



RESEARCH CENTER
Bordeaux - Sud-Ouest

FIELD

Activity Report 2014

Section Application Domains

Edition: 2015-03-24

ALGORITHMICS, PROGRAMMING, SOFTWARE AND ARCHITECTURE	
1. LFANT Project-Team (section vide)	4
APPLIED MATHEMATICS, COMPUTATION AND SIMULATION	
2. BACCHUS Team (section vide)	5
3. CAGIRE Team	6
4. CQFD Project-Team	8
5. GEOSTAT Project-Team	9
6. MC2 Team	10
7. REALOPT Project-Team	12
DIGITAL HEALTH, BIOLOGY AND EARTH	
8. CARMEN Team	16
9. MAGIQUE-3D Project-Team	17
10. MAGNOME Project-Team	19
11. MNEMOSYNE Project-Team	21
12. SISTM Team	22
NETWORKS, SYSTEMS AND SERVICES, DISTRIBUTED COMPUTING	
13. HIEPACS Project-Team	23
14. PHOENIX Project-Team	26
15. RUNTIME Team	28
PERCEPTION, COGNITION AND INTERACTION	
16. FLOWERS Project-Team	29
17. MANAO Project-Team (section vide)	30
18. POTIOC Project-Team	31

LFANT Project-Team (section vide)

BACCHUS Team (section vide)

CAGIRE Team

4. Application Domains

4.1. Effusion cooling of aeronautical combustion chambers walls

The industrial applications of our project is the cooling of the walls of the combustion chambers encountered in the helicopter engines, and more precisely, we wish to contribute to the improvement of effusion cooling.

Effusion cooling is nowadays very widespread, especially in the aeronautical context. It consists in piercing holes on the wall of the combustion chamber. These holes induce cold jets that enter inside the combustion chamber. The goal of this jet is to form a film of air that will cool the walls of the chamber, see Figure 2 .

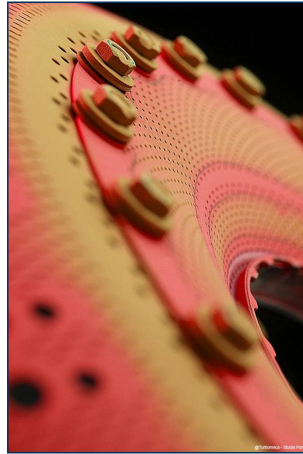


Figure 2. Effusion cooling of aeronautical combustion chambers: close view of a typical perforated chamber wall

Effusion cooling in a combustion chamber takes at the wall where thousands of small holes allow cool air to enter inside the combustion chamber. This induces jets in crossflow in charge of cooling the walls, whatever the heat and the acoustic waves present inside the chamber. Nevertheless, this technique is not straightforward to put in practice: the size, design and position of the holes can have an important effect on the cooling efficiency. For a safe and efficient functioning of the combustion chamber, it is required that the cooling jets and the combustion effects be as much independent as possible. For example, this means that

- The jets of cool air should not mix too much with the internal flow. Otherwise it will decrease the efficiency of the combustion.
- The jets should be as much stable as possible when submitted to waves emitted in the combustion chamber, e.g. acoustic waves induced by combustion instabilities. Otherwise the jets may not cool enough the walls of the combustion chamber which can then undergoes severe damages.

The first point is what we aim at simulate in this project. As the model chosen is the fully compressible Navier Stokes system, there should not be any problem in the future for being able to simulate the effect of an acoustic forcing on the jet in crossflow.

Having a database of Direct Numerical Simulations is also fundamental for testing closure laws that are used in turbulence models encountered in RANS and LES models. With such models, it is possible for example to perform optimisation.

An important aspect that we began to address in this project is the interaction between the flow and the wall. The aim is to understand the effect of coupling between the heat propagation in the wall and the flow near the wall. A careful study of this interaction can allow to determine the exchange coefficients, and so the efficiency of the cooling by the jet. Such determination may be particularly useful to develop one or multidimensional models of wall-fluid interaction [27]. The large eddy simulation performed by Florenciano [26] clearly put into evidence the strong effect of the presence of an acoustic wave in the crossflow on the dynamics of the heat transfer coefficient at the wall.

From the application point of view, compressibility effects must be taken into account since the Mach number of the flow can reach values equal to 0.3, hence/or acoustic waves may be present inside the combustion chamber. This can raise a problem, because upwind numerical schemes are known to be less accurate in the low Mach limit.

CQFD Project-Team

4. Application Domains

4.1. Dependability and safety

Our abilities in probability and statistics apply naturally to industry in particular in studies of dependability and safety.

An illustrative example which gathers all the topics of team is a collaboration started in May 2010 with Thales Optronique on the subject of *optimization of the maintenance of a digital camera equipped with HUMS* (Health Unit Monitoring Systems). This subject is very interesting for us because it combines many aspects of our project. Classification tools will be used to select significant variables as the first step in the modeling of a digital camera. The model will then be analysed and estimated in order to optimize the maintenance.

A second example concerns the optimization of the maintenance date for an aluminum metallic structure subject to corrosion. It is a structure of strategic ballistic missile that is stored in a nuclear submarine missile launcher in peace-time and inspected with a given periodicity. The requirement for security on this structure is very strong. The mechanical stress exerted on the structure depends on its thickness. It is thus crucial to control the evolution of the thickness of the structure over time, and to intervene before the break.

A third example is the minimization of the acoustic signature of a submarine. The submarine has to chose its trajectory in order to minimize at each time step its observability by a surface ship following an unknown random trajectory.

However the spectrum of applications of the topics of the team is larger and may concern many other fields. Indeed non parametric and semi-parametric regression methods can be used in biometry, econometrics or engineering for instance. Gene selection from microarray data and text categorization are two typical application domains of dimension reduction among others. We had for instance the opportunity via the scientific program PRIMEQUAL to work on air quality data and to use dimension reduction techniques as principal component analysis (PCA) or positive matrix factorization (PMF) for pollution sources identification and quantization.

GEOSTAT Project-Team

4. Application Domains

4.1. Application domains

As mentioned above, applicative aspects in GEOSTAT encompass biomedical data (heartbeat signal analysis with IHU LIRYC, biomedical applications in speech signal analysis) and the study of universe science datasets. GEOSTAT's objectives in analysis of biomedical data hinge on the following observations:

- The analysis and detection of cardiac arrhythmia and pathological voice disorders is a paradigm in nonlinear methodologies applied to these types of signals.
- The classical hypothesis under linear approaches are confronted with strong nonlinearities, aperiodicity and chaotic phenomena present in these signals.
- Existing nonlinear approaches are lacking physiological interpretation.

Our objective in this part is to propose new measures based on low-level transition characteristics, these transition phenomena being related to general concepts associated to predictability in complex systems.

MC2 Team

4. Application Domains

4.1. Introduction

We now present our contribution to these above challenges concerning interface problem for complex fluids, direct simulations and analysis, flow control and optimization. From the technical point of view, many productions are common to the different parts of the project. For example, level-set methods, fast-marching procedure are used for shape optimization and for microfluidics, penalization methods are used for high Reynolds flows and for tumor growth. This leads to a strong politic of development of numerical modules.

4.2. Multi-fluid flows

- computation of bifluid flows : see the thesis of S. Tancogne ([90]) and P. Vigneaux ([93]). Stability of an interface, shape of droplets, formation of a jet. Study of the Plateau-Rayleigh instability. Behaviour of diphasic fluids evolving in square microchannels.
- mixing in micro-channel : see the thesis of J. Dambrine [73]. Passive mixing strategies involving boundary conditions. Enhanced oil recovery (study of mixing oil-water-polymer in a microchannel).
- emulsions and foam : see the thesis of S. Benito [56]. Applications in biology : behaviour of tissues, of tumor,...
- polymer nanotube conglomerate wire : it was the subject of a talk in the following conference "WCCM8-ECCOMAS2008" and of the talk [70].

4.3. Cancer modeling

- Specific models : investigation of particular cancers : gliomas (brain tumors), meningioma, colorectal cancers lung and liver metastasis, breast cancer. This is one part of the PhD works of P. Berment, J. Jouganous, G. Lefebvre and post-doc of J. Joie.
- Modelling of electrochemotherapy
- Parameter estimations with the help of low order models : see the PhD of J. Jouganous
- Patient-specific simulations
- Theoretical biology of the metastatic process: dynamics of a population of tumors in mutual interactions, dormancy, pre-metastatic and metastatic niche, quantification of metastatic potential and differential effects of anti-angiogenic therapies on primary tumor and metastases.
- Mathematical models for preclinical cancer research: description and prediction of tumor growth and metastatic development, effect of anti-cancerous therapies

4.4. Newtonian fluid flows simulations and their analysis

- Simulation of a synthetic or pulsed jet. This is an ongoing project with Renault and PSA inside a PREDIT project.
- Vortex dynamics : see [75].
- Simulation of compressible flows on cartesian grids : see the thesis of Gabriele Ottino's Thesis [86], who underwent his doctoral studies in conjunction in the MC2 team and at the Politecnico di Torino, and defended in April 2009. He had a grant of the French-Italian university.
- 3D turbulent flows through DESGRIVRE contract with AIRBUS. Thesis of C. Wervaecke [95]. The goal is to use Detached-Eddy Simulation to model turbulent flows around iced bodies.

- Porous media : Numerical study of coupling between Richards and transport-diffusion equations in permeable sediment affected by tidal oscillation. See the thesis of R. Chassagne [68]
- Modeling and numerical simulation of the flow around a real wind turbine. Phd thesis of Xin Jin. This includes reduced order model to design more efficient blades.

4.5. Flow control and shape optimization

- passive control : the idea is to put a porous interface between the solid body and the fluid. See the D. Depeyras thesis [74] and Yong-Liang Xiang [97] and CH Bruneau and Iraj Mortazavi) [60]. See also project [65] founded by the European Community.
- active control : see the three PhD thesis: M. Buffoni, J. Weller [94], E. Lombardi and FFAST project funded by EU and led by the University of Bristol and AIRBUS UK.
- shape optimization for turbo-machines : See [91].
- reduced order models : it consists in designing a non-linear observer that estimates the state of the flow field from a limited number of measurements in the field. The challenge is to reduce as much as possible the information required and to take it from the boundary. See J. Weller [94] and E. Lombardi.
- passive control of flows with porous media : see [62], [59], [58], [85], [63].
- inverse problems in imagery : see [67].

REALOPT Project-Team

4. Application Domains

4.1. Introduction

Our group has tackled applications in logistics, transportation and routing [72], [71], [67], [69], in production planning [93] and inventory control [67], [69], in network design and traffic routing [49], [58], [65], [96], [47], [59], [79], [86], in cutting and placement problems [74], [75], [90], [91], [92], [94], and in scheduling [5], [80], [45].

4.2. Network Design and Routing Problems

We are actively working on problems arising in network topology design, implementing a survivability condition of the form “at least two paths link each pair of terminals”. We have extended polyhedral approaches to problem variants with bounded length requirements and re-routing restrictions [58]. Associated to network design is the question of traffic routing in the network: one needs to check that the network capacity suffices to carry the demand for traffic. The assignment of traffic also implies the installation of specific hardware at transient or terminal nodes.

To accommodate the increase of traffic in telecommunication networks, today’s optical networks use grooming and wavelength division multiplexing technologies. Packing multiple requests together in the same optical stream requires to convert the signal in the electrical domain at each aggregation of disaggregation of traffic at an origin, a destination or a bifurcation node. Traffic grooming and routing decisions along with wavelength assignments must be optimized to reduce opto-electronic system installation cost. We developed and compared several decomposition approaches [98], [97], [96] to deal with backbone optical network with relatively few nodes (around 20) but thousands of requests for which traditional multi-commodity network flow approaches are completely overwhelmed. We also studied the impact of imposing a restriction on the number of optical hops in any request route [95]. We also developed a branch-and-cut approach to a problem that consists in placing sensors on the links of a network for a minimum cost [65], [66].

We studied several time dependent formulations for the unit demand vehicle routing problem [51], [50] [30]. We gave new bounding flow inequalities for a single commodity flow formulation of the problem. We described their impact by projecting them on some other sets of variables, such as variables issued of the Picard and Queyranne formulation or the natural set of design variables. Some inequalities obtained by projection are facet defining for the polytope associated with the problem. We are now running more numerical experiments in order to validate in practice the efficiency of our theoretical results.

We also worked on the p-median problem, applying the matching theory to develop an efficient algorithm in Y-free graphs and to provide a simple polyhedral characterization of the problem and therefore a simple linear formulation [85] simplifying results from Baiou and Barahona.

We considered the multi-commodity transportation problem. Applications of this problem arise in, for example, rail freight service design, "less than truckload" trucking, where goods should be delivered between different locations in a transportation network using various kinds of vehicles of large capacity. A particularity here is that, to be profitable, transportation of goods should be consolidated. This means that goods are not delivered directly from the origin to the destination, but transferred from one vehicle to another in intermediate locations. We proposed an original Mixed Integer Programming formulation for this problem which is suitable for resolution by a Branch-and-Price algorithm and intelligent primal heuristics based on it.

For the problem of routing freight railcars, we proposed two algorithmes based on the column generation approach. These algorithmes have been testes on a set of real-life instances coming from a Russian freight real transportation company. Our algorithmes have been faster on these instances than the current solution approach being used by the company.

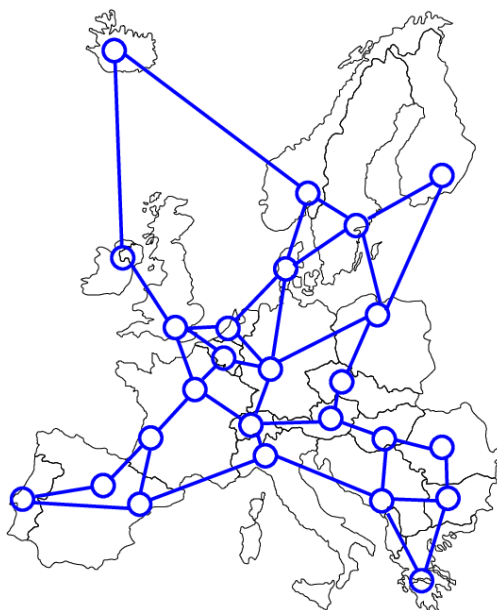


Figure 1. Design of a SDH/SONET european network where demands are multiplexed.

4.3. Packing and Covering Problems

Realopt team has a strong experience on exact methods for cutting and packing problems. These problems occur in logistics (loading trucks), industry (wood or steel cutting), computer science (parallel processor scheduling).

We developed a branch-and-price algorithm for the Bin Packing Problem with Conflicts which improves on other approaches available in the literature [84]. The algorithm uses our methodological advances like the generic branching rule for the branch-and-price and the column based heuristic. One of the ingredients which contributes to the success of our method are fast algorithms we developed for solving the subproblem which is the Knapsack Problem with Conflicts. Two variants of the subproblem have been considered: with interval and arbitrary conflict graphs.

We also developed a branch-and-price algorithm for a variant of the bin-packing problem where the items are fragile. In [43] we studied empirically different branching schemes and different algorithms for solving the subproblems.

We studied a variant of the knapsack problem encountered in inventory routing problem [69]: we faced a multiple-class integer knapsack problem with setups [68] (items are partitioned into classes whose use implies a setup cost and associated capacity consumption). We showed the extent to which classical results for the knapsack problem can be generalized to this variant with setups and we developed a specialized branch-and-bound algorithm.

We studied the orthogonal knapsack problem, with the help of graph theory [62], [60], [64], [63]. Fekete and Schepers proposed to model multi-dimensional orthogonal placement problems by using an efficient representation of all geometrically symmetric solutions by a so called *packing class* involving one *interval graph* for each dimension. Though Fekete & Schepers' framework is very efficient, we have however identified several weaknesses in their algorithms: the most obvious one is that they do not take advantage of the different possibilities to represent interval graphs. We propose to represent these graphs by matrices with consecutive

ones on each row. We proposed a branch-and-bound algorithm for the 2d knapsack problem that uses our 2D packing feasibility check. We are currently developing exact optimization tools for glass-cutting problems in a collaboration with Saint-Gobain. This 2D-3stage-Guillotine cut problems are very hard to solve given the scale of the instance we have to deal with. Moreover one has to issue cutting patterns that avoid the defaults that are present in the glass sheet the are used as raw material. They are extra sequencing constraints regarding the production that make the problem even more complex.

Finally, let us add that we are now organizing a european challenge on packing with society Renault: see <http://challenge-esicup-2015.org/>. This challenge is about loading trucks under practical constraints. The final results will be announced in March 2015.

4.4. Planning, Scheduling, and Logistic Problems

Inventory routing problems combine the optimization of product deliveries (or pickups) with inventory control at customer sites. We considered an industrial application where one must construct the planning of single product pickups over time; each site accumulates stock at a deterministic rate; the stock is emptied on each visit. We have developed a truncated branch-and-price algorithm: periodic plans are generated for vehicles by solving a multiple choice knapsack subproblem; the global planning of customer visits is generated by solving a master program. Confronted with the issue of symmetry in time, we used a state-space relaxation idea. Our algorithm provides solutions with reasonable deviation from optimality for large scale problems (260 customer sites, 60 time periods, 10 vehicles) coming from industry [70]. We previously developed approximate solutions to a related problem combining vehicle routing and planning over a fixed time horizon (solving instances involving up to 6000 pick-ups and deliveries to plan over a twenty day time horizon with specific requirements on the frequency of visits to customers [72].

Together with our partner company GAPSO from the associate team SAMBA, we worked on the equipment routing task scheduling problem [78] arising during port operations. In this problem, a set of tasks needs to be performed using equipments of different types with the objective to maximum the weighted sum of performed tasks.

We participated to the project on an airborne radar scheduling. For this problem, we developed fast heuristics [57] and exact algorithms [45]. A substantial research has been done on machine scheduling problems. A new compact MIP formulation was proposed for a large class of these problems [44]. An exact decomposition algorithm was developed for the NP-hard maximizing the weighted number of late jobs problem on a single machine [80]. A dominant class of schedules for malleable parallel jobs was discovered in the NP-hard problem to minimize the total weighted completion time [82]. We proved that a special case of the scheduling problem at cross docking terminals to minimize the storage cost is polynomially solvable [83], [81].

Another application area in which we have successfully developed MIP approaches is in the area of tactical production and supply chain planning. In [42], we proposed a simple heuristic for challenging multi-echelon problems that makes effective use of a standard MIP solver. [41] contains a detailed investigation of what makes solving the MIP formulations of such problems challenging; it provides a survey of the known methods for strengthening formulations for these applications, and it also pinpoints the specific substructure that seems to cause the bottleneck in solving these models. Finally, the results of [46] provide demonstrably stronger formulations for some problem classes than any previously proposed.

We have been developing robust optimization models and methods to deal with a number of applications like the above in which uncertainty is involved. In [53], [52], we analyzed fundamental MIP models that incorporate uncertainty and we have exploited the structure of the stochastic formulation of the problems in order to derive algorithms and strong formulations for these and related problems. These results appear to be the first of their kind for structured stochastic MIP models. In addition, we have engaged in successful research to apply concepts such as these to health care logistics [48]. We considered train timetabling problems and their re-optimization after a perturbation in the network [55], [54]. The question of formulation is central. Models of the literature are not satisfactory: continuous time formulations have poor quality due to the presence of discrete decision (re-sequencing or re-routing); arc flow in time-space graph blow-up in size (they can only handle a single line timetabling problem). We have developed a discrete time formulation that strikes a

compromise between these two previous models. Based on various time and network aggregation strategies, we develop a 2-stage approach, solving the contiguous time model having fixed the precedence based on a solution to the discrete time model.

Currently, we are conducting investigations on a real-world planning problem in the domain of energy production, in the context of a collaboration with EDF. The problem consists in scheduling maintenance periods of nuclear power plants as well as production levels of both nuclear and conventional power plants in order to meet a power demand, so as to minimize the total production cost. For this application, we used a Dantzig-Wolfe reformulation which allows us to solve realistic instances of the deterministic version of the problem [31]. In practice, the input data comprises a number of uncertain parameters. We deal with a scenario-based stochastic demand with help of a Benders decomposition method. We are working on Multistage Robust Optimization approaches to take into account other uncertain parameters like the duration of each maintenance period, in a dynamic optimization framework. The main challenge addressed in this work is the joint management of different reformulations and solving techniques coming from the deterministic (Dantzig-Wolfe decomposition, due to the large scale nature of the problem), stochastic (Benders decomposition, due to the number of demand scenarios) and robust (reformulations based on duality and/or column and/or row generation due to maintenance extension scenarios) components of the problem [28].

CARMEN Team

4. Application Domains

4.1. Scientific context: the LIRYC

Our fields of application are naturally: electrophysiology and cardiac physiopathology at the tissue scale on one side; medical and clinical cardiology on the other side.

The team's research project is part of the IHU LIRYC project, initiated by Pr. M. Haissaguerre. It is concerned by the major issues of modern electrocardiology: atrial arrhythmias, sudden death due to ventricular fibrillation and heart failure related to ventricular dyssynchrony.

We aim at bringing applied mathematics and scientific computing closer to biomedical research applied to cardiac rhythmology and clinical cardiology. It aims at enhancing our fundamental knowledge of the normal and abnormal cardiac electrical activity, of the patterns of the electrocardiogram; and we will develop new simulation tools for training, biological and clinical applications.

4.2. Basic experimental electrophysiology

Our modeling is carried out in coordination with the experimental teams from the LIRYC. It will help to write new concepts concerning the multiscale organisation of the cardiac action potentials and will serve our understanding in many electrical pathologies:

At the atrial level, we apply our models to understand the mechanisms of complex arrhythmias and the relation with the heterogeneities at the insertion of the pulmonary vein.

At the ventricula level, we focus on (1) modeling the complex coupling between the Purkinje network and the ventricles and (2) modeling the structural heterogeneities at the cellular scale, taking into account the complex organisation and disorganisation of the myocytes and fibroblasts. Point (1) is supposed to play a major role in sudden cardiac death and point (2) is important in the study of infarct scars for instance.

MAGIQUE-3D Project-Team

4. Application Domains

4.1. Seismic Imaging

The main objective of modern seismic processing is to find the best representation of the subsurface that can fit the data recorded during the seismic acquisition survey. In this context, the seismic wave equation is the most appropriate mathematical model. Numerous research programs and related publications have been devoted to this equation. An acoustic representation is suitable if the waves propagate in a fluid. But the subsurface does not contain fluids only and the acoustic representation is not sufficient in the general case. Indeed the acoustic wave equation does not take some waves into account, for instance shear waves, turning waves or the multiples that are generated after several reflections at the interfaces between the different layers of the geological model. It is then necessary to consider a mathematical model that is more complex and resolution techniques that can model such waves. The elastic or viscoelastic wave equations are then reference models, but they are much more difficult to solve, in particular in the 3D case. Hence, we need to develop new high-performance approximation methods.

Reflection seismics is an indirect measurement technique that consists in recording echoes produced by the propagation of a seismic wave in a geological model. This wave is created artificially during seismic acquisition surveys. These echoes (i.e., reflections) are generated by the heterogeneities of the model. For instance, if the seismic wave propagates from a clay layer to sand, one will observe a sharp reflected signal in the seismic data recorded in the field. One then talks about reflection seismics if the wave is reflected at the interface between the two media, or talks about seismic refraction if the wave is transmitted along the interface. The arrival time of the echo enables one to locate the position of this transition, and the amplitude of the echo gives information on some physical parameters of the two geological media that are in contact. The first petroleum exploration surveys were performed at the beginning of the 1920's and for instance, the Orchard Salt Dome in Texas (USA) was discovered in 1924 by the seismic-reflection method.

4.2. Modeling of Multiperforated plates in turboreactors

In the turbo-engine, the temperature can reach 2000 K inside the combustion chamber. To protect its boundary, "fresh" air at 800 K is injected through thousands of perforations. The geometry of the network of perforations is chosen in order to optimize the cooling and the mechanical properties of the chamber. It has been experimentally observed that these perforations have a negative impact on the stability of the combustion. This is due to the interaction with an acoustic wave generated by the combustion. Due to the large number of holes (2000) and their small sizes (0.5 mm) with respect to the size of the combustion chamber (50 cm), it is not conceivable to rely on numerical computations (even with supercomputers) to predict the influence of these perforations.

In collaboration with ONERA, we develop new models which allow to take into account these multiperforated plates at the macroscopic scale.

4.3. Helioseismology

This collaboration with the Max Planck Institute for solar system, which started in 2014, aims at designing efficient numerical methods for the wave propagation problems that arise in helioseismology in the context of inverse problems. The final goal is to retrieve information about the structure of the sun i.e. inner properties such as density or pressure via the inversion of a wave propagation problem. Acoustic waves propagate inside the sun which, in a first approximation and regarding the time scales of physical phenomena, can be considered as a moving fluid medium with constant velocity of motion. Some other simplifications lead to computational saving, such as supposing a radial or axisymmetric geometry of the sun. Aeroacoustic equations must be

adapted and efficiently solved in this context, this has been done in the finite elements code Montjoie 5.2 . In other situations, a full 3D simulation is required and demands large computational resources. Ultimately, we aim at modeling the coupling with gravity potential and electromagnetic waves (MHD equations) in order to be able to recover sun spots.

MAGNOME Project-Team

4. Application Domains

4.1. Function and history of genomes

Yeasts provide an ideal subject matter for the study of eukaryotic microorganisms. From an experimental standpoint, the yeast *Saccharomyces cerevisiae* is a model organism amenable to laboratory use and very widely exploited, resulting in an astonishing array of experimental results. From a genomic standpoint, yeasts from the hemiascomycete class provide a unique tool for studying eukaryotic genome evolution on a large scale. With their relatively small and compact genomes, yeasts offer a unique opportunity to explore eukaryotic genome evolution by comparative analysis of several species. MAGNOME applies its methods for comparative genomics and knowledge engineering to the yeasts through the ten-year old Génolevures program (GDR 2354 CNRS), devoted to large-scale comparisons of yeast genomes with the aim of addressing basic questions of molecular evolution.

We developed the software tools used by the CNRS's <http://www.genolevures.org/> web site. For example, MAGNOME's Magus system for simultaneous genome annotation combines semi-supervised classification and rule-based inference in a collaborative web-based system that explicitly uses comparative genomics to simultaneously analyse groups of related genomes.

4.2. Alternative fuels and bioconversion

Oleaginous yeasts are capable of synthesizing lipids from different substrates other than glucose, and current research is attempting to understand these conversions with the goal of optimizing their throughput, production and quality. From a genomic standpoint the objective is to characterize genes involved in the biosynthesis of precursor molecules which will be transformed into fuels, which are thus not derived from petroleum. MAGNOME's focus is in acquiring genome sequences, predicting genes using models learned from genome comparison and sequencing of cDNA transcripts, and comparative annotation. Our overall goal is to define dynamic models that can be used to predict the behavior of modified strains and thus drive selection and genetic engineering.

4.3. Winemaking and improved strain selection

Yeasts and bacteria are essential for the winemaking process, and selection of strains based both on their efficiency and on the influence on the quality of wine is a subject of significant effort in the Aquitaine region. Unlike the species studied above, yeast and bacterial starters for winemaking cannot be genetically modified. In order to propose improved and more specialized starters, industrial producers use breeding and selection strategies.

Comparative genomics is a powerful tool for strain selection even when genetic engineering must be excluded. Large-scale comparison of the genomes of experimentally characterized strains can be used to identify quantitative trait loci, which can be used as markers in selective breeding strategies. Identifying individual SNPs and predicting their effect can lead to better understanding of the function of genes implicated in improved strain performance, particularly when those genes are naturally mutated or are the result of the transfer of genetic material from other strains. And understanding the combined effect of groups of genes or alleles can lead to insight in the phenomenon of heterosis.

4.4. Knowledge bases for molecular tools

Affinity binders are molecular tools for recognizing protein targets, that play a fundamental role in proteomics and clinical diagnostics. Large catalogs of binders from competing technologies (antibodies, DNA/RNA aptamers, artificial scaffolds, etc.) and Europe has set itself the ambitious goal of establishing a comprehensive, characterized and standardized collection of specific binders directed against all individual human proteins, including variant forms and modifications. Despite the central importance of binders, they presently cover only a very small fraction of the proteome, and even though there are many antibodies against some targets (for example, > 900 antibodies against p53), there are none against the vast majority of proteins. Moreover, widely accepted standards for binder characterization are virtually nonexistent. Alongside the technical challenges in producing a comprehensive binder resource are significant logistical challenges, related to the variety of producers and the lack of reliable quality control mechanisms. As part of the ProteomeBinders and Affinomics projects, MAGNOME works to develop knowledge engineering techniques for storing, exploring, and exchanging experimental data used in affinity binder characterization.

MNEMOSYNE Project-Team

4. Application Domains

4.1. Overview

One of the most original specificities of our team is that it is part of a laboratory in Neuroscience (with a large spectrum of activity from the molecule to the behavior), focused on neurodegenerative diseases and consequently working in tight collaboration with the medical domain. As a consequence, neuroscientists and the medical world are considered as the primary end-users of our researches. Beyond data and signal analysis where our expertise in machine learning may be possibly useful, our interactions are mainly centered on the exploitation of our models. They will be classically regarded as a way to validate biological assumptions and to generate new hypotheses to be investigated in the living. Our macroscopic models and their implementation in autonomous robots will allow an analysis at the behavioral level and will propose a systemic framework, the interpretation of which will meet aetiological analysis in the medical domain and interpretation of intelligent behavior in cognitive neuroscience.

The study of neurodegenerative diseases is targeted because they match the phenomena we model. Particularly, the Parkinson disease results from the death of dopaminergic cells in the basal ganglia, one of the main systems that we are modeling. The Alzheimer disease also results from the loss of neurons, in several cortical and subcortical regions. The variety of these regions, together with large mnemonic and cognitive deficits, require a systemic view of the cerebral architecture and associated functions, very consistent with our approach.

Of course, numerical sciences are also impacted by our researches, at several levels. At a global level, we will propose new control architectures aimed at providing a higher degree of autonomy to robots, as well as machine learning algorithms working in more realistic environment. More specifically, our focus on some cognitive functions in closed loop with a real environment will address currently open problems. This is obviously the case for planning and decision making; this is particularly the case for the domain of affective computing, since motivational characteristics arising from the design of an artificial physiology allow to consider not only cold rational cognition but also hot emotional cognition. The association of both kinds of cognition is undoubtedly an innovative way to create more realistic intelligent systems but also to elaborate more natural interfaces between these systems and human users.

At last, we think that our activities in well-founded distributed computations and high performance computing are not just intended to help us design large scale systems. We also think that we are working here at the core of informatics and, accordingly, that we could transfer some fundamental results in this domain.

SISTM Team

4. Application Domains

4.1. Systems Biology and Translational medicine

Biological and clinical researches have dramatically changed because of the technological advances, leading to the possibility of measuring much more biological quantities than previously. Clinical research studies can include now traditional measurements such as clinical status, but also thousands of cell populations, peptides, gene expressions for a given patient. This has facilitated the transfer of knowledge from basic to clinical science (from "bench side to bedside") and vice versa, a process often called "Translational medicine". However, the analysis of these large amounts of data needs specific methods, especially when one wants to have a global understanding of the information inherent to complex systems through an "integrative analysis". These systems like the immune system are complex because of many interactions within and between many levels (inside cells, between cells, in different tissues, in various species). This has led to a new field called "Systems biology" rapidly adapted to specific topics such as "Systems Immunology" [47], "Systems vaccinology" [44], "Systems medicine" [35]. From the statistician point of view, two main challenges appear: i) to deal with the massive amount of data ii) to find relevant models capturing observed behaviors.

4.2. The case of HIV immunology

The management of HIV infected patients and the control of the epidemics have been revolutionized by the availability of highly active antiretroviral therapies. Patients treated by these combinations of antiretrovirals have most often undetectable viral loads with an immune reconstitution leading to a survival which is nearly the same to uninfected individuals [39]. Hence, it has been demonstrated that early start of antiretroviral treatments may be good for individual patients as well as for the control of the HIV epidemics (by reducing the transmission from infected people) [34]. However, the implementation of such strategy is difficult especially in developing countries. Some HIV infected individuals do not tolerate antiretroviral regimen or did not reconstitute their immune system. Therefore, vaccine and other immune interventions are required. Many vaccine candidates as well as other immune interventions (IL7, IL15) are currently evaluated. The challenges here are multiple because the effects of these interventions on the immune system are not fully understood, there are no good surrogate markers although the number of measured markers has exponentially increased. Hence, HIV clinical epidemiology has also entered in the era of Big Data because of the very deep evaluation at individual level leading to a huge amount of complex data, repeated over time, even in clinical trials that includes a small number of subjects.

HIEPACS Project-Team

4. Application Domains

4.1. Material physics

Participants: Pierre Blanchard, Olivier Coulaud, Arnaud Etcheverry, Matthias Messner.

Due to the increase of available computer power, new applications in nano science and physics appear such as study of properties of new materials (photovoltaic materials, bio- and environmental sensors, ...), failure in materials, nano-indentation. Chemists, physicists now commonly perform simulations in these fields. These computations simulate systems up to billion of atoms in materials, for large time scales up to several nanoseconds. The larger the simulation, the smaller the computational cost of the potential driving the phenomena, resulting in low precision results. So, if we need to increase the precision, there are two ways to decrease the computational cost. In the first approach, we improve algorithms and their parallelization and in the second way, we will consider a multiscale approach.

A domain of interest is the material aging for the nuclear industry. The materials are exposed to complex conditions due to the combination of thermo-mechanical loading, the effects of irradiation and the harsh operating environment. This operating regime makes experimentation extremely difficult and we must rely on multi-physics and multi-scale modeling for our understanding of how these materials behave in service. This fundamental understanding helps not only to ensure the longevity of existing nuclear reactors, but also to guide the development of new materials for 4th generation reactor programs and dedicated fusion reactors. For the study of crystalline materials, an important tool is dislocation dynamics (DD) modeling. This multiscale simulation method predicts the plastic response of a material from the underlying physics of dislocation motion. DD serves as a crucial link between the scale of molecular dynamics and macroscopic methods based on finite elements; it can be used to accurately describe the interactions of a small handful of dislocations, or equally well to investigate the global behavior of a massive collection of interacting defects.

To explore i.e. to simulate these new areas, we need to develop and/or to improve significantly models, schemes and solvers used in the classical codes. In the project, we want to accelerate algorithms arising in those fields. We will focus on the following topics (in particular in the currently under definition **OPTIDIS** project in collaboration with CEA Saclay, CEA Ile-de-france and SIMaP Laboratory in Grenoble) in connection with research described at Sections 3.4 and 3.5 .

- The interaction between dislocations is long ranged ($O(1/r)$) and anisotropic, leading to severe computational challenges for large-scale simulations. In dislocation codes, the computation of interaction forces between dislocations is still the most CPU time consuming and has to be improved to obtain faster and more accurate simulations.
- In such simulations, the number of dislocations grows while the phenomenon occurs and these dislocations are not uniformly distributed in the domain. This means that strategies to dynamically construct a good load balancing are crucial to achieve high performance.
- From a physical and a simulation point of view, it will be interesting to couple a molecular dynamics model (atomistic model) with a dislocation one (mesoscale model). In such three-dimensional coupling, the main difficulties are firstly to find and characterize a dislocation in the atomistic region, secondly to understand how we can transmit with consistency the information between the two micro and meso scales.

4.2. Co-design for scalable numerical algorithms in scientific applications

Participants: Pierre Brenner, Jean-Marie Couteyen, Mathieu Faverge, Luc Giraud, Xavier Lacoste, Guillaume Latu, Salli Moustapha, Pierre Ramet, Fabien Rozar, Jean Roman, Pablo Salas, Xavier Vasseur.

The research activities concerning the ITER challenge are involved in the Inria Project Lab (IPL) **C2S@EXA**.

4.2.1. MHD instabilities edge localized modes

The numerical simulations tools designed for ITER challenges aim at making a significant progress in understanding active control methods of plasma edge MHD instabilities Edge Localized Modes (ELMs) which represent particular danger with respect to heat and particle loads for Plasma Facing Components (PFC) in ITER. Project is focused in particular on the numerical modeling study of such ELM control methods as Resonant Magnetic Perturbations (RMPs) and pellet ELM pacing both foreseen in ITER. The goals of the project are to improve understanding the related physics and propose possible new strategies to improve effectiveness of ELM control techniques. The tool for the nonlinear MHD modeling (code **JOREK**) will be largely developed within the present project to include corresponding new physical models in conjunction with new developments in mathematics and computer science strategy in order to progress in urgently needed solutions for ITER.

The fully implicit time evolution scheme in the **JOREK** code leads to large sparse linear systems that have to be solved at every time step. The MHD model leads to very badly conditioned matrices. In principle the **PaStiX** library can solve these large sparse problems using a direct method. However, for large 3D problems the CPU time for the direct solver becomes too large. Iterative solution methods require a preconditioner adapted to the problem. Many of the commonly used preconditioners have been tested but no satisfactory solution has been found. The research activities presented in Section 3.3 will contribute to design new solution techniques best suited for this context.

4.2.2. Turbulence of plasma particules inside a tokamak

In the context of the ITER challenge, the **GYSELA** project aims at simulating the turbulence of plasma particules inside a tokamak. Thanks to a better comprehension of this phenomenon, it would be possible to design a new kind of source of energy based of nuclear fusion. Currently, **GYSELA** is parallalized in a MPI/OpenMP way and can exploit the power of the current greatest supercomputers (e.g., Juqueen). To simulate faithfully the plasma physic, **GYSELA** handles a huge amount of data. In fact, the memory consumption is a bottleneck on large simulations (449 K cores). In the meantime all the reports on the future Exascale machines expect a decrease of the memory per core. In this context, mastering the memory consumption of the code becomes critical to consolidate its scalability and to enable the implementation of new features to fully benefit from the extreme scale architectures.

In addition to activities for designing advanced generic tools for managing the memory optimisation, further algorithmic research will be conduced to better predict and limit the memory peak in order to reduce the memory footprint of **GYSELA**.

4.2.3. SN Cartesian solver for nuclear core simulation

As part of its activity, EDF R&D is developing a new nuclear core simulation code named COCAGNE that relies on a Simplified PN (SPN) method to compute the neutron flux inside the core for eigenvalue calculations. In order to assess the accuracy of SPN results, a 3D Cartesian model of PWR nuclear cores has been designed and a reference neutron flux inside this core has been computed with a Monte Carlo transport code from Oak Ridge National Lab. This kind of 3D whole core probabilistic evaluation of the flux is computationally very demanding. An efficient deterministic approach is therefore required to reduce the computation effort dedicated to reference simulations.

In this collaboration, we work on the parallelization (for shared and distributed memories) of the DOMINO code, a parallel 3D Cartesian SN solver specialized for PWR core reactivity computations which is fully integrated in the COCAGNE system.

4.2.4. 3D aerodynamics for unsteady problems with moving bodies

ASTRIUM has developed for 20 years the FLUSEPA code which focuses on unsteady phenomenon with changing topology like stage separation or rocket launch. The code is based on a finite volume formulation with temporal adaptive time integration and supports bodies in relative motion. The temporal adaptive integration

classifies cells in several temporal levels, zero being the level with the slowest cells and each level being twice as fast as the previous one. This repartition can evolve during the computation, leading to load-balancing issues in a parallel computation context. Bodies in relative motion are managed through a CHIMERA-like technique which allows building a composite mesh by merging multiple meshes. The meshes with the highest priorities recover the least ones, and at the boundaries of the covered mesh, an intersection is computed. Unlike classical CHIMERA technique, no interpolation is performed, allowing a conservative flow integration. The main objective of this research is to design a scalable version of FLUSEPA in order to run efficiently on modern parallel architectures very large 3D simulations.

PHOENIX Project-Team

4. Application Domains

4.1. Introduction

Building on our previous work, we are studying software development in the context of communication services, in their most general forms. That is, going beyond human-to-human interactions, and covering human-to-machine and machine-to-machine interactions. Software systems revolving around such forms of communications can be found in a number of areas, including telephony, pervasive computing, and assisted living; we view these software systems as coordinating the communication between networked entities, regardless of their nature, human, hardware or software. In this context, our three main application domains are pervasive computing, avionics and cognitive assistance.

4.2. Pervasive Computing

Pervasive computing systems are being deployed in a rapidly increasing number of areas, including building automation and supply chain management. Regardless of their target area, pervasive computing systems have a typical architectural pattern. They aggregate data from a variety of distributed sources, whether sensing devices or software components, analyze a context to make decisions, and carry out decisions by invoking a range of actuators. Because pervasive computing systems are standing at the crossroads of several domains (e.g., distributed systems, multimedia, and embedded systems), they raise a number of challenges in software development:

- **Heterogeneity.** Pervasive computing systems are made of off-the-shelf entities, that is, hardware and software building blocks. These entities run on specific platforms, feature various interaction models, and provide non-standard interfaces. This heterogeneity tends to percolate in the application code, preventing its portability and reusability, and cluttering it with low-level details.
- **Lack of structuring.** Pervasive computing systems coordinate numerous, interrelated components. A lack of global structuring makes the development and evolution of such systems error-prone: component interactions may be invalid or missing.
- **Combination of technologies.** Pervasive computing systems involve a variety of technological issues, including device intricacies, complex APIs of distributed systems technologies and middleware-specific features. Coping with this range of issues results in code bloated with special cases to glue technologies together.
- **Dynamicity.** In a pervasive computing system, devices may either become available as they get deployed, or unavailable due to malfunction or network failure. Dealing with these issues explicitly in the implementation can quickly make the code cumbersome.
- **Testing.** Pervasive computing systems are complicated to test. Doing so requires equipments to be acquired, tested, configured and deployed. Furthermore, some scenarios cannot be tested because of the nature of the situations involved (e.g., fire and smoke). As a result, the programmer must resort to writing specific code to achieve ad hoc testing.

4.3. Smart Cities

The Internet of Things (IoT) has become a reality with the emergence of Smart Cities, populated with large amounts of smart objects which are used to deliver a range of citizen services (e.g., security, well being, etc.) The IoT paradigm relies on the pervasive presence of smart objects or “things”, which raises a number of new challenges in the software engineering domain.

4.3.1. The Object's World project

There are an abundance of research and industry initiatives that have been undertaken with the aim of promoting the emergence of IoT [1]. In line with this goal, the Object's World project brings together stakeholders from different domains to build and support the emergence of an IoT sector in France and beyond. The project is lead by SIGFOX, the world's first cellular network operator dedicated to low-bandwidth wireless objects. The cooperation between industry and research partners (e.g., sensor manufacturers, computer science and electrical engineering research labs) is of uttermost importance in overcoming technological barriers. This issue is currently hindering the development of an IoT sector. The main objectives of this project are the development of:

- expertise in the low-bandwidth network sector,
- low-cost transmitter/receiver chips,
- low-energy autonomous sensors, and
- software frameworks which cover the entire lifecycle of IoT applications.

Network infrastructures which support huge numbers of objects open up a range of opportunities for innovative services. Critically, these new opportunities rely on the ability to address the software engineering challenges of this new sector. We promote an approach that revolves around software frameworks. In areas such as mobile and web development, this approach has already been shown to facilitate software development by abstracting over implementation details and guiding the programmer.

4.4. Assistive Technology for Cognition

Cognitive impairments (memory, attention, time and space orientation, etc) affect a large part of the population, including elderly, patients with brain injuries (traumatic brain injury, stroke, etc), and people suffering from cognitive disabilities, such as Down syndrome.

The emerging industry of assistive technologies provide hardware devices dedicated to specific tasks, such as a telephone set with a keyboard picturing relatives (<http://www.doro.fr>), or a device for audio and video communication over the web (<http://www.technosens.fr>). These assistive technologies apply a traditional approach to personal assistance by providing an equipment dedicated to a single task (or a limited set of tasks), without leveraging surrounding devices. This traditional approach has fundamental limitations that must be overcome to significantly improve assistive technologies:

- they are not adaptable to one's needs. They are generally dedicated to a task and have very limited functionalities: no networking, limited computing capabilities, a limited screen and rudimentary interaction modalities. This lack of functionality may cause a proliferation of devices, complicating the end-user life. Moreover, they are rarely designed to adapt to the cognitive changes of the user. When the requirements evolve, the person must acquire a new device.
- they are often proprietary, limiting innovation. As a result, they cannot cope with the evolution of users' needs.
- they have limited or no interoperability. As a result, they cannot rely on other devices and software services to offer richer applications.

To break this model, we propose to offer an assistive platform that is open-ended in terms of applications and entities. (1) An online catalog of available applications enables every user and caregiver to define personalized assistance in the form of an evolving and adapted set of applications; this catalog provides a community of developers with a mechanism to publish applications for specific daily-activity needs. (2) New types of entities (whether hardware or software) can be added to a platform description to enhance its functionalities and extend the scope of future applications.

RUNTIME Team

4. Application Domains

4.1. Application Domains

The RUNTIME group is working on the design of efficient runtime systems for parallel architectures. We are currently focusing our efforts on High Performance Computing applications that merely implement numerical simulations in the field of Seismology, Weather Forecasting, Energy, Mechanics or Molecular Dynamics. These time-consuming applications need so much computing power that they need to run over parallel machines composed of several thousands of processors.

Because the lifetime of HPC applications often spreads over several years and because they are developed by many people, they have strong portability constraints. Thus, these applications are mostly developed on top of standard APIs (e.g. MPI for communications over distributed machines, OpenMP for shared-memory programming). That explains why we have long standing collaborations with research groups developing parallel language compilers, parallel programming environments, numerical libraries or communication software. Actually, all these “clients” are our primary target.

Although we are currently mainly working on HPC applications, many other fields may benefit from the techniques developed by our group. Since a large part of our efforts is devoted to exploiting multicore machines and GPU accelerators, many desktop applications could be parallelized using our runtime systems (e.g. 3D rendering, etc.).

FLOWERS Project-Team

4. Application Domains

4.1. Applications

Personal robotics. Many indicators show that the arrival of personal robots in homes and everyday life will be a major fact of the 21st century. These robots will range from purely entertainment or educative applications to social companions that many argue will be of crucial help in our aging society. For example, UNECE evaluates that the industry of entertainment, personal and service robotics will grow from 5.4Bn to 17.1Bn over 2008-2010. Yet, to realize this vision, important obstacles need to be overcome: these robots will have to evolve in unpredictable homes and learn new skills while interacting with non-engineer humans after they left factories, which is out of reach of current technology. In this context, the refoundation of intelligent systems that developmental robotics is exploring opens potentially novel horizons to solve these problems.

Human-Robot Collaboration. Robots play a vital role for industry and ensure the efficient and competitive production of a wide range of goods. They replace humans in many tasks which otherwise would be too difficult, too dangerous, or too expensive to perform. However, the new needs and desires of the society call for manufacturing system centered around personalized products and small series productions. Human-robot collaboration could widen the use of robot in this new situations if robots become cheaper, easier to program and safe to interact with. The most relevant systems for such applications would follow an expert worker and works with (some) autonomy, but being always under supervision of the human and acts based on its task models. Video games. In conjunction with entertainment robotics, a new kind of video games are developing in which the player must either take care of a digital creature (e.g. Neopets), or tame it (e.g. Nintendogs), or raise/accompany them (e.g. Sims). The challenges entailed by programming these creatures share many features with programming personal/entertainment robots. Hence, the video game industry is also a natural field of application for FLOWERS.

Environment perception in intelligent vehicles. When working in simulated traffic environments, elements of FLOWERS research can be applied to the autonomous acquisition of increasingly abstract representations of both traffic objects and traffic scenes. In particular, the object classes of vehicles and pedestrians are of interest when considering detection tasks in safety systems, as well as scene categories ("scene context") that have a strong impact on the occurrence of these object classes. As already indicated by several investigations in the field, results from present-day simulation technology can be transferred to the real world with little impact on performance. Therefore, applications of FLOWERS research that is suitably verified by real-world benchmarks has direct applicability in safety-system products for intelligent vehicles.

Automated Tutoring Systems. Optimal teaching and efficient teaching/learning environments can be applied to aid teaching in schools aiming both at increase the achievement levels and the reduce time needed. From a practical perspective, improved models could be saving millions of hours of students' time (and effort) in learning. These models should also predict the achievement levels of students in order to influence teaching practices.

MANAO Project-Team (section vide)

POTIOC Project-Team

4. Application Domains

4.1. Popularization of science, education, art, entertainment

Our project aims at providing 3D digital worlds to all, including the general public, to stimulate understanding, learning, communication and creation. Our scope of applications encompasses

- popularization of science
- education
- art
- entertainment

See "Objective 3: Exploring new applications and usages" for the detailed description [3.4](#) .