is based on a numerical approximation of the Two-Scale model. These are called of order 0. A Two-Scale Numerical Method is called of order 1 if it incorporates information from the corrector and from the equation of which this corrector is a solution. If the oscillation period varies between very small values and values of order 1, it is necessary to have new types of numerical schemes (Two-Scale Asymptotic Preserving Schemes of order 1 or TSAPS) that preserve the asymptotics between the model parameterized by the oscillation period and the Two-Scale model with its corrector. A first work in this direction has been initiated by Crouseilles et al. [26].

3.1.2. Semi-Lagrangian schemes

The Strasbourg team has a long and recognized experience in numerical methods for Vlasov-type equations. We are specialized in both particle and phase space solvers for the Vlasov equation: Particle-in-Cell (PIC) methods and semi-Lagrangian methods. We also have a long-standing collaboration with CEA Cadarache for the development of the GYSELA software for gyrokinetic tokamak plasmas.

The Vlasov and the gyrokinetic models are partial differential equations that express the transport of the distribution function in the phase space. In the original Vlasov case, the phase space is the six-dimension position-velocity space. For the gyrokinetic model, the phase space is five-dimensional because we consider only the parallel velocity in the direction of the magnetic field and the gyrokinetic angular velocity instead of three velocity components.

A few years ago, Eric Sonnendrücker and his collaborators introduced a new family of methods for solving transport equations in the phase space. This family of methods are the semi-Lagrangian methods. The principle of these methods is to solve the equation on a grid of the phase space. The grid points are transported with the flow of the transport equation for a time step and interpolated back periodically onto the initial grid. The method is then a mix of particle Lagrangian methods and Eulerian methods. The characteristics can be solved forward or backward in time leading to the Forward Semi-Lagrangian (FSL) or Backward Semi-Lagrangian (BSL) schemes. Conservative schemes based on this idea can be developed and are called Conservative Semi-Lagrangian (CSL).

GYSELA is a 5D full gyrokinetic code based on a classical backward semi-Lagrangian scheme (BSL) [31] for the simulation of core turbulence that has been developed at CEA Cadarache in collaboration with our team [28].

More recently, we have started to apply the semi-Lagrangian methods to more general kinetic equations. Indeed, most of the conservation laws of physics can be represented by a kinetic model with a small set of velocities and relaxation source terms [4]. Compressible fluids or MHD equations have such representations. Semi-Lagrangian methods then become a very appealing and efficient approach for solving these equations.

3.1.3. PIC methods

Historically PIC methods have been very popular for solving the Vlasov equations. They allow solving the equations in the phase space at a relatively low cost. The main disadvantage of this approach is that, due to its random aspect, it produces an important numerical noise that has to be controlled in some way, for instance by regularizations of the particles, or by divergence correction techniques in the Maxwell solver. We have a long-standing experience in PIC methods and we started implementing them in Selalib. An important aspect is to adapt the method to new multicore computers. See the work by Crestetto and Helluy [25].

3.2. Fluid and reduced kinetic models for plasmas

As already said, kinetic plasmas computer simulations are very intensive, because of the gyrokinetic turbulence. In some situations, it is possible to make assumptions on the shape of the distribution function that simplify the model. We obtain in this way a family of fluid or reduced models.

Assuming that the distribution function has a Maxwellian shape, for instance, we obtain the MagnetoHydroDynamic (MHD) model. It is physically valid only in some parts of the tokamak (at the edges for instance). The fluid model is generally obtained from the hypothesis that the collisions between particles are strong.

But the reduction is not necessarily a consequence of collisional effects. Indeed, even without collisions, the plasma may still relax to an equilibrium state over sufficiently long time scales (Landau damping effect).

In the fluid or reduced-kinetic regions, the approximation of the distribution function could require fewer data while still achieving a good representation, even in the collisionless regime.

Therefore, a fluid or a reduced model is a model where the explicit dependency on the velocity variable is removed. In a more mathematical way, we consider that in some regions of the plasma, it is possible to exhibit a (preferably small) set of parameters α that allows us to describe the main properties of the plasma with a generalized "Maxwellian" M. Then

$$f(x, v, t) = M(\alpha(x, t), v).$$

In this case it is sufficient to solve for $\alpha(x,t)$. Generally, the vector α is the solution of a first order hyperbolic system.

Another way to reduce the model is to try to find an abstract kinetic representation with an as small as possible set of kinetic velocities. The kinetic approach has then only a mathematical meaning. It allows solving very efficiently many equations of physics.

3.2.1. Numerical schemes

As previously indicated, an efficient method for solving the reduced models is the Discontinuous Galerkin (DG) approach. It is possible to make it of arbitrary order. It requires limiters when it is applied to nonlinear PDEs occurring for instance in fluid mechanics. But the reduced models that we intend to write are essentially linear. The nonlinearity is concentrated in a few coupling source terms.

In addition, this method, when written in a special set of variables, called the entropy variables, has nice properties concerning the entropy dissipation of the model. It opens the door to constructing numerical schemes with good conservation properties and no entropy dissipation, as already used for other systems of PDEs [32], [24], [30], [29].

3.2.2. Matrix-free implicit schemes

In tokamaks, the reduced model generally involves many time scales. Among these time scales, many of them, associated to the fastest waves, are not relevant. In order to filter them out, it is necessary to adopt implicit solvers in time. When the reduced model is based on a kinetic interpretation, it is possible to construct implicit schemes that do not impose solving costly linear systems. In addition the resulting solver is stable even at a very high CFL (Courant Friedrichs Lax) number.

3.3. Electromagnetic solvers

Precise resolution of the electromagnetic fields is essential for proper plasma simulation. Thus it is important to use efficient solvers for the Maxwell systems and its asymptotics: Poisson equation and magnetostatics.

The proper coupling of the electromagnetic solver with the Vlasov solver is also crucial for ensuring conservation properties and stability of the simulation.

Finally, plasma physics implies very different time scales. It is thus very important to develop implicit Maxwell solvers and Asymptotic Preserving (AP) schemes in order to obtain good behavior on long time scales.

3.3.1. Coupling

The coupling of the Maxwell equations to the Vlasov solver requires some precautions. The most important one is to control the charge conservation errors, which are related to the divergence conditions on the electric and magnetic fields. We will generally use divergence correction tools for hyperbolic systems presented for instance in [23] (and the references therein).

3.3.2. Implicit solvers

As already pointed out, in a tokamak, the plasma presents several different space and time scales. It is not possible in practice to solve the initial Vlasov-Maxwell model. It is first necessary to establish asymptotic models by letting some parameters (such as the Larmor frequency or the speed of light) tend to infinity. This is the case for the electromagnetic solver and this requires implementing implicit time solvers in order to efficiently capture the stationary state, the solution of the magnetic induction equation or the Poisson equation.

TOSCA Team

3. Research Program

3.1. Research Program

Most often physicists, economists, biologists and engineers need a stochastic model because they cannot describe the physical, economical, biological, etc., experiment under consideration with deterministic systems, either because of its complexity and/or its dimension or because precise measurements are impossible. Therefore, they abandon trying to get the exact description of the state of the system at future times given its initial conditions, and try instead to get a statistical description of the evolution of the system. For example, they desire to compute occurrence probabilities for critical events such as the overstepping of a given thresholds by financial losses or neuronal electrical potentials, or to compute the mean value of the time of occurrence of interesting events such as the fragmentation to a very small size of a large proportion of a given population of particles. By nature such problems lead to complex modelling issues: one has to choose appropriate stochastic models, which require a thorough knowledge of their qualitative properties, and then one has to calibrate them, which requires specific statistical methods to face the lack of data or the inaccuracy of these data. In addition, having chosen a family of models and computed the desired statistics, one has to evaluate the sensitivity of the results to the unavoidable model specifications. The TOSCA team, in collaboration with specialists of the relevant fields, develops theoretical studies of stochastic models, calibration procedures, and sensitivity analysis methods.

In view of the complexity of the experiments, and thus of the stochastic models, one cannot expect to use closed form solutions of simple equations in order to compute the desired statistics. Often one even has no other representation than the probabilistic definition (e.g., this is the case when one is interested in the quantiles of the probability law of the possible losses of financial portfolios). Consequently the practitioners need Monte Carlo methods combined with simulations of stochastic models. As the models cannot be simulated exactly, they also need approximation methods which can be efficiently used on computers. The TOSCA team develops mathematical studies and numerical experiments in order to determine the global accuracy and the global efficiency of such algorithms.

The simulation of stochastic processes is not motivated by stochastic models only. The stochastic differential calculus allows one to represent solutions of certain deterministic partial differential equations in terms of probability distributions of functionals of appropriate stochastic processes. For example, elliptic and parabolic linear equations are related to classical stochastic differential equations (SDEs), whereas nonlinear equations such as the Burgers and the Navier–Stokes equations are related to McKean stochastic differential equations describing the asymptotic behavior of stochastic particle systems. In view of such probabilistic representations one can get numerical approximations by using discretization methods of the stochastic differential systems under consideration. These methods may be more efficient than deterministic methods when the space dimension of the PDE is large or when the viscosity is small. The TOSCA team develops new probabilistic representations in order to propose probabilistic numerical methods for equations such as conservation law equations, kinetic equations, and nonlinear Fokker–Planck equations.

VERIDIS Project-Team

3. Research Program

3.1. Automated and Interactive Theorem Proving

The VeriDis team gathers experts in techniques and tools for automatic deduction and interactive theorem proving, and specialists in methods and formalisms designed for the development of trustworthy concurrent and distributed systems and algorithms. Our common objective is twofold: first, we wish to advance the state of the art in automated and interactive theorem proving, and their combinations. Second, we work on making the resulting technology available for the computer-aided verification of distributed systems and protocols. In particular, our techniques and tools are intended to support sound methods for the development of trustworthy distributed systems that scale to algorithms relevant for practical applications.

VeriDis members from Saarbrücken are developing the SPASS [10] workbench. It currently consists of one of the leading automated theorem provers for first-order logic based on the superposition calculus [56] and a theory solver for linear arithmetic.

In a complementary approach to automated deduction, VeriDis members from Nancy work on techniques for integrating reasoners for specific theories. They develop veriT [1], an SMT [0] solver that combines decision procedures for different fragments of first-order logic. The veriT solver is designed to produce detailed proofs; this makes it particularly suitable as a component of a robust cooperation of deduction tools.

Finally, VeriDis members design effective quantifier elimination methods and decision procedures for algebraic theories, supported by their efficient implementation in the Redlog system [4].

An important objective of this line of work is the integration of theories in automated deduction. Typical theories of interest, including fragments of arithmetic, are difficult or impossible to express in first-order logic. We therefore explore efficient, modular techniques for integrating semantic and syntactic reasoning methods, develop novel combination results and techniques for quantifier instantiation. These problems are addressed from both sides, i.e. by embedding decision procedures into the superposition framework or by allowing an SMT solver to accept axiomatizations for plug-in theories. We also develop specific decision procedures for theories such as non-linear real arithmetic that are important when reasoning about certain classes of (e.g., real-time) systems but that also have interesting applications beyond verification.

We rely on interactive theorem provers for reasoning about specifications at a high level of abstraction when fully automatic verification is not (yet) feasible. An interactive proof platform should help verification engineers lay out the proof structure at a sufficiently high level of abstraction; powerful automatic plug-ins should then discharge the resulting proof steps. Members of VeriDis have ample experience in the specification and subsequent machine-assisted, interactive verification of algorithms. In particular, we participate in a project at the joint Microsoft Research-Inria Centre on the development of methods and tools for the formal proof of TLA$^+$ [66] specifications. Our prover relies on a declarative proof language, and calls upon several automatic backends [3]. Trust in the correctness of the overall proof can be ensured when the backends provide justifications that can be checked by the trusted kernel of a proof assistant. During the development of a proof, most obligations that are passed to the prover actually fail – for example, because necessary information is not present in the context or because the invariant is too weak, and we are interested in explaining failed proof attempts to the user, in particular through the construction of counter-models.

[0] Satisfiability Modulo Theories [58]

3.2. Formal Methods for Developing and Analyzing Algorithms and Systems

Theorem provers are not used in isolation, but they support the application of sound methodologies for modeling and verifying systems. In this respect, members of VeriDis have gained expertise and recognition in making contributions to formal methods for concurrent and distributed algorithms and systems [2], [9], and in applying them to concrete use cases. In particular, the concept of *refinement* [55], [57], [70] in state-based modeling formalisms is central to our approach because it allows us to present a rational (re)construction of system development. An important goal in designing such methods is to establish precise proof obligations, many of which can be discharged by automatic tools. This requires taking into account specific characteristics of certain classes of systems and tailoring the model to concrete computational models. Our research in this area is supported by carrying out case studies for academic and industrial developments. This activity benefits from and influences the development of our proof tools.

In this line of work, we investigate specific development and verification patterns for particular classes of algorithms, in order to reduce the work associated with their verification. We are also interested in applications of formal methods and their associated tools to the development of systems that underlie specific certification requirements in the sense of, e.g., Common Criteria. Finally, we are interested in the adaptation of model checking techniques for verifying actual distributed programs, rather than high-level models.

Today, the formal verification of a new algorithm is typically the subject of a PhD thesis, if it is addressed at all. This situation is not sustainable given the move towards more and more parallelism in mainstream systems: algorithm developers and system designers must be able to productively use verification tools for validating their algorithms and implementations. On a high level, the goal of VeriDis is to make formal verification standard practice for the development of distributed algorithms and systems, just as symbolic model checking has become commonplace in the development of embedded systems and as security analysis for cryptographic protocols is becoming standard practice today. Although the fundamental problems in distributed programming are well-known, they pose new challenges in the context of modern system paradigms, including ad-hoc and overlay networks or peer-to-peer systems, and they must be integrated for concrete applications.