LaBRI
Laboratoire Bordelais de Recherche en Informatique
UMR 5800 du CNRS - Université Bordeaux 1
ENSEIRB MATHMECA - Université Bordeaux 2
Laboratoire associé à l’INRIA

PROJECT
2011-2014

351, Cours de la Libération, 33405 Talence - France
direction@labri.fr – http://www.labri.fr
# Contents

## Project of the Lab

1. Self-analysis ................................................. 6
2. Scientific projects and objectives of the laboratory .......... 8
   2.1 To tackle large-scale scientific problems ................. 8
   2.2 LaBRI's project on the Bordeaux campus ................. 10
   2.3 Team structure ......................................... 11
3. Implementation .............................................. 12

## Team COMBALGO

1. Self-analysis ................................................. 17
2. Project and perspectives .................................... 18
   2.1 Distributed algorithms ................................ 18
   2.2 Enumerative and algebraic combinatorics ................. 19
   2.3 Graphs and applications ................................ 20
   2.4 EPI CEPAGE ............................................. 21

## Team IS

1. Self assessment ............................................. 23
   1.1 Strong Points ......................................... 23
   1.2 Weak points ........................................... 23
   1.3 Opportunities ......................................... 23
   1.4 Risks .................................................. 23
2. Project and Prospects ...................................... 23
   2.1 Theme: Analysis and Indexing of Video .................. 24
   2.2 Theme: 3-D Modeling, Visualization, and Interaction ... 25
   2.3 Theme: Modeling Sound and Music ....................... 26
   2.4 Theme: Image Analysis and Structuring ................. 28

## Team LSR

1. Self assessment ............................................. 31
   1.1 Positives ............................................. 31
   1.2 Negatives ............................................. 31
   1.3 Opportunities ......................................... 31
   1.4 Risks .................................................. 31
2. Project and prospects ..................................... 32
   2.1 Comet ................................................ 34
   2.2 Distributed Systems and Objects \rightarrow Muse ........ 35
   2.3 Phoenix .............................................. 36
   2.4 MTSC ................................................. 37
5 Team MABIOVIS
5.1 Self Assessment ....................................................... 40
  5.1.1 MABioVis Strengths ............................................. 40
  5.1.2 Weaknesses ...................................................... 41
  5.1.3 Opportunities .................................................. 42
  5.1.4 Risks ......................................................... 42
5.2 Project and prospects .............................................. 42
  5.2.1 Comparative genomics ........................................ 43
  5.2.2 Scalability is mandatory ...................................... 43
  5.2.3 Algorithmics and analysis of metabolic and metabolomic networks .......... 43
  5.2.4 Dynamic systems and models .................................. 44
  5.2.5 Multiscale tree graphs models for plants and self-similarity ................. 45
  5.2.6 Text algorithmics .............................................. 45
5.3 Evolution of our team ............................................. 45

6 Team MF ..................................................................... 47
6.1 Self Assessment ....................................................... 47
  6.1.1 Forces ............................................................ 47
  6.1.2 Weaknesses ..................................................... 48
  6.1.3 Opportunities .................................................. 48
  6.1.4 Risks ............................................................ 48
6.2 Principal research directions ...................................... 49

7 Team SATANAS .......................................................... 54
7.1 Self Assessment ....................................................... 54
7.2 Project and prospects .............................................. 55

8 Transversal Projects .................................................... 61
8.1 Project and prospects for VISIDIA .................................. 61
  8.1.1 Self assessment ................................................... 61
  8.1.2 Project and prospects .......................................... 61
8.2 Project and prospects for SIMBALS ................................ 63
  8.2.1 Self assessment ................................................... 63
  8.2.2 Project and prospects .......................................... 64
8.3 Project and prospects for RHOBAN : PERSONAL ROBOTICS .................. 66
Project of the Lab

During the last 4-year contract, the scientific and institutional environment of LaBRI has undergone much change. Computer science kept evolving rapidly, both in its internal developments and in its connections with other sciences. Section 1.2.1 below describes the main features of our scientific project in this shifting context, and more details can be found in the projects formulated by the different teams.

LaBRI is already a reference laboratory at the national level – and in some fields at the international level –, enjoying the support of its tutelles (mother institutions). The further development of our visibility requires a continued effort in terms of scientific quality, and an effort to engage our tutelles and to participate in their projects.

Our institutional environment is profoundly different from what it was four years ago. At the national level, for instance, the LRU (the 2007 law modifying the functioning of Universities) has brought numerous changes in the procedures to define faculty positions and to hire new faculty. CNRS has transformed its internal organisation, the way it interacts with laboratories and even its project as far as computer science is concerned. For either tutelle, the situation is probably not yet stabilized and we are still in a period of adjustment. Because these changes are occuring at the level of our tutelles and because the University is the employer of most LaBRI members, they already have, and will continue to have a deep impact on our daily work environment. These transformations have mobilized much energy in the last few years, and we hope that the upcoming years will be more peaceful.

At the local level, LaBRI must learn to work with the Bordeaux Polytechnic Institute (IPB), a new structure within the PRES which houses ENSEIRB. The creation of the Pole Information Technologies, which will associate our laboratory with the Institut de Mathématiques de Bordeaux and the Institut des Matériaux et Systèmes, is another element which we need to take into account. The Optics Institute, centered on Imaging and related problems, will be yet another partner for the laboratory. Finally, INRIA has brought substantial support to certain work groups since its arrival in the Region a few years ago. Yet interaction with INRIA requires a constant effort of harmonization, due to a different type of organisation and to the fact that its action is inserted in a scientific strategy that is defined at the national level.

In this context, LaBRI must anticipate the changes occuring in its environment, to fully participate in the extension of the competences of the University, and to the actions concerning computer science within the framework of the PRES Université de Bordeaux. At the same time, it must maintain its level of excellence and its traditional role of the laboratory of all the Bordeaux campus computer scientists.

LaBRI’s many tutelles all have their own strategic preoccupations, as do all its partner institutions. The regulating role of the laboratory is all the more essential, to preserve the internal dynamics of the community of Bordeaux computer scientists. It is a matter of guaranteeing a fair distribution of ressources, between sub-fields which are differently positioned to be supported by the different partners of the laboratory, and which have different types of needs at different moments. We must for instance bring the necessary support to groups that are engaged in applied research or in technology transfer, while taking care also to maintain sufficient funding for the groups that are further removed from industrial contracts. All this must be done without endangering the attractivity of the laboratory and the excellence of its research in all its domains of activity – and without missing opportunities to develop new themes or to encourage emerging activities. This is the spirit in which LaBRI has built, and will continue to build its scientific policy.
1.1 Self-analysis

Strengths

- Our capacity to tackle computer science problems, from their fundamental to their applied aspects, to create transverse projects and to put forward new research themes.

- The great diversity of the research groups with strong national or international visibility, and the excellent visibility of certain researchers; attractiveness toward French and international researchers.

- Continued support from the University, the IPB and the CNRS, including a substantial flow of researcher positions; the support from Region Aquitaine is precious as well.

- The constitution, already formalized, of a large number of joint projects with INRIA. Similarly, we have shared scientific projects within the Pole Information Technologies, with IMB and IMS, that are already operational; strong insertion in the Pole of competitiveness Aerospace Valley.

- The tradition of unity of the laboratory, across institutional boundaries. The unanimous voice of the laboratory, in its dialogue with the tutelles and the local authorities, is one of our most precious assets.

- Strong national collaborations, which allow us to receive important funds through ANR projects.

- The renovation of the laboratory’s governance since 2007, with the construction of a larger group of vice directors.

- The laboratory is deeply involved in teaching programs, in particular with the Masters programs and the Doctoral School. Our system to select PhD topics and allocate fellowships allows us to play our role within the logics of the different organisations (University, CNRS, INRIA, Region, Ministry, CIFRE, ...).

Weaknesses

- Shortage of administrative and technical staff; especially research engineers to keep the software developments up to date, and administrators for the doctoral school.

  * An effort has been made in the last few years regarding research engineers, but much remains to be done, to better leverage our production. We have insistently queried our tutelles on this count, as well as about administrative staff. We are also exploring the possibility to share contractual engineers, financed by ANR projects, to use these credits more efficiently and to improve the long-term follow-up of our softwares.

- The ratio rank A / rank B (full professors / assistant professors) is too low, and the age pyramid is not favorable in some of the high-visibility themes.

  * This message has been passed on to our tutelles, so that certain MC positions could be transformed into PR positions.

- The number of industrial contracts is too low – a paradoxical effect of the increasing weight of ANR-funded projects.

  * The effort we plan in terms of technology transfer, especially in the hiring of research and development engineers, should make things easier in this domain. Also, the constitution of the IT Pole will increase our visibility in the industry, and the sharing of our contacts with the other labs of the Pole will have an impact on our industrial contracts.

- The number of European contracts is too low.

  * In this domain, one should distinguish the coordination and the participation in European contracts. Participation, which is more easily organised, will continue to be strongly encouraged by the laboratory. One should however remark that participation in a contract, European or otherwise, is not just a resource, it is also an obligation of research in a given field. The multiplication of such contracts with all kinds of agencies finds its limits in the availability of human resources (that is : us).
The setting-up and the coordination of European projects has become an extremely onerous affair. We will have to change our cultural habits and adopt the same attitude as our colleagues in other fields and other countries: to entrust the setting-up of these projects to specialized private agencies. But even in these conditions, a European project, submitted or accepted, represents a heavy load. We hope that support tools will keep being developed, for instance at the level of the PRES.

- It is difficult to attract students to research in theoretical computer science.
  * No research group in the laboratory can any more feed its flow of PhD students exclusively with students trained in Bordeaux. In fundamental informatics, we will continue our effort to attract students from MPRI, but also to attract foreign students. We already started with a few Indian students (from the IITs) and Chinese students. The financing available through ANR and European projects is an asset in this endeavour.

- The publishing strategy of certain teams ought to be more ambitious.
  * The situation and the dynamics of the different teams are very diverse, as are the publishing cultures of the different branches of computer science. One must also take into account the volume of the software or prototype production of each team. Nevertheless, publishing in international journals is a necessary component of any strategy and all groups are aware of this fact. All teams have accomplished major progress in this domain since our last evaluation, and these efforts will be continued.

- Possibly due to the size of the laboratory, a large number of younger colleagues find the governance mechanisms of the laboratory, and beyond of the Universities and the research organisations, rather opaque.
  * We think that this problem is not specific to the Bordeaux campus. We will explore the way in which comparable laboratories familiarize our newer colleagues with the functioning of the institutions in which they work. It is traditional in our milieu to leave the transmission of this type of information, as well as career mentoring, to individual initiatives. It may not be expedient to proceed in a centralized fashion, although such initiatives as CNRS’s Journées des Entrants are praiseworthy. We might want to develop a mentoring system for our new recruits, whatever their status (University, IPB or research organisations).

**Opportunities**

- The Campus operation will endow the Université de Bordeaux with considerable means to improve its international visibility.

- The Information Technologies Pole (the laboratories IMB, IMS and LaBRI group 1000 persons, of which 400 permanent researchers) will boost the activities of several research groups in LaBRI. The creation, in the framework of the IT Pole, of joint laboratories with large industrial groups, will open new windows to industrial problems. The IT Pole will also give us a new framework to participate in new Pôles de compétitivité, currently being created.

- Similarly, our partnership with the INRIA Bordeaux-Sud-Ouest Center, and in 2010 with the Optics Institute, will have a positive impact on the activities of LaBRI.

- The development of international cooperation, notably the more recent projects with China, India and Japan, will broaden our opportunities to collaborate and to recruit.

**Threats**

- The cultural differences between the mathematical, computer science, electronics, engineering and biological communities are not easy to overcome. This can threaten the development of transversal projects between the laboratories of the IT Pole.

- In certain domains (e.g. communications, imaging), under intense international competition, it is essential for the development of industrial partnerships to master all the competences along the information processing chain. This makes the corresponding research groups particularly sensitive to the flow of departures and arrivals of researchers and PhD students.
- The scientific strategies of our tutelles and partners seem, at times, to be antagonistic. This creates an unpleasant atmosphere for the actors of research, which can threaten the unity of the laboratory, in its dealings with the tutelles and the local authorities.
- The development of isolated research themes is a risk for an expanding laboratory.
- Assistant professors with a Habilitation may feel demotivated in a laboratory with a high ratio of external recruiting.
- The long-term evolution of the French system of research and higher education, the multiplication of contractual research operations with a very large number of regional, national and European agencies, and the increasing complexity of the budgetary and administrative regulations, all concur to sharply increase the time spent by the faculty in administrative tasks. In contrast, the French system has not modified its restrictive approach in terms of administrative and technical support staff. This represents a definite disadvantage for French research in the context of international competition, and a serious long-term threat.
- The evolution of the strategy of the ANR is becoming a crucial element of strategy for research laboratories. The ANR sometimes tends to imitate more specialized or regional agencies, and to accept financing only starting or breakthrough projects; but the life of research is not made only of fresh starts: who should finance research programs that are already successful?

1.2 Scientific projects and objectives of the laboratory

The strategy of LaBRI is centered around two fundamental ideas:

1. To tackle full-scale scientific problems from the level of fundamental research to the level of technology transfer;
2. To grow on the Bordeaux campus, by federating the efforts of all its tutelles and partners.

1.2.1 To tackle large-scale scientific problems

During the last few years, a number of organisations and scientific societies have published their vision of research in computer science for the decade to come: the Computing Research Association\(^1\), the British Computer Society\(^2\), the European Union\(^3\), or in France the MEFI\(^4\), the INRIA\(^5\), the CNRS\(^6\). Common orientations emerge from these documents, emphasizing in particular

- the new modes of interaction, be it the interaction between the real and the virtual worlds, or the interaction between persons and systems – extended to the notion of ubiquitous computing, whereby the notion of a user-spectator of the system is replaced by the notion of a user-actor-producer;
- the notion of safety of systems, which includes the safety of software or of critical systems, but also the protection of data and transactions, or the quality of the services rendered by ever more performing networks;
- the notion of individual services in every field, from daily life to medical services, to the assistance of senior or handicapped citizens; all the way to the notion of a safety.net, which will allow a better management of crisis situations;
- the extension of the domain of ever more pervasive, but also more efficient computation, made possible by the evolution of technology, with considerable applications to all fields of science, technology and industry.

---

Finally, all insist on the challenge represented by the conquest of complexity, and on the fecundity and the necessity of the interactions between computer science and other fields of science, whether by the exchange of tools, especially in by-now traditional domains of applications of computer science, or by the exchange of fundamental scientific concepts.

What is at stake in information processing, as depicted above, holds exciting perspectives for a laboratory like ours, and of course considerable challenges. To quote the report recently completed by A. Petit for the direction of CNRS\textsuperscript{7}, \textit{La complexité des problèmes à traiter, en particulier ceux en lien fort avec le monde réel, nécessite une approche globale système faisant appel à plusieurs concepts devant interagir}. This report identifies several sets of problems in which LaBRI possesses first-rate expertise:

- Data storing and mining: encoding, representation, organisation;
- Communication: encoding, compression, transmission, protection, communicating objects;
- Processing: computing grids, distributed systems, parallel algorithms, programming, including the issues of efficiency, verification, proof, safety and security;
- Analysis: data-specific models and algorithms: audio, speech, image, video, biological data, combinatorial structures;
- Interaction: natural language processing, image synthesis, augmented and virtual reality.

The activity of LaBRI will continue to unfold at the best possible level around these problems, and around the underlying foundational issues of informatics (combinatorics, logics, automata and languages). We will continue our efforts to master the applications of our research, and to develop our operations in such a way that they span the complete arc, from foundations to operational realizations and to technology transfer.

During the upcoming 4-year period, we will particularly support federative research actions, that will implicate several groups, within LaBRI or with other laboratories from the IT Pole. Within LaBRI, we already identified a first (unordered) set of transversal problems.

- \textit{Distributed systems and computation} - The importance of this paradigm in informatics is growing. It is of course central in the preoccupations of the Distributed Algorithms theme (CA team), of the VISIDIA project and of the Cepage EPI, which focus on the foundations of distributed computing and on the definition of models that can be instantiated on large-scale platforms. Another approach, combining foundations and applications, concerns the modeling, verification, synthesis and control of distributed system (MV theme, MeF team). Distributed computation is also at the core of the activity of the SATANAS team, in the context of high-performance computing. Finally, the SOD and Phoenix themes (LSR team) work on the simulation of mobile worlds that are intrinsically distributed, and for which they have developed technological platforms and software engineering tools.

- \textit{Networks and communications of the future} - The development of ever more efficient and pervasive networks poses fascinating problems, theoretical as well as applied, concerning routing, testing, communication protocols, quality of service, etc. They are at the heart of the research of the Comet theme (LSR team), which tackles numerous questions such as the optimisation of multimedia services in mobile networks or the convergence of networks and services in tomorrow’s domestic networks. Networks are also studied in the GA and LGL themes (CA and MeF team respectively), around the problems of routing and labeling.

- \textit{Information processing} - This theme is so very transversal that it could be used as a definition of informatics! What we have in mind here is more specific, and concerns the processing chain of the specific kinds of data provided by imaging, video or sound (IS team), natural language (IL theme, MeF team) or biology (MABioVis team), as well as the challenges of visualisation (MABioVis) and high-performance computing (SATANAS team). Common problems run across these different modalities, in particular the handling of very large quantities of data, or the articulation between form and meaning, syntax and semantics. These questions are essential to understand the data and what they represent of the world around us, but also to make this understanding operational, in the form of intelligent systems.

- \textit{Foundations of informatics} - Logics, combinatorics, all the tools which allow the finite des-

\textsuperscript{7} A. Petit, Rapport ST2I, 2009.
cription and the efficient manipulation of infinite objects, are the building blocks of the representation and the processing of data. They are studied in the CA and MeF teams with varying angles of attack, depending on the themes: combinatorial, algorithmic, algebraic or logical. This research uses notions that have become classical (automata, languages, graphs, generating series and many others) or notions that have emerged more recently in that field (such as percolation or game theory). The boundaries between these two teams, which often share a common language, are already porous, but there is a vast potential for more transversal actions.

1.2.2 LaBRI’s project on the Bordeaux campus

This dimension of LaBRI’s project relies on two essential principles: we want to actively participate in the new partnerships that are being constructed on the Bordeaux campus, and to benefit from the dynamics they will generate; and we want to continue to be the laboratory, the common house of all the computer scientists of the campus.

Information Technologies Pole The constitution of the Information Technologies Pole, by Université Bordeaux-1, around IMS, IMB and LaBRI is both an opportunity and a challenge for our laboratory. An opportunity because it encourages the development of collaborative projects between laboratories that are already close; and a challenge because the Pole will lead us to share more resources with communities that have different cultures, are organised differently, and evaluate research according to different keys. Let us however observe that the boundaries between these communities are not watertight: a number of our teams have already developed joint projects with IMS (around imaging and signal, but also around safety) and with IMB (especially through projects involving INRIA, in high-performance computing and in combinatorial optimization).

LaBRI is determinedly involved in the development of this Pole, which will provide the framework for the upcoming joint laboratory with Thales. We have also started thinking about sharing our transfer activities. Cooperative actions within the IT Pole will be essential building blocks of LaBRI’s position. Three research directions are currently identified by the laboratories of the Pole:
- cryptography, safety, security,
- simulation,
- signal, image, video.

Soon the Optics Institute Before the next 4-year contract starts, the University will host the Optics Institute, with which LaBRI will actively collaborate in all fields concerning imaging.

INRIA INRIA has been present on the campus since 2002, through joint projects with LaBRI, and it has had an autonomous and fully operational research center since 2008. The institutional differences between INRIA and the tutelles of LaBRI make our collaboration with INRIA a different development tool than that offered by the IT Pole. One can already measure the very positive impact of INRIA on the development of certain research themes within LaBRI, on their visibility and on the resources at their disposal. The laboratory, which actively sought the implantation of INRIA in Bordeaux, looks forward to the further development of this collaboration, in a shared and balanced dynamic.

One should not disguise, however, that the different modes of functioning of the partners in presence, the University, the IPB, CNRS and INRIA, sometimes pose concrete problems, in terms of circulation of information, of financial transparency and in fine, of allocation of resources. Major progress has already been accomplished, thanks to the conventions that have been signed between INRIA and the University but also – and that is crucial – thanks to an improvement of the dialogue between INRIA and the direction of LaBRI. This improvement is the result of a long and deep effort, which cannot be considered as completed. Our objective will be to reach a balanced, peaceful and harmonious mode of functioning, which will conciliate the laboratory’s vocation to
brings together all the computer scientists of the Bordeaux campus, and the necessary respect of different organizing principles, whose definition escapes local control, and which must be allowed to co-exist.

Indeed, we are convinced that the unity of the laboratory, the fact that it gathers all the Bordeaux computer scientists, is an essential element of its scientific quality, of its capacity to bring different groups to collaborate and to allow new ones to emerge, of its visibility on the campus and of the credibility it enjoys vis-à-vis the tutelles and the regional authorities.

We are aware that this objective, however simple to state, is very ambitious. Its success will depend on the fashion in which the parties construct an open dialogue, with the support of their respective tutelles, and on the degree of awareness of collective issues that is shared by the actors of research involved in existing or planned joint projects.

**SCRIME and CBiB**  LaBRI relies on its partnership with the SCRIME (Studio of Creation and Research in Informatics and Electroacoustic Music) for its activities related to the modeling of sound and music, and with the CBiB (Center of Bioinformatics of Bordeaux) for its research in bio-informatics. It will continue to actively support both these entities.

The bio-informatics theme, which interacts with a large number of partners (Bordeaux-1, Bordeaux-2, CBiB, INRIA and CNRS, sections 07 and 22 of the National Committee), provides a good example of a situation where the regulating role of LaBRI is crucial: the definition of a pluri-disciplinary scientific policy remains a difficult task (arbitrages leading to the earmarking of positions, financing of large-scale actions such as Génolevures, etc.) and the laboratory cannot limit its ambitions to adjusting to the politics defined by its partners independently of one another.

### 1.2.3 Team structure

LaBRI is attached to a structure in a small number of teams, that must therefore be rather large, themselves structured in themes. The heads of the teams form the backbone of the Scientific Council of LaBRI in its reduced perimeter; they are joined by the heads of themes for plenary meetings. The heads are therefore the primary partners of the direction of the laboratory for all issues of scientific choices and allocation of resources. At the same time, the Scientific Council strives to function with a global view of the laboratory and its development (including the emergence of new themes), and not as if it was a simple federation of teams.

We will not modify the number or the perimeter of the teams during the next 4-year period. They were restructured (passing from 5 to 6 teams and deeply transforming their boundaries) in January 2007 and thirty months do not provide sufficient distance to measure and capitalize change. We can however state that the integration work within the more recently created teams (LSR, MABioVis et MeF) has taken a very good start. We recall that LaBRI is constituted in 6 teams, all of which develop joint projects with INRIA:

- **Combinatorics and Algorithms (CA)**
- **Image and Sound (IS)**
- **Languages, Systems and Networks (LSR)**
- **Formal Methods (MeF)**
- **Models and algorithms for Bioinformatics and the Visualisation of information (MABioVis)**
- **Supports and Algorithms for Hi-Performance Numerical Applications (SATANAS)**

This structure allows numerous collaborations between the teams, and transversal projects regularly emerge. A dense schedule of working groups feeds these collaborations: there are a dozen of them, plus a general colloquium and a PhD students’ seminar. The dynamical aspect of our team structure also allows the creation of new research themes *ex nihilo*, on the basis of external hires: it happened that way for Video at the end of the 90s or *Computational Linguistics* during the current decade. Such dynamics also require constant adjustments within the teams.
1.3 Implementation

**Human resources – teacher/researchers and researchers** In order to meet our objectives of visibility and scientific excellence, we count on the continuation of the effort made by our tutelles (the Universities – including Bordeaux-2, IPB and CNRS) and by INRIA in terms of hiring of researchers and teacher/researchers.

In view of the evolution of the rank A / rank B ratio in the laboratory, we will ask our tutelles for a marked effort to recruit more often at the Professor or Research Director level, so as to guarantee better guidance for the young researchers and doctoral students.

The diversity of the research themes developed in the laboratory is an asset in our interaction with the Universities and the IPB. The majority of teaching hours are indeed directed to students who are beginners in computer science: in Licence or at the IUT, but also during the first semesters at ENSEIRB, where most students come straight out of preparatory classes. It is as important for this public to be exposed early to different points of view on contemporary informatics, as to teach them advanced techniques.

The selection of researchers by CNRS or INRIA largely escapes the control of the laboratory, and we can only encourage the best candidates to apply to join us, and continue our efforts to offer them a supportive working environment.

**Human resources – administrative and technical staff** The human resources allocated to the administration of the laboratory are at a worrisome level. In the last 4 years, LaBRI received

- an ADT, specialized in budget management, on an open competitive examination of CNRS (the position was won by a person already on a temporary contract with us)
- a lab manager, allocated to us by Université Bordeaux 1
- un ADT specialized in budget management, as the restitution of a position that was cancelled five years ago after an internal re-allocation.

In the meantime, LaBRI has grown by 27% (between 2006 and 2009). LaBRI already functions with too few administrative staff members for approximately 280 teacher/researchers, researchers and PhD students: one laboratory manager (Bordeaux-1), 1 secretary of direction (CNRS), 3 persons managing budgets (2 CNRS and 1 Bordeaux-1) and 1 person in charge of the organisation of conferences and other events (contractual, IPB). The laboratory currently survives thanks to 3.5 CAE, financed on our own budget.

Our tutelles must urgently decide to reinforce our administrative staff in a permanent fashion. We need at least three ADT-level agents to reinforce:

- the secretariat,
- the financial pole,
- the reception.

As for the technical staff, we urgently need to hire a development engineer, to help the LaBRI Systems and Networks team administer the technological platforms, whose extension is crucial for the development of the laboratory and its transfer projects.

**Involvement in teaching curricula, hiring of PhD students** The laboratory participates in all the computer science programs on the campus.

We are looking for a stronger involvement of the research teams with the masters students, and the master’s accreditation documents plan for a development and/or research project in the third semester. A project, chosen by a group of students, will be guided by members of a research team and the students will be integrated in the team, e.g. they will participate in the working groups. We want to immerse all of the students on the campus in the laboratory, to help them form an opinion of what research can be like.

In a similar vein, IPB will delimit some space dedicated to technology transfer projects. Every year, two LaBRI teams will be called to organise this space, and to teach a class to the ENSEIRB students. In that fashion, ENSEIRB students will have elbowed during their three years at the school, all of the teams of LaBRI.
The doctoral program performs well and its current functioning suits our objectives, notably in terms of the earmarking of high-priority research projects and the hiring and mentoring of students. We do not anticipate deep changes during the next four years.

However, our experience with the hiring of doctoral students for BDI-CNRS fellowships shows how useful it is to audition candidates, to complete the informations given in their records. The Doctoral School considers generalising this procedure to the awarding of Ministry fellowships.

In addition, and together with the Doctoral School, we plan to experiment with a mid-way defence, to supplement the factual information provided in the mid-way report. It would allow all doctoral students to present their research in front of a small jury, whose constructive remarks would be beneficial for the sequel of their work.

Finally, we will continue to encourage students’ initiatives, including the PhD students’ seminar, and we will keep an open ear to the requests they express through their association, AFODIB.

Governance of the research unit The new organisation of the direction, with an enlarged team of vice-directors, made it possible for the laboratory to better manage its activity, and to be better represented in more circles. The direction has relied on the work of two councils, the Laboratory Council and the Scientific Council, to handle the more important questions (requests for and allocation of resources, of doctoral fellowships, etc) in a collegial and transparent fashion.

However, we still need to work on the articulation between these councils. During the next four years, it will be necessary to ensure mutual representations within each council, so that their decisions become more understandable for all our colleagues.

In our governance model, the role of heads of teams is crucial, in the framework of the Scientific Council but also ahead of the meetings of that Council. This point must be further debated within the laboratory, so we can improve our collective functioning at the level of each team.

Financial resources, investments and buildings The volume of resources of LaBRI and the balance between recurring grant and contractual resources are essentially satisfactory. We will work to preserve this balance – it would be a negative evolution if recurring grants would decrease, or if ANR would concentrate more of its funding on narrowly defined programs.

Concerning our investment in heavy equipment, we must already think of renewing certain machines:
- *Génolevures* Platform (jointly financed with the GdR of the same name);
- Grid (jointly financed with CNRS and INRIA, who supervise this project);
- Massive data storage for video and imaging (joint financing under negotiation, including via industrial contracts).

Extensions of the technological platforms for our activities concerning sound, image, video and networks have become necessary, and we plan heavy infrastructure investments (as part of Plan Campus) to house them.

Finally, our building will be full as early as the fall of 2009, and we will convert two meeting rooms into offices. As early as 2011, we will run out of space to accommodate students in their last semester of Masters, precisely at the time when we want to integrate them into the life of the laboratory. The Masters interns are often put up by the teams, in the hope that some will continue in the PhD program, and it is important to be able to give them some working space within our walls.

Strategy for the allocation of resources We approach this question with a notion of the laboratory as an interdependent and federative collectivity, emphasizing its role in optimizing resources and anticipating needs.

In this spirit, we have begun to work towards an optimization of financial flows, by maximizing the circulation of information between project heads, and by grouping purchases and expenditures. An example of this approach can be found in the hiring of research engineers in the context of ANR projects: we have observed in this domain that it is sometimes difficult to hire engineers, which can endanger the development of the project and lead to the loss of engineer-months. We
are investigating the possibility to pool this part of the ANR contracts and to hire permanently a number of engineers, who will be allocated successively to different projects.

Allocating resources includes allocating faculty positions and doctoral fellowships. In the same state of mind, we have instituted an efficient discussion and arbitrage mechanism in the Scientific Council and the Laboratory Council, to produce a ranked list of fellowships or faculty profile applications, which is then distributed as best as possible between the University and IPB (for faculty positions) and the financing agencies (for fellowships) to fit their specific requirements.

**Developments and technology transfer**  LaBRI wants to continue and expand its development and technology transfer activity. Four main directions are considered.

1. **To maintain a substantial activity supporting SMEs and micro-entreprises in Aquitaine.** Even though the uncertainty on the life expectancy of these companies represents a liability, this activity is important for LaBRI, who is perceived by regional authorities as an important support for regional economic development.

2. **To increase our activity within the Pôles de compétitivité.** The laboratory will keep a very high degree of involvement in the Aerospace Valley Pole, notably through the joint laboratory with Thales, and in the Route des Lasers Pole. We will increasingly work with the Logiciel libre (free software) and Avenia (énergie) Poles, currently under construction.

3. **To increase technology transfer towards large corporations.** The laboratory's activity report shows that most of the collaborations with large corporations take the form of CIFRE doctoral contracts: these are rather narrow-scoped collaborations, which are not part of a long-term strategy. The advent of the IT Pole will allow LaBRI to benefit of the already well-established contacts between IMS and large corporations. The joint laboratory with Thales seems to be an interesting model of collaboration, which might be adopted with and adapted to other partners. One of our objectives in collaborating with large corporations, is to be more immediately identified in the set-up phase of projects within Pôles de compétitivité.

4. **To support and leverage software development.** In the last two years, LaBRI has devoted part of its Carnot funding to develop its software prototypes, in order to facilitate their transfer. In the last year, we have conducted an in-depth reflection on the patenting of software. Currently, 4 software systems have been declared, and preliminary studies are on-going for 5 more. These software developments are an important element of the scientific production of LaBRI and the next 4-year period will witness a more determined action to leverage and utilize these softwares (presentation and demonstration to industrial contacts, systematic declaration of software systems in v.1, etc).

5. **The valorisation office, LaBRI-Transfert** was created in 1998. A number of events have occurred since that date: the creation by the PRES of Aquitaine Valo, whose main mission is technology transfer; the arrival of INRIA and its technology transfer office on the campus; the constitution of the IT Pole; the integration of LaBRI within MIB (the Carnot branch of Université Bordeaux-1),... All contribute to a deep modification of the context of technology transfer for LaBRI. The mission of our office will continue to be guided by the directions outlined above, but we will have to reconsider its organisation and institutional set-up to take into account these new facts.

**Staff training policy**  We want to devote a more important share of our training effort, whether formalized in external training sessions or exercised internally, to career development for the technical and administrative staff of the laboratory. This requires a more stable staff composition – as each new arrival, whether contractual or CAE (short-term contractual), must be trained to become operational, at a high time cost for their colleagues already in the laboratory – and a strengthening of the permanent staff. Indeed, any training session requires an investment, in time on the part of the staff directly involved in the training, and in organisation on the part of the services in which they work. This effort is hardly possible in the posture of scarcity management in which LaBRI's administrative and technical teams find themselves.
Dissemination of scientific and technical information  The laboratory will continue to participate enthusiastically in Sciences en Fête, as organised on the campus for a highschool public, in MathEnJeans and other open door type events. It will support as much as possible its members’ activities in terms of dissemination and popularization, be it the publication of books destined to the larger public (e.g. A. Zvonkin’s recent book) or interventions in the media.

Hygiene and safety  The laboratory will pursue the effort it undertook around fire and first-aid training, as these fill its main needs in this domain.
CNRS – UNIVERSITÉ BORDEAUX 1 – UNIVERSITÉ BORDEAUX 2 – INSTITUT POLYTECHNIQUE DE BORDEAUX

Pascal WEIL, directeur
Directeurs adjoints
Maylis DELEST, Eric SOPENA, X, Y

Gestionnaire d’unité
Magali HINNENBERGER, IGE Bx1

Gestion financière – Contrats
Philippe BIAIS, A1 CNRS
Non connu à ce jour, A1 Bx1

Assistante de direction
Cathy ROUBINEAU, TCH (CS) CNRS

Secrétariat de Direction

Assistante d’équipes de recherche
Brigitte CUBEVILLE, TECH, CDD ENSEIRB – Budget Propre
Lena MEZANI, TECH, CDD ENSEIRB – Budget Propre

Accueil
Magali LABROUSSE, CDD, Bx1 – Budget Propre

Appui à la Recherche
IR affectés sur projet aux équipes de recherche
Nathalie FURMENTO, IR2 CNRS
Patrick HENRY, IR1 CNRS
Bois MANSENCAL, IR2 ENSEIRB
Tiphaine MARTIN, IR CNRS
Patrick MARY, IR2 Bx1
Génaud Point, IR2 CNRS

Recherche
Équipes de Recherche

Cellule Labri – Transfert
Valorisation et transfert
Véronique BOGATI, IGE, CDI ADERA
Budget propre

Équipe Technique « Systèmes – Réseaux »
Pascal UNG, IR2 CNRS
Responsable de l’équipe

Combinatoire et Algorithmique
Olivier BEAUMONT

Image et Son
Henti NICOLAS

Langages, Systèmes et Réseaux
Serge CHAUMETTE

Méthodes Formelles
Igor WALUKIEWICZ

Mathématiques et algorithmes pour la Bioinformatique et la Visualisation d’informations
David SHERMAN

Supports et Algorithmes pour les Applications Numériques hautes performances
Raymond NAMYST

Maintenance du bâtiment
Bernard DUFLO, TCH (CN)

Chargés de mission
Non connu à ce jour
Communication
Non connu à ce jour
Informatique
Non connu à ce jour
Labri Hebdo
Non connu à ce jour
Formation doctorale
Non connu à ce jour
Web
Non connu à ce jour
2.1 Self-analysis

Strong points
• **Attractiveness and Spin-off** : During the evaluation period, the team has recruited 7 CNRS and INRIA “chargés de recherche”. Moreover 7 of our doctoral graduates have been recruited by the CNRS (1 locally) and 7 others have obtained academic posts (enseignant chercheur).
• **Visibility** : The team enjoys an excellent international visibility : international collaboration, local organisation of major conferences in the field (DISC, EuroComb, ICALP), participation, in total, in 57 programme and editorial committees.
• **Habilitations** : 2 habilitations have been defended in 2008 and 4 others are expected in the coming months.
• **Theme evolutions** : Moreover, each team theme has a sufficient critical mass to bring out new research themes (Analysis of probabilistic algorithms, algorithmics of large distributed networks, self-stabilising distributed algorithms, combinatorial optimisation, algebraic combinatorics) without sacrificing the working groups’ dynamism or the teams unity (for example, 9 of the 31 group members are involved in more than one theme).

Weak points
• **Teaching in M2** : The team’s effort in M2 teaching is currently insufficient, which creates problems for local recruitment of doctoral students. Rather than mount a specific M2, we envisage greater involvement in existing or soon to be created streams (Systems and networks, Operations Research, IPB).
• **Doctoral students** : Even though the global numbers of doctoral students are reasonable (26 for the evaluation period), there are significant disparities between themes. Nevertheless it is important to note that the record of doctoral graduates is excellent.
• **Industrial contracts and software production** : The team’s technology transfer activity is not sufficient. Nevertheless it has improved since the preceding period (transfer to Alcatel of compact routing expertise, transfer to 4SH of distributed algorithmics expertise for example). It is important to continue this effort in the upcoming period.
• **Non-publishers** : The team has a significant number of “non-publishers” according to AERES criteria. However this label covers contrasting situations and 3/4 of the “non-publishers” have a research activity (working groups, participation in team contracts).

Opportunities
• **Relations with INRIA** : The team is currently involved in two INRIA projects (CEPAGE and RealOpt), even though the centre of gravity of RealOpt is more in the IMB, particularly since the departure of Arnaud Pêcher to Toulouse as a professor. Nevertheless INRIA Bordeaux Southwest’s development allows us to envisage the development of existing project-teams and offers opportunities to widen the range of research themes covered by INRIA project-teams.
• **Recruitment opportunities** : The team’s age pyramid lets us hope for many opportunities for recruitment at the professor level in the coming years. We must profit from this opportunity to develop and widen our expertise in the current themes.
• **Teaching** : The introduction of new Masters streams (Operations Research at the IMB in particular) offers us an opportunity to increase and diversify our recruitment of doctoral students.

Risks
• **Evolution of our supervision capacity** : The team will lose a significant fraction of its senior members (professor and DR) in a few years. It is important that we manage to obtain the opening of level A posts in the team’s research themes.
• **Expertise management** : In parallel, several local habilitation defences will take place. It is essential for the team to maintain the tradition of external appointments so as to develop new expertise while maintaining that which has been developed locally.
2.2 Project and perspectives

In the last four years the Combinatorics and Algorithms team has undergone significant changes: the departure of the “Bio-informatics” and “Visualisation” projects to form the new Mabiovis team, development of the “Distributed algorithms” theme, with the arrival of the INRIA CEPAGE project, and, of course, variations in our research themes with the arrival of new colleagues. The links between themes have strengthened in these four years: a quarter of team members contribute to several themes, the three working groups draw widely among members, techniques arising in enumeration are used in algorithm analysis and the study of combinatorial structures is used in the design of compact encodings. We wish to continue to reinforce these links, for instance by recruitment of colleagues with transverse expertise. One important event in the coming years will be a massive retirement of group members: three of them have already retired in the last four years but as many as seven retirements are due in the next four years, mainly of rank A members. This will certainly have a significant impact on the group, in its supervision capabilities and in its scientific orientation. This situation may weaken us but it also offers an opportunity to reflect on our recruitment policy and research directions. We mention some possibilities: strengthening the “Algebraic Combinatorics” group (in the short term it will have only 2 rank B members), putting more emphasis on graph algorithms which are not currently well represented, reinforcing our links with the Mathematics institute through the INRIA RealOpt project, as well as the transverse subjects which have appeared in recent years (probabilistic algorithms, algorithm analysis etc.).

2.2.1 Distributed algorithms

The LaBRI’s “Distributed algorithms” theme has, over several years, assembled and recruited researchers with different approaches and complementary expertise in areas covered by the keywords: combinatorics of words and graphs, logic, parallelism, probability, complexity, algorithms. We will profit from this diversity in our fields of expertise to continue to develop this area in studies, already under way or new, which combine these fields. The theme’s weekly working group is and will continue to be particularly useful for attaining our objectives.

Modeling  We study an original formalism in this domain (called the « local computation » model): the rewriting operations and the rewriting rules applied correspond to local computation steps.

In the same way that we have established a link between the message passing and local computation models, we must study the links between the local computation and communication by shared memory models.

An important direction to develop is the modeling in this formalism of WIFI networks, to have a useful model which would allow us to determine the tasks realisable or not (for instance leader election).

The study of local computation through the mobile agent paradigm may allow us to develop solutions resistant to certain types of fault and functioning in dynamic networks.

In the local computation framework, various results on the termination problem have been obtained. A natural follow up of this study is the problem of computation of a global state in particular for fault tolerant algorithms, such as self stabilising algorithms.

Probabilistic algorithms  We are interested in the development of efficient (even optimal), in terms of messages, probabilistic algorithms in the context of anonymous networks of arbitrary topology where the vertices have no knowledge of the network structure. We have proposed and analysed algorithms for two classical graph problems: computation of a maximal independent set of vertices and vertex coloring. Similar studies should be undertaken for the matching problem and the computation of a dominating set.
Self stabilising and robust algorithms  A self stabilising protocol guarantees automatic return to normal functioning after a sequence of perturbation, without external intervention. Unfortunately, no property can be guaranteed during the convergence to correct functioning. An approach complementary to that of self stabilisation is robustness. A protocol which is self stabilising and robust guarantees return to a minimal service in a very short time lapse compared to that for stabilisation.

Our current objective is to design, verify and evaluate a self stabilising and robust hierarchic routing protocol. We hope to confirm that this approach is really promising and well adapted to the management of very dynamic networks, those whose topology changes very frequently because of faults, movement etc. If the usefulness of our approach confirmed in this initial stage, we will then work on the design of self stabilising and robust protocols for other management tasks in very dynamic networks (for example : resource allocation protocols, computation sharing, spectral allocation).

Peer to peer networks : BitTorrent  It is specific to BitTorrent that each file is divided into segments. Thus the uploading of a file is accomplished by different peers, and moreover the segments of one file may be uploaded in an arbitrary order. Using these two properties we have defined uploading rules (segment ordering and peer server choice) tolerant of malicious users. This protocol is \(d\)-safe : it is only after loading \(d\) copies of the segment that the peer can become « source » of this segment. This type of strategy is tolerant of certain peers halting during the uploading or peers having corrupted segments. We must study the theoretical limits of this approach.

2.2.2 Enumerative and algebraic combinatorics

A characteristic of the “enumerative and algebraic combinatorics” theme is a wish to attack major problems which can be described as “long term” because they demand long periods of work from researchers (in our group and at the global level). A certain stability in its structure enables our group to be at the forefront of work on these major open questions.

Alternating sign matrices (ASMs)  In the continuity of the ANR MARS project, we intend to continue the study of these objects which arise in difficult problems in combinatorics, probability and statistical physics. Two major objectives stand out : enumeration of ASMs invariant by symmetry of both diagonals (the last great enumeration conjecture) and a study of the Razumov-Stroganov conjecture describing the stationary distribution of a Markov chain defined on planar matchings, which involves the number of ASMs.

Counting colored planar maps  The question is to determine the generating functions of families of planar maps, weighted by their size (for example the number of edges) and by their chromatic polynomial \(P(q)\), or more generally by their Potts polynomial, which counts the proper or improper colorings according to the number of improper (monochrome) edges. We have obtained a first family of results : for certain values of \(q\), the function that we are looking for is algebraic. We are now attempting to prove that, for any \(q\), the function is differentially algebraic. This would constitute a major step forward compared with the models solved up to now, all of them algebraic.

Planar paths  We have recently conducted a classification of problems concerning paths restricted to a quadrant. Each of these problems has an associated group and we have found an approach which solves all the cases for which this group is finite, except one! Solving this last case is a priority. Beyond that, we have discovered more general conjectures for the number of paths restricted to an open angle.

Two-stack sortable permutations  We have made progress on an old (dating back about 40 years) problem : counting permutations which can be sorted using two stacks acting in parallel.
We have determined a system of functional equations defining the generating function of these permutations, and from this we are able to derive the asymptotic behaviour of their number.

**Gas models and directed animals**  It is known that finding counting formulas for directed animals on the square grid according to their area is equivalent to calculating the density in a gas model with inelastic particles. In the same way, it is possible to show the equivalence between counting directed animals according to their area and perimeter (a quantity linked to site percolation) and the computation of the density of a gas. In order to compute this density we are looking for a complete description of this gas.

**Algebraic combinatorics**  The recruitment of a CNRS researcher (CR) has strengthened the position of this research theme within the group.

Among our objectives, we would like to understand the structure of the ideal of symmetric polynomials in non-commutative variables. Although the analogue in commutative variables is well known, the case of non-commutative variables remains to be elucidated. In particular, we conjecture that the quotient of the polynomial ring by this ideal is of finite dimension, which was unexpected and deserves an explanation.

Another line of research is the study of formulae connected with the characters of the symmetric group. Among our primary objectives are a better understanding of the families of maps appearing in certain coefficients of Stanley’s character formula and the extension of this formula to other groups.

The moment problem for Laurent polynomials is to give a description of the Laurent polynomials orthogonal to all the powers of one given Laurent polynomial on the unit circle. A more complex behaviour than in the case of polynomials seems to be linked to the decomposition into irreducible components of the natural representation of the monodromy group of the Laurent polynomials in question; the nature of this link is an important question.

Hypermaps are branched coverings of the sphere minus three points. But the torus minus one point has the same fundamental group as the sphere minus three points. Consequently, the so-called “combinatorial” hypermaps should also serve as descriptions of the coverings of the torus minus one point. We wish to study the correspondence between the two models.

### 2.2.3 Graphs and applications

Certain “difficult” problems or conjectures in graph theory have always been “attraction points” for our team, generating work which has enabled us to make progress in the understanding of the objects and tools concerned. We intend to follow this line which has so far proved to be very fruitful.

Moreover, our expertise in these fields and the evolution of our discipline lead us regularly to investigate new application areas. We naturally wish to consolidate the work already completed in this direction and to consider new research lines, notably in the areas of network communications and algorithms treating large scale data.

**Havel’s and Steinberg’s conjectures**  It is an important challenge to determine sufficient conditions for a planar graph with a given number of triangles to be colorable (Steinberg’s conjecture and Havel’s problem). We have proposed a relaxation of Havel’s problem suggesting a new angle on the 3-colorability problem and the results obtained to date extend to list colorability.

Recently a new approach based on an extension of “good pre-colorings” on “good cycles” has been proposed by Borodin et al. This approach is a real turning point in the techniques proposed and their results represent a step forward towards a resolution of Steinberg’s conjecture. We propose to develop this new approach for Havel’s problem.

**Oriented colorings, colorings of colored mixed graphs**  Colorings of oriented graphs have always been the source of many open questions (notably determining the oriented chromatic
number of the family of planar graphs, known to lie between 18 and 80). The tools used up to
now turn out not be sufficiently appropriate and it seems that it will be necessary to study more
closely the links between structural properties of planar graphs and those of possible universal
target graphs. We equally study many other questions in the field of colorings of colored mixed
graphs which generalise classical and oriented colorings.

Identifying codes Our objectives in this work (ANR IDEA 2009-2011) principally concern
the investigation of dynamic versions of identifying codes: on the fly construction of identifying
codes and adaptive approximation of identifying codes. In this approach the vertices can be tested
sequentially and the sequence of vertices to interrogate will be constructed gradually in depending
on the answers of vertices already interrogated. This leads to an identification strategy and so to
a decision tree, rather than the determination of a static code.

Topological graphs In the framework of the ANR GRATOS (GRAph and TOpological Struc-
tures), we intend to develop a project whose aim is principally to define a complexity measure
on graphs based only on topological properties as opposed to approaches based on tree width
(Courcelle) or excluded minors (Robertson and Seymour).

Navigable graphs It is an enormous challenge to manage to treat high dimension data without
the corresponding geometrical graphs (bounded stretch metric space spanners) exploding in size. In
fact, current constructions propose spanners whose average degree is exponential in the dimension
of the data. We hope to propose navigable spanners whose stretch factors and average degrees are
polynomial sized.

Graphs and data bases Many data mining algorithms can be re-interpreted as graph algo-
rithms, typically in terms of dominating sets. This approach has not been exploited and seems
to be very promising because the best algorithms in the literature are based on integer linear
programming and are too slow for practical application.

2.2.4 EPI CEPAGE

Computation on distributed platforms We are particularly interested in defining and va-
lidating models which are pertinent and achievable for large scale platforms (topology, latency,
bandwidth, dynamicity, etc.). The aim is to obtain a model with relatively few parameters so
as to be efficiently embeddable on dynamic platforms, and to determine particular topological
properties permitting the definition of guaranteed efficient algorithms.

In this framework, we have recently introduced the bounded degree multi-port model. Our
objective is to use resource augmentation techniques to obtain guaranteed sequencing and to use
adversary model based techniques to demonstrate approximation results in the dynamic platform
setting. It is also very interesting in the definition of algorithms to dispose of more qualitative
models expressing structural properties of the underlying network. In this setting we are also
interested in the construction of geometric spanners in metric spaces which efficiently model the
topology of the Internet (e.g. 2+1D used by Vivaldi, δ-hyperbolic space, etc.).

In the longer term, our objective is to propose, validate and diffuse a set of high level services
allowing efficient programming of large scale platforms (a decentralised clustering mechanism,
procedures to carry out computations using resources close to the data).

Fundamentals of mobile agent computation The objective is to study the computation
capabilities of systems of mobile agents under different hypotheses. A major parameter of mobile
agent models is the agents’ interaction capability. In addition we can distinguish synchronous and
asynchronous environments, the presence or absence of identifiers, bounded or unbounded memory,
etc.
Our study will concentrate on some classical mobile agent computation problems. The most natural is the exploration or mapping problem, in which one or several mobile agents must explore or map their environment. In the rendez-vous problem, two agents must meet at a node of the network which is not specified in advance. Two other fundamental problems concern security (graph exploration in the presence of malicious nodes and capture of a malicious agent) which is often the principal concern of real mobile agent systems.
3.1 Self Assessment

3.1.1 Strong Points

- The Image and Sound team is one of the rare examples in the world to summon in a same group researchers working on the different instantiations of digital multimedia information: from image to sound, through video. This ensemble of competences allows the members of this group to share their visions and experiences, and to work on common problems with rich and various approaches.
- The research themes of this team consider every area related to digital multimedia information: capture, modeling, representation, indexing, processing, analysis, synthesis. Handling the whole processing chain allows a better comprehension, and thus leads to better results, for example to efficiently manage huge multimedia databases that are now commonplace.
- The research team has a broad spectrum: from fundamental research to applications, in collaboration with companies, through software engineering.
- For all these reasons, the team is well-known at both the national and international levels, as shown by our participation to journals, conferences (steering committees, organisation committees, best paper awards), to research networks of the European Union, ANR, Region (participation and expertise), to actions for the structuring, animation, diffusion of the research, and to academic and industrial collaborations.

3.1.2 Weak points

- We face difficulties for the development of software applications because of insufficient means in terms of software engineers.
- The Video Analysis and Indexing theme lacks an associate professor, for example to handle further collaboration opportunities.
- The Sound and Music Modeling theme has industrial contacts mainly with small companies. Job opportunities in this domain are thus restricted, and the students are reluctant to join the research theme.

3.1.3 Opportunities

- Collaborations among the different themes of the team already exist, and shall be encouraged and developed.
- Thanks to the digital revolution, many opportunities for industrial collaboration exist (new replacement format for MP3, new applications in video, etc.).

3.1.4 Risks

- The reduction of the number of students in the Master of computing science, mainly in the multimedia, reduces the potential number of future PhD students in our research team.

3.2 Project and Prospects

In recent years, the group Image and Sound has developed innovative methods allowing the analysis and modeling of many different types of multimedia data (sound, 2-D/3D images, video, 3-D data for virtual reality). Based on the acquired competence, the group has identified many different prospects for the next four years. These prospects can be roughly divided into four main groups:
- **Continued research into fundamental aspects**: The analysis and modeling of multimedia data is faced with complex scientific problems for which the research group has suggested new solutions in the past. Despite this work, in many cases, for instance when it comes to structuring and representing discrete partitions in imaging, 3-D modeling, analysis of audio or video signals possibly with the help of meta-data, or extraction of characteristic and relevant parameters for indexing multimedia data, many open problems remain. Continuing the research on these fundamental aspects is essential in order to succeed with the prospects in the following groups.

- **Development of complex processing chains**: The concrete realization of an application based on the manipulation of multimedia data often requires the use of a compete processing chain going from the acquisition of the data to the presentation of the result, with intermediate steps for complex analysis, modeling, and interpretation. The links of these chains are often not independent. For this reason, developing a link without considering the entire chain can easily lead to sub-optimal solutions. The research group has already made considerable efforts in developing complete chains in the past few years. Despite these efforts, much work remains to be done such as developing a complete chain for digitizing 2-D and 3-D images from the acquisition of digitized documents to the distribution of digitized and analyzed documents, or taking into account the acquisition of video streams in 3-D modeling. Controlling the entire processing chain is also an important aspect when it comes to joint projects with industry.

- **Obtaining high-level semantic information**: The analysis phases mention above supply important information about the analyzed signals, but still often supply only low-level information that is not directly useful in many applications. The problem of obtaining higher-level semantic information is a vast research prospect for the research group in a domain that has been little investigated in the past. Obtaining such information might for instance involve interpreting behavior of objects in a video, or creating a visual summary of it, or to propose methods for transforming musical data represented as parameters extracted from a preceding analysis phase.

- **Development of interactivity**: Many applications assume the interaction between an end user and the data. The way in which this interaction can be realized often requires the creation of complex procedures. As an example, one might consider the modification of musical parameters, or the way a user might interact with a virtual-reality model.

The next following paragraphs give a more detailed description of the research directions for each one of the themes represented in the research group. Note that the group also suggests reinforcing the collaboration between the different themes in order to take better advantage of complementarity and existing competence, in particular when it comes to developing complex processing chains.

### 3.2.1 Theme: Analysis and Indexing of Video

In this theme, there are a number of research prospects. When it comes the paradigm of primary and scalable indexing, we will continue applying our tools on a large scale in the context of international evaluation campaigns such as TREC Video where we coordinate the participation of the national consortium IRIM GDR-ISIS. The convergence of the tools is very enriching - tracking tools are to be used in scalable object-based video segmentation. Furthermore, we have proposed a first scalable statistical descriptor for objects in the transform domain. At the moment it is intra image. In the future we plan to explore the temporal dimension and orient our work towards descriptors of type SIFT space/time coupled with interesting regions and objects. Another research direction consists of developing scalable tools for creating visual summaries. The main components of these tools, i.e., classification and extraction of indices, are becoming mature in the supervised case, and should be further developed for supervised classification. It will be necessary to develop these tools depending on the taxonomy of scenarios defined by the users of the contents, for example in the case of embedded monitoring. For that particular case, we have started collaborative research in a joint project with the IMS (UMR5218 CNRS, ENSEIRB, ENSCPB, University of
CHAPTER 3. TEAM IS 25

Bordeaux 1, Department LAPS/Signal and Image). This research will continue in the context of joint projects between the LabRI and the IMS.

In the context of analysis applied to video-surveillance problems, the first prospect consists of developing the most generic approaches possible, allowing semantic interpretation of the behavior of objects in the scene based on motion analysis.

The second direction consists of refining methods based on the use of more than one camera for video surveillance.

3.2.2 Theme : 3-D Modeling, Visualization, and Interaction

Evolution

This theme has existed for more than twenty years. Initially focused on modeling (based on topology either by interpolation surfaces or implicit surfaces) and rendering (photo realistic, real time), it evolved some ten years ago towards virtual reality and 3-D interaction. In November of 2003, the INRIA project IPARLA was created from this theme. This project handles problem of visualization and interaction for mobile users. This continuing evolution of our interests is the result of spontaneous discussion and debates and of answers to external evaluations. During the reference period, we have organized two research group seminars, each one two days long, where we mixed on the one hand current activities that should be reinforced, and on the other hand activities yet to be initiated. From this reflection resulted a number of new staff members:

- in 2005 : Interaction (CR INRIA), Modeling/Rendering (lecturer Bordeaux 2),
- in 2007 : Expressive rendering (CR INRIA, after the departure of a lecturer), Video acquisition (lecturer Bordeaux 1),
- in 2008 : Modeling/Rendering (CR INRIA),

that today allow us to have two permanent staff members on each main project. It has now become necessary to recruit a senior researcher (rendering) in order to balance the composition of our theme. We have also established a plan for PhD students and post-doctoral students using the same strategy so that each new permanent staff member immediately can work with a student on his or her project. Unfortunately, because of a lack of candidates, we have not managed to recruit as many PhD students as we would have liked; we will of course continue this effort. Note that our PhD graduates have found positions as teacher/researcher (1), researcher (1), engineers (5) or post-doctoral student (1). The hiring of engineers has been organized in terms of projects (ANR, Pôle de compétitivité) in order to create pairs consisting of a permanent scientific staff member and an engineer for each project.

Visibility

The period has been fruitful in terms of publications, both concerning quantity and quality, in particular with three “Best paper” in major conferences (Eurographics, 3-D symposium on User Interfaces). For that reason, we will continue our research in this domain. We have also significantly increased our presence in program committees of conferences and journals. Thanks to the results we have obtained, we have been awarded a large number of grants (8 ANR, 1 Pôle de compétitivité). Note for instance, the ContInt 2009 program with 108 submissions, 23 selected proposals, two of which are from our theme and one from the theme MSM. On the average, we would like to maintain three or four projects per year, which requires us to apply again in 2011. During the evaluation period, we have developed a large number of recognized international collaborations: two INRIA research themes, one STIC-Asia project, one JST-CNRS, one LIA CNRS-Japan. These relations are mainly focused on Asia (China, Japan) and Canada. As a natural consequence, most of our foreign visitors come from these laboratories. Thanks to the international visibility that we have acquired, we have organized one of the two most important conferences in Virtual Reality (ACM VRST) with a record number of submissions and participants. Together with the Bunraku group (Rennes), we co-organized two conferences in China with the goal of developing scientific relations between our countries. We have also contributed to national events by organizing first the AFIG
days and then the AFRV days. During the next evaluation period we will continue this work and we are already thinking about organizing a Eurographics symposium. We have developed relations of various kinds (Industry, CIFRE PhDs, ANR projects, Pôle de compétitivité, PhD consultants), with several very different companies (small and large, French and foreign). We plan to continue and reinforce these collaborations.

Scientific Prospects

In this section, we present a non-exhaustive vision of the future evolution of the themes. **Acquisition from video flow** : This theme started during the evaluation period with a PhD and was made permanent with the hiring of a lecturer. This theme is growing in importance and is now at the center of the growth of our activities. We are considering hiring a post-doctoral student in the short term, and perhaps a permanent staff member somewhat later. **Acquisition of appearance** : We would like to develop this subject rapidly in order to complement our work in the domain of geometry acquisition. This work has already begun in collaboration with the CEA and the initial results have encouraged us to submit a proposal for a European project. **Modeling** : This the historical core of our activity, and it is evident that we will continue our research in this domain that was also reinforced recently by the hiring of a new researcher. **Expressive Rendering** : This theme started immediately before the beginning of the reference period, and has been reinforced by the hiring of a research, and we insist on continuing our work in this very active domain. **Tactile Interaction** : We initiated this activity at the end of the period and we intend to make this one of the core activities of our work. This desire has become very concrete for instance by the recognition in the form of the ANR project Instinct. **Augmented Reality** : This activity was explicitly initiated with the ANR project RaXenV, and is becoming increasingly interesting to use. For that reason it will be further developed in the future.

### 3.2.3 Theme : Modeling Sound and Music

The researches conducted in the MSM theme deal with the elaboration of models according to two levels of representation : the sound level and the music level. Designed first for batch processing, these models now aim more and more at real time capabilities for live performance. Thus, we want to include in the representation some information related to the kind of interactions that will be allowed in real time.

Regarding written scores, the interactive score model is integrating temporal information allowing the interpretation, and it will be extended to other kinds of interactions including browsing in open compositions.

Regarding sound models, the study of the watermarking of the sound representation will be conducted in order to facilitate the separation of the sound sources as well as the interactive manipulation of these sources.

On a scale measuring the model expressiveness, one can situate the applications we will focus on during the next 4 years :

1. **Active Listening** : the aim is to allow the listener to change other parameters than the classic volume and balance. Thus, the listener is invited to go from a passive to an active listening of the music, however staying in the limits of the possibilities allowed by the composer and still with a limited expressiveness.

2. **Musical Games** : the aim is to allow musicians to play and interpret orchestral scores with other musicians or with the computer, with rich and various expression modes.

**Active Listening**

From a sound signal, such as the sound fixed on a physical support like the audio Compact Disc (CD), we want to analyze the different sound entities perceived within this musical sound, in order to be able to transform each entity separately. If this analysis and these transformations are performed in real time, we can offer the listener an experience of **active listening** : modifying
the music while it is played, for example by cancelling the singing voice (karaoke effect), by changing the volumes of the different instruments (for example by fading the drums out), or even by changing their positions (for example by moving a guitar from the left to the right). Although these manipulations are common practices of the composers of electroacoustic music we are collaborating with in the context of the SCRIME, they are quite revolutionary for the public, used to a more passive listening experience.

**Informed Source Separation** Sound source separation is a major scientific challenge. Indeed, we perceive the musical mix with our 2 ears only (the reason for the 2 channels on audio CDs), whereas in a musical sound the number of sound entities present simultaneously (polyphony) is generally far greater than 2. We are in a “degenerated” case (mathematically speaking), with less equations than unknowns, that is less channels than sources. However, the brain is able to find a solution. Mimicking the mechanisms of our perception, we perform a frequency (spectral) decomposition of the signals and we consider that for each signal there is only one dominant sound entity present at each point (atom) of the time / frequency plane (hypothesis of orthogonality of the short-term spectra).

Then, we must cluster the spectral atoms into sound entities, for example by using perceptive criteria based on laws of acoustics and psychoacoustics: remarkable spectral structures, simultaneity and similarity in time, spatial coincidence, etc. In order to help this classic CASA analysis (*Computational Auditory Scene Analysis*), we propose (in collaboration with GIPSA-lab, Grenoble and Télécom ParisTech) an original approach consisting in taking advantage of annex information added to the sounds in an inaudible way (audio watermarking).

**Musical Transformation** Active listening could also concern in the near future musical parameters such as the metric, tempo, rhythm, tonality, etc. The user might be able to modify the tempo of a song, its tonality or its chord grid. For this reason, we wish not only to continue the quest for analysis methods (possibly informed) of the musical information from the audio signal, but also to propose transformation methods for some of these parameters. Thus, we consider working on algorithms for the synthesis of musical pieces (or accompaniment) based on an automatic pre-analysis of the music.

**Interaction** Moreover, we consider modeling the interaction by a constraint optimization problem: from a realization proposed by the composer, the listener will be able to perturb some of the parameters, leading to a constraint propagation.

**Musical Games, Tools for Musical and Instrumental Pedagogy**

Musical games constitute a privileged application domain for our researches. Indeed, we plan to extend our research results for a notable amelioration of such applications, but especially an innovation regarding expressiveness and musical pedagogy.

**Multi-Player Temporal Modeling** A multi-player score has to be adapted to different playing modes while preserving the expressiveness given to the musicians as well as the of the accuracy of the interpretation by the computer. From the model for interactive scores, we will have to study the several strategies for the transformation of the affectation of the interaction points and of the temporal relations in function of the configuration of the events for triggering / stopping each voice of the score.

**Interpretation Comparison** In the context of the SIMBALS project, we have designed and evaluated algorithms for the estimation of musical similarity. We wish to extend this work and find applications to instrumental pedagogy and musical games. Indeed, we consider proposing an accurate system for the estimation of the interpretation quality of an instrumentalist / player, and also
CHAPTER 3. TEAM IS

for the comparison of musical interpretations, in order for example to allow the automatic coordination of an ensemble playing. Other research perspectives deal with automatic accompaniment, based on various musical properties (rhythm, harmony, etc.).

**Pedagogic Evaluation in the Context of the SCRIME** The collaboration with the composers of the SCRIME has lead us to new pedagogical methods for the teaching of music in schools. However, we face the problem of the evaluation of these methods: How to measure the progress of a pupil and how to adapt the pedagogy in function of his/her results? We think that, in the context of the pedagogy by musical dialogs, the search for similarities might bring some answers to this problem, that will also be useful in the context of musical games.

**Control Interface for Interpretation** We wish to continue our researches on 3D control interfaces and opportunistic interfaces while aiming them towards musical interpretation, for which very few interesting results exist. Indeed, the graphical representation shall allow the musician to locate in a moving temporal context, where the time locator moves in the score (or the opposite) with possibly a symbolic notation of the gestures as well as the expressiveness induced by the intention of the composer. But the main challenge consists in integrating in a unique graphical representation a mix of static events (written on a time line) and dynamic events (in a stream) that change the score itself.

### 3.2.4 Theme: Image Analysis and Structuring

**Prospects**

For problems related to finalized research, we rely on the mastery of the foundations of imaging, mostly in the discrete domain. For that, we intend to continue investing in research in the domains of structuring and representation of discrete partitions, or the extraction of discrete geometric parameters. This, in association with the richness of our multi-disciplinary collaborations (physicists, physicians, anthropologists, mathematicians, etc) allows us to handle problems based on the acquisition of real images in 2-D, 3-D, and dynamically. The main illustration of these dynamics is our work on the document digitization. The hiring of Nicholas Journet as lecturer gave us new competence in this domain. In addition, the tools for segmentation and indexing developed by the theme are now sufficiently mature and robust to be directly used in this context. Our theme is involved in the POLINUM project (collaborative project financed by FEDER. The overall budget is 4M Euros, and that of the LaBRI 500k Euros), the goal of which is to develop a “laboratory factory” dedicated to image digitization. The ambition of this project is to produce a complete digitization chain from the acquisition to the recomposition and distribution of digitized documents. In collaboration with small companies (I2S, Arkenum, etc.) and CEA LIST, the research activities of our theme will concern the correspondence between digitization quality and usage, and the logic segmentation of the document. Following initial contacts, the laboratories of La Rochelle and Tours will like be integrated into this project very soon.

A second prospect is related to the 3-D imaging chain. Initial research related to the study of this chain, from acquisition to production, has allowed us to make progress on each of the links of the chain. At the moment, we have a first complete version of this chain that we can apply to different models of volume (MRI, X-ray scanner or microscanner, microscopy, terahertz imaging, etc) or surface (laser scanner) acquisition. The objective is now to further develop each link while paying attention to their coherence and precision. In this prospect, we will particularly concentrate on the links “acquisition/reconstruction” and “3-D segmentation”. In the first case, we will concentrate on tomography techniques an on the possibility of integrating the image processing at the reconstruction level. In the second case, we will in particular integrate tools for discrete geometry in the segmentation procedure in order to guide that procedure or to extract characteristics required for analysis. With respect to this case, we are confronted with an open research problem consisting of the extraction of geometry characteristics in a noisy context (i.e.,
“real” images). The goal is to supply a platform for 3-D segmentation that is equivalent to the one developed for 2-D in terms of robustness and functionality.

An other goal is to be able to use dynamic 3-D partitioning techniques, i.e., to be able to extract and track characteristics in a partition that evolves dynamically over time, while still maintaining a coherent partitioning. In particular, we need to formalize this evolution and establish a correspondence between different discrete chunks of this partitioning on images acquired at different moments. For instance, we just started work on the correspondence between bronchus trees obtained from inhalations and exhalations. This work, in collaboration with U885 unit if the CHU (Hospital) in Bordeaux, will allow a better understanding of pathologies of the bronchus walls during the entire respiratory cycle.
Team LSR

4.1 Self assessment

4.1.1 Positives

- The LSR research group has many national and international collaborations, with both academics and private companies. In particular, we are quite successful in proposing and/or participating in European projects; several such projects will begin during the next four years period (see below).
- The group is strongly linked to high-level curricula at the University (Master) and at the Enseirb School of Engineering.
- The group has developed a number of software that have national and sometimes international visibility (Shaadhoc, strategic map sharing in MANETs; JCatools, software suite for smart cards; TGSE, simulation, emulation and test generation for time constrained systems; the Caliph platform, timed automata manipulation environment; etc.).

4.1.2 Negatives

- Even if the number of publications is good, the number of papers in magazines is probably less than what one might expect.
- Efforts have been made to develop relationships between the different topics of LSR and joint projects are being set up but these are not yet sufficiently effective.

4.1.3 Opportunities

In terms of opportunities we have identified the following facts:
- The willingness of Thales, the University and the INRIA to set up a joint laboratory should allow the relationships that exist with LSR to develop.
- The introduction of the DAS (Strategic Activity Domain) UAVs Systems within the World Competitiveness Cluster Aerospace Valley will make it possible to increase the relationships of the team with the industry in this domain. Several labeled projects are already underway.
- The redefinition of several network related curricula which is currently being done will be the opportunity to lean them even closer to the team. It should be noted that these are managed by members of LSR.
- The activities with the DGA (french army), which is funding PhD thesis, may develop further thanks, for instance, to the DAS UAVs Systems (see above), and more generally thanks to the topics related to mobile communication, to embedded systems and security that are developed in LSR. This will also be made possible because UAVs and security have been declared the major focus of the Defense White Paper, which is a report that defines the priorities of the french army for the years to come.
- The discussions around the creation of an Information Technology Cluster within the University should provide a framework and thus give more visibility to many activities of LSR, especially to the many existing collaborations that are already running with the IMS laboratory.

4.1.4 Risks

- There is a certain imbalance in term of number of member researchers between the components of the LSR research group. The recruitment policy of the group will have to take this fact into account. We may also note that there is no INRIA or CNRS member in LSR.
– An issue is raised by the retirement of a Professor, in charge of LSR, who is also the single professor of the MTSC group. This will have to be dealt with during the next evaluation period.

### 4.2 Project and prospects

The Languages, Systems and Networks research group was created in 2006 to bring people working on a set of consistent activities to work together. The synergies developed in the previous period, even if they are not yet perfect, provide opportunities that will be developed further in the years to come. Projects and prospects of the team are thus about the development of these internal synergies, but also of synergies with other laboratories, private companies or institutional partners. This will be achieved by means of specific projects like ANR and European projects. This is described below.

**Emergence of joint projects**

– Phoenix / SOD : simulation tool for mobile worlds (mobile phones, devices for ubiquitous computing, fleets of UAVs) in collaboration with the Image and Sound research group of LaBRI. This project is still in its infancy, but it should eventually involve a number of other partners such as the DGA and the Fly-n-Sense company.

– Comet / SOD : optimizing the flow of information and energy consumption, with the support of co-design, within a fleet of containers (with the IMS laboratory). The project has been labeled by the World Competitiveness Cluster Aerospace Valley; it is currently in search of funding.

**Collaborations around common technology platforms**

The Laboratory and the INRIA have fostered the development of technology platforms within the LaBRI and across research groups, among which LSR :

– SOD and Comet have a large room, a sort of technology workshop, which makes it possible for them to deploy sensitive pieces of equipment (sensors, mobile phones, a smart cards grid, satellite receivers, etc.) and to host researchers, students and guests visiting the laboratory.

– Phoenix runs the Smart Space : a platform to experiment applications the goal of which is to coordinate distributed heterogeneous devices. This platform allows to deploy and test innovative applications in different areas such as home automation, health care and building management. Scenarios of home automation coming from a collaboration with France Telecom have been deployed successfully.

These platforms which have been set up in the previous evaluation period will support synergies between the groups of LSR (and others) in the years to come.

– A new platform comprising a hundred phones/PDAs running the Android operating system is being deployed by SOD. It is based on both an educational component and a research component, and will make it possible for the LSR research group to conduct large/real scale experiments. It is also expected to host experimentation of colleagues from Vannes and Le Havre and to be used in the framework of the Smart Urban Spaces European project.

**Relationship with IMS**

The various actions of LSR have developed strong relationships with the IMS laboratory. Because of the creation of a IT (Information Technology) Cluster at the University of Bordeaux, we are led to further strengthen these relationships. Among the actions that are already undertaken, we can note the following :

– Since 2007, Phoenix has been working with the IMS on aspects related to embedded systems and hardware. This collaboration takes place in the framework of a contract with
France Telecom (HomeSIP) and a funding of the Région Aquitaine. In 2009, a co-supervised (IMS/LaBRI) PhD thesis began; it is partly funded by the Région Aquitaine.

- Since 2008, a co-supervised (Comet) PhD thesis focuses on the study and the establishment of an embedded multi-service platform for the management of multimedia streams and of network sensors at different software and hardware levels to create a « Multi-Play » home network.

- Since 2009, LSR (SOD) and the IMS develop the SYMM project. The contributions of SOD are the algorithms for UAVs fleets and the associated security features that it has developed and the IMS provides its expertise on hardware aspects. SYMM is labeled by the Aerospace Valley World Competitiveness Cluster.

- LSR and some groups of the IMS co-supervise several student projects combining software, network, security and hardware reconfiguration based on sensor networks or FPGA processors.

Relationship with the industry and with large organizations

The LSR research group will continue to open to the industrial world. It runs many contracts involving companies, direct contracts or contracts managed by institutions (Région Aquitaine, etc.). Among the industrial partners we can quote: Thales, France Telecom, Gemalto, Orange, Airbus, RATP, EDF, Thomson, DGA, Geenesys, Montimage, Viotech, Coronis Systems, etc.

Many of these relationships take the form of projects labeled by the World Competitiveness Cluster Aerospace Valley. This has also helped to develop our relationships with the Toulouse area. For instance: a (CIFRE) Thales PhD thesis, co-directed by the LaBRI (Phoenix) and the LAAS is expected to start in the autumn of 2009 in the domain of dependability in software architectures; the SYMM Project (SOD), labeled by the Cluster and involving, among others, the ENAC (Toulouse) and Cygfox (Toulouse) is in the process of being funded; the VISTA project (MTSC) targeting the test of aircraft systems is supported by EADS, based in Toulouse. The dual localization Midi-Pyrenees and Aquitaine of the Competitiveness Cluster will make it possible to further develop these relationships between the Bordeaux and Toulouse areas.

Phoenix has developed very close collaborations with Thales as evidenced by two running PhD thesis and one PhD thesis that has already been defended. These collaborations are in the field of software engineering platforms for communicating objects. They will develop within the common laboratory that is going to be created.

SOD shares a special relationship with the DGA around algorithmic aspects in mobile fleets (infantry, UAVs, etc.). This takes the form of a PhD thesis that has already been defended and one that will continue during the next evaluation period. A PEI project is also being discussed. Moreover, this collaboration also involves the IMS and the Aerospace Valley Cluster. We hope to give more visibility to this relationship.

ANR Projects

The following ANR projects will begin or continue during the next four years: TESTEC (MTSC), a project to enforce the use of test and verification techniques for critical industrial systems (EDF) designed by control engineers; Webmovie (MTSC), a project on testing Web services, new types of services requiring new techniques and test architectures; DIAFORUS (Comet), Distributed Applications and Functions Over Redundant Unattended Sensors, which began in March 2009; ARDMAHN (Comet), which studies dynamically reconfigurable architectures and methodologies for self-adaptation in the context of home networking. The LSR research group is also working on several additional proposals.

European Projects
CHAPTER 4. TEAM LSR

Three European projects will begin during the next four years period: the IP project ALICANTE (Comet), media ecosystem through the deployment of content-aware ubiquitous network environments; the STREP project ENVISION (Comet), Enriched-aware Network Video Services over Internet Overlay Networks; the ITEA2 project SUS (SOD), Smart Urban Spaces.

Of course the research activities of LSR are driven by the major issues and challenges related to: future generation networks, modeling and associated checking techniques, ubiquity, mobility and security. This is developed in the following pages, by describing the specific projects of each subgroup of LSR.

4.2.1 Comet

Theme Evolution

The recruitment of a senior lecturer in 2007, then a professor in 2008 has enabled the strengthening of the ongoing activities of COMET theme. The current promotion of an associate professor, recruited in 2004 to the rank of university professor assures also, the sustainability of research activities within the group.

Continuation and strengthening of ongoing activities

The COMET theme will continue to develop the research directions Service level Negotiation, autonomic networking and QoS Control of multimedia delivery, concerning the networks of future and the Content-Aware Networks and Network-aware applications approaches. Concerning the research area Service Level negotiation, the effort will focus in particular on the issue of vertical handover and on maintaining the service level in mobility scenarios. The autonomic networking research area will cover mainly the following aspects:

- Proposal of an autonomic and distributed architecture for the management of virtual sessions in order to make flexible communications in the Internet (a thesis on this subject starts in 2009)
- Management of user QoS within converged mobile networks for the efficient broadcast of ambient services (contract Thomson, thesis with Orange Labs)
- Proposal of an autonomic architecture for heterogeneous ad hoc sensor networks (contract ANR DIAFORUS). The QoS Control of multimedia delivery research area addresses in particular the following issues:
- Optimization of multimedia services for mobile terminals
- Optimization of traffic in P2P networks (two current Ph.D thesis are ongoing on this subject among which one with Orange Labs)
- Multimedia ad hoc sensor networks (two theses will start on this subject in the autumn 2009).

Moreover, the COMET theme is participating actively in the project STREP ENVISION (Enriched-aware Network Video Services over Internet Overlay Networks), which is under negotiation with the European community and aims, in particular, to bring together P2P users and network operators to a common collaboration.

Emerging Themes

A new focus was introduced in 2008. This is a transverse area of research that has for application framework the Home networking. It deals, in particular, with the convergence of networks and next generation services and its implication within home networks and extensions to the end users.

This new research theme includes also the study and the implementation of a multiservice platform for management of multimedia streams and sensor networks at software and hardware levels. Two theses started on the home networking in 2008, one in collaboration with the IMS laboratory. Furthermore, two new projects on this subject have been accepted: the ANR ARD-MAHN project (Dynamically Reconfigurable Architecture and Methodology for Self-adaptation in Home Networking) and the IP ALICANTE project (Media Ecosystem Deployment Through Content-Aware Ubiquitous Network Environments).
4.2.2 Distributed Systems and Objects — Muse

Evolution

We have narrowed our field of activities what also strengthened the cohesion with the other topics of LSR. This was necessary because the group had one single permanent member. Today SOD counts three and a half permanent members (1 Professor and 2.5 Assistant Professors) but focusing in term of research domain remains essential. In the next four years, the group will construct his research on its three core competencies, as described in the next section. Distributed systems and grid aspects that target computation intensive environments are no longer key topics of SOD. To clearly state this break, the name of the group is changed to Muse.

Strengthening of ongoing activities

The algorithmic and applications for fleets of mobile communicating devices. We have achieved significant results in terms of modeling and characterization (2009 Sirocco [Cs15]) of algorithms and mobile applications, based on graph re-labeling approaches that we have extended [Cs14] to take into account dynamic environments ([Cs3] and Milcom 2006 [Cs4] and 2007 [Cs5], DGA funded PhD thesis and ANR project Sarah). The goal is now to extend these results to the original field of fleets of UAVs (Milcom 2009 [Cs18]), an area that has been little studied in France and in Europe. This has recently begun within the framework of a new project called SYMM that has been labeled by the Aerospace Valley World Competitiveness Cluster. The interaction of the dynamic evolution of the system imposed by the algorithms (designed to achieve a specific mission) with the flight plans of the UAVs will be a major topic of our work.

Middleware for fleets of mobile communicating devices. One of the goals of this activity is to define universal primitives that can run on any platform [Ca4]. This includes the design of an associated model. The current possible candidate that is still in an early definition stage, is a CCS-like original event based model that could deal with both synchronous and asynchronous communication, unicast and broadcast, with or without loss of messages. The heterogeneity of the target platforms (in terms of code and of communication protocols) and the management of this heterogeneity is also an important topic [Cs9]. This work takes place within the framework of the Adapt action of the GDR ASR of the CNRS. It will also be conducted in collaboration (that has already begun) with the IMS laboratory for aspects regarding co-design, which can help to support context adaptation (safety, software radio, etc.).

Securing mobile systems. The activities in this area will continue through (1) a new ITEA2 European project called Smart Urban Spaces (involving gemalto, the world leader of the smart card industry) and (2) the SYMM project (algorithms in a fleet of UAVs and associated secure systems [Cs18]), which has been labeled by the Aerospace Valley World Competitiveness Cluster.

Full-scale deployment. We are currently setting up a platform comprising 100 mobile phones running the Android operating system. These will be made available to students and colleagues who will in return allow us to run our applications and perform our experiments. The first goal will be to collect mobility traces, the anonymity of which will be guaranteed by the use of embedded smart cards (possibly the SIMs of the phones). We will also deploy our algorithms on a fleet of UAVs provided by the Fly-n-Sense company within the framework of the SYMM project (that also involves the DGA).

Emerging topics

A number of topics emerge, even though they are already part to some extent of our work. These are : (1) distributed algorithms and associated models in a fleet of mobile communicating devices (for instance drones in collaboration with the DGA) and interaction with mobility models; (2) heterogeneity management, adaptation of communication and security to the context (in collaboration with the IMS regarding hardware-software co-design) by using generic/specialised
middleware. We decided on exploring these questions further by applying for a number of resources which we have obtained: 2 BQR (1 University and 1 Enseirb) and 1 PhD thesis starting Q4 2009.

4.2.3 Phoenix

Evolution

A host of networked devices are populating smart environments that become prevalent in an increasing number of areas, including supply chain management (e.g., parcel tracking), monitoring (e.g., building surveillance and patient monitoring) and home and building automation (e.g., control of energy consumption). This situation raises a number of challenges concerned with the interweaving of these smart environments in our daily life, a high demand of applications matching the wide range of user needs, and a large spectrum of expertise covering the combination of areas involved in smart environments.

Over the next four years, the research efforts of Phoenix will address these challenges. To do so, our goal is to develop a software engineering approach that is domain oriented, extending our work on domain-specific languages.

Strengthening current research efforts

We propose to design languages dedicated to the development of software for smart spaces, targeting key aspects such as the modeling of a smart space, the coordination logic of entities, safety and security.

Domain-specific languages. We propose to create languages dedicated to the software development of smart spaces, targeting each key dimension of this domain: modeling of a smart space, expressing the coordination logic, defining security policies, and specifying fault tolerance policies. The specificities of the target dimensions will lead to different design strategies for these languages: specification, configuration, or programming. Because they are domain specific, these languages will offer notations and concepts that are well suited to each dimension.

Verification. The dedicated nature of the proposed languages will enable domain-specific properties to be decidable. To do so, the languages will be designed in conjunction with these properties and will take into account the power of the existing verification techniques and tools.

Compilation. We will develop compilation and optimization strategies that are dedicated to the domain of smart spaces. The dedicated compilers will encapsulate the techniques and technologies of the target domain.

Runtime environment. We will develop new runtime support to manage the dynamic changes of a smart space in conformance with its modeling. This runtime support will be efficient enough to manage thousands of networked entities.

Applications. We will validate our approach in the area of home/building automation. This area gives us an opportunity to generalize our research on telephony. This work on home/building automation will be done in the context of the project to develop a “smart building” named SmartImmo, funded by the “pôle de compétitivité Solutions Communicantes Sécurisées”. Additionally, we are participating in a proposal for European project and one directly submitted to France Telecom.
Emerging topics

We plan to cover more topics that are related to smart spaces. A new thesis starts in October on the topic of fault tolerance, in collaboration with the LAAS in Toulouse and Thales.

Simulating smart spaces has turned out to be a key enabler of the pervasive computing approach. It plays a key role in assembling new concepts about smart spaces, create new services coordinating entities before they are actually deployed, and visualizing the execution of new scenarios.

Last, we would want to apply our work to the area of assisted living. A European proposal is being prepared on this topic.

4.2.4 MTSC

Computer systems become increasingly complex, it becomes necessary to test them in an orderly fashion. The theme MTSC (Modeling and Test of Communicating Systems) acquired a long experience in this field, mainly in the test of reactive systems (communication protocols, constrained systems). A notable feature of this theme is to combine the practical aspects and the theoretical ones.

Theme Evolution

This theme gathers currently 4 permanent researchers (1PR, 3MC). The prospects of research have recently focused in the fields of protocol engineering, communicating and distributed systems, critical industrial systems, embedded and constrained systems. Our activities are mainly oriented in the fields of conformance testing (test of timed systems in particular), interoperability testing and robustness testing. The tool, named TGSE, testing timed systems has been developed.

Two projects ANR Webmov and TESTEC are still ongoing. They will develop research on critical industrial systems and testing Web services. They finance a PhD, an engineer and a post-doc on these subjects. The impulse given by the pole of competitiveness AESE must also be an opportunity to adapt our methods to the test of embedded systems.

Continuation and strengthening of ongoing activities

New models are taken into account, particularly for real-time systems. Our approaches combine the techniques of model checking, symbolic computation and generation on the fly to reduce the combinatorial explosion inherent in most techniques of automatic test generation.

The test techniques (conformance, interoperability and robustness) and automatic generation of test sequences (with calculation of coverage) from the formal description are to deepen to take into account aspects of real-time, distribution and mobility. In the same way, testing embedded systems, systems with components or Web services bring problems on observability and controllability that also need more deeper researches. All of this research has a a practical aspect with implementation of methods and a theoretical aspect with formal studies.

An other aspect still little studied in this theme is the the implementation of automatic generated tests in a given architecture. The consideration of time constraints and the distribution of components warrant further research.

Emerging Themes

Two main areas (not disjoint) emerge from the perspectives of the theme :

- test Web services. The problems of controllability and observability are already under study. Some early work on test generation for such systems propose an integrated approach starting with a high level description of a composition of a Web service to the test this composition : test architecture, test generation and execution of the generated tests. The proposed test architecture can incorporate an approach based on fault injection : passive testing techniques will be studied based on "monitoring" mechanisms already existing.
testing critical constrained systems. Research works on generating test sequences from timed systems with shared variables began. Combining verification techniques and testing will improve combinatorial explosion problem. This work should be widespread in test systems based on components communicating.
Team MABIOVIS

Our four year activity report underlines the links our team has with fundamental research conducted at LaBRI (combinatorial mathematics, formal methods, graph theory). That being said, our research activities opens the scope of research taking place in our laboratory and covers areas such as data mining and data integration – as show our results in the area of databases [Cs14, Ca49, Au33, Ca8, Rs32, Rs17, Rs21] or results concerned with clustering and classification [Rs34, Ca18, Ca7, Ca34]. We could also mention our contributions to human-computer interaction [Cs4, Ca24, Ca23], although LaBRI belonging to the Image et Son team already invest this area (more particularly the MVI3D group).

The first part of this chapter is devoted to self-criticism. The second part lists a number of questions we plan to investigate in the coming years.

5.1 Self Assessment

5.1.1 MABioVis Strengths

An ambitious and wide scope research program One of the strength of our team is to address theoretical as well as applied problems, while including applications to real life problems (see [Rs41, Ra3]), sometimes directly linked with the industry. This wide scope strategy is coherent with the ambitions of our team : systems biology must answer biologists with models of the cell that are relevant at all levels, from the biochemical to the cell and the tissues. “Visual Analytics Science and Technology” in information visualization calls for integrated approaches combining analytical and visualization techniques to help explore and analyze massive, dynamics and sometimes uncertain data.

On the theoretical front, we study objects from discrete mathematics : permutations with/without forbidden patterns [Au23, Au25], algebraic languages and tree structures [Rs15], the coding and compression of multiscale tree graphs as well as self-similarity properties [Ra16, Rs29, Rs16] ; graph hierarchies computed based on optimization criteria have always been at the heart of our concern [Rs4, Rs5, Ca3]. We focus on proving properties, as well as designing efficient algorithms to generate or decompose them; establishing the behavior of our algorithms; finding astute uses of the objects and algorithms in various application domains.

Maintaining strong links with our environment Because we conduct multiple activities, we come across numerous opportunities to build collaborations with actors of the scientific community at all levels : local, national or international as shows our list of research grants (see our accompanying research report). The emergence of the SIMBALS transverse project can be seen as a direct consequence of the links we locally develop with our colleagues. At a national level, we are involved in a number of research groups (GDR) concerned with our research areas (Génélectures, BiM – bioinformatique moléculaire, MABEM – modélisation mathématique en biologie et médecine, IM – informatique mathématique, I3 – information, intelligence, interaction). Visits of foreign collaborators, together with our involvement in European projects reflect how we link with the international community.

A well structured software prototyping and development activity Because we ultimately seek to transfer our research in various application domains, we need to build prototypes validating our results [Rs22, Ca52, Ca80, Ca5, Rs2]. Indeed, the community is convinced of the value of a novel algorithm only after it has been confronted to a benchmark dataset [Rs25, Rs11, Ca77, Ca68]. The efficiency of an integrated approach or system often is confirmed by showing how the system behaves in real life case studies. As a consequence, our team has acquired some experience in

1. See for instance [Rs13, Rs2, Ra16, Rs25, Rs28, Ra18].
2. See for instance [Rs4, Rs6, Rs8, Rs22, Rs23, Rs34, Rs35, Rs36, Rs40, Ra11].
3. Voir le site du NVAC et l’ouvrage fondateur "Illuminating the path". http://nvac.pnl.gov/agenda.stm
software engineering. Many of our software realization who began as simple prototypes became widespread software used by a vast community of users. The Génolevures database, the BlastSets tool hosted by the CBiB and the graph visualization framework Tulip are examples we are proud of. Users report their activity using these tools, providing us with an interesting feedback on this part of our activities.

This experience and know-how we develop in software engineering comes together with a real strategy to keep it alive and ongoing. Université Bordeaux I and CNRS both invested a full-time engineer to back up the activities of our team (we should also mention a young engineer hired under a two-year INRIA contract working with us). We use the resources we get from research grants and hire young software developers, working under the direction of our permanent staff, building short time small development teams working within MABioVis.

Technological transfer  By nature, a team working in bioinformatics must devote part of its time to transfer technology towards its collaborators, so one could ask why it should be mentioned as part of the strong points of the team. We want to emphasize here the efforts we put into developing results and tools of the most generic nature, of which the SIMBALS transverse project is an excellent example: algorithms used to compare genome sequences were used and adapted to the comparison and identification of melodies. The ANR SPANGEO project is another good example focusing on the visual analyzes of graphs: heuristics to compare and identify changes in graphs has also been applied successfully to video indexation (Fanny Chevalie’s Ph.D. thesis, supervised by two professors from MABioVis and “Image et Son”). We also were able to fund PH.D. thesis through the CIFRE program.

In the same vein, we want to make one statement. Our research places at the forefront of all applications of our results – this is quite natural since life sciences motivate most of the efforts and resources devoted to bioinformatics. But the new questions we are facing – large-scale sequencing, and users becoming more demanding in terms of systems simulation and easy access to growing knowledge – do form a challenge that must be addressed by computer scientists. The questions brought to us by biology and life sciences still remain more or less the same.

5.1.2 Weaknesses

A quite low critical mass  The critical mass of our team still is globally low, with an even more crying situation for the project “Algorithmique pour l’Analyse de Structures Biologiques” which actually gathers only one Ph.D. student, 2 assistant professors and 1 full-time professor.

Our project, because it is multidisciplinary, requires to bring many people of different horizons and sharing different knowledge and competencies – which is coherent with the goals we had at the origin of MABioVis four years ago.

We work with collaborators from various geographical sites, at a national and international levels, almost on a daily basis, as we mentioned above as one of our strong points. We often come across invitations to enter projects and contracts and we must respond favourably to those to insure our integration to the community and the evolution of our team. Grants capture a large part of our energy and human resources keeping them busy with administration. The time required to write grants, manage money and part-time human resources consume precious time we would rather invest doing or supervising research. All professors (rangs A) of our team are deeply involved in the administrative life of the university and various institutes.

A somehow unstable team seminar  The internal dynamics of our group is penalized because our staff is so minimal. Our members should be able to spend more time exchanging ideas and doing actual work together.

We do realize that we must install a periodic and stable working session gathering the whole team. For now, we have had weekly subgroups working quite well but gathering only a subset of all MABioVis members, probably favoured by the creation of the MAGNOME (comparative genomics) and GRAVITÉ (graph visualization) INRIA teams. It is mandatory that MABioVis
CHAPTER 5. TEAM MABIOVIS

has a forum of its own. This is even more important due to the fact that our members spread over different geographical sites around the campus (some of us have their daily office on University Bordeaux 2 campus to keep a close link with the CBiB).

5.1.3 Opportunities

A number of opportunities have been seized to help our research go forward, as assessed by all grants listed in our report. The creation of the INRIA teams MAGNOME and GRAVITÉ can also be seen as a testimony of our will to benefit from all opportunities our team is offered. The long-term stay of members from the “Structures biologiques” (P. Ferraro in Calgary and J. Allali in Vancouver) with the PIMS (Pacific Institute for the Mathematical Sciences), and invitations of foreign collaborators are efficient ways to stimulate our research. We also took advantage of our grants to fund Ph.D. thesis; as a matter of fact, funds seem much easier to find than high quality applicants. Collaboration with the industry, even within the context of public research grants (ANR DIVEONI, ANR FIVE, ANR TANGUY), have been used as a mean to enrich our research and enter new application domains.

We certainly need to identify in our local environment partners with competencies that complement our team, so we can rest on their geographical proximity and intensify local collaborations. The regional “Pôle Technologies de l’Information” certainly is an initiative we should try to benefit from. Colleagues from the IMB (Institut Mathématiques de Bordeaux) or from INRIA teams (EPI ALEA, for instance) work on issues that are relevant for us as they link to data mining. Other collaborations with plant biologists from Université de Bordeaux would certainly be welcomed. The European FET initiative VisMaster from the FP7 will certainly trigger interesting opportunities to disseminate our results and forge new collaborations and hopefully benefit from European funding.

5.1.4 Risks

The large scope of our research actually is a risk for our team, that because members work on different topics we lack unity – having a larger number of people makes it easier to find continuum. Low critical mass also impedes our working sessions to keep with a weekly rhythm. Another tangible risk is that the team cannot easily recover from the loss of a member (long term stay abroad, promotion, etc.) and keep a research theme alive.

Our software development projects also potentially contain risks. Because our software are being used by a growing community, our team must maintain high quality standards, and insure ascending compatibility. The Génolevures project, for one, puts us in a delicate situation; because human resources are obviously lacking for this project although quality is a must, our members work at tasks keeping them away from fundamental research and for which they do not possess all competencies. We are faced here with problems typical of multi-disciplinary projects involving different institutions, which moreover develops at a national level (GDR).

Moreover, engineers hired on the basis of short-term contracts support funded projects. There is a risk here of being seen by our collaborators as “service providers”. This danger is real and we must keep in mind that our software development task force is primarily devoted to support our fundamental research.

5.2 Project and prospects

We now describe our research projects for the coming years, organized around specific topics. Generally speaking, we clearly see how Antoine Petit’s priorities fit with our project. Research priorities such as “Stockage et fouille : codage, représentation, organisation”, or “Analyse : modèles et algorithmes propres aux données traitées” clearly fits with our research agenda. Systems biology

4. We link to this project through the INRIA teams GRAVITÉ (Bordeaux) and AVIZ (Paris, Jean-Daniel Fekete). See the URL www.vismaster.eu.
calls for the development of global approaches covering all levels, from concepts to down-to-earth implementations and technologies. Our methodology is motivated by questions raised by biologie and bioinformatics. But we also work at transferring concepts and methods in other fields and application domains, bringing a contribution to answer the challenges posed by the complexity of problems computer science must help to solve.

5.2.1 Comparative genomics

One of the strengths of our team is its close collaboration with experimental groups in the life sciences (as assessed by the links with life sciences laboratories or researchers listed in the summary part of our report), and over the next four years we will use these collaborations as a way to focus our effort in developing new methodological tools. Two main application challenges will provide this focus: realization of comprehensive comparative analyses of groups of yeast genomes, to elucidate their history and to identify mechanisms of molecular evolution; and development of hierarchical models of cell behavior linking genotype to phenotype.

Scaling up to next-generation sequencing is an important objective. This will require algorithmic improvements in our genome comparison techniques, and software improvements in our Magus annotation platform [Rs32]. The Magus genome annotation system integrates genome sequences and sequences features, unsupervised classifications, in silico analyses, and views of external data resources into a web-based collaborative platform for simultaneous annotation of related eukaryote genomes. In order to meet the increasing need for greater throughput, we must integrate new data-mining techniques for accelerating simultaneous annotation of homologous genes.

BioRica is a platform for hierarchical mixed-formalism modeling of complex systems in the life sciences, developed over the past four years as an internal tool [Ca80]. It explicitly addresses model reusability, repurposing and other engineering best practices that are necessary for sustainable, incremental development of comprehensive models incorporating individually validated components. Current modeling methodologies allow either global solutions (stoichiometric models and their kernel space solution), or local quantitative solutions (continuous models and precise kinetics solutions), under a steady-state hypothesis. We will move forward by building a methodology that can combine existing models with different granularities, and define time-dependent transitions from one steady-state to another.

5.2.2 Scalability is mandatory

As we have mentioned earlier, easier access to biological data offered by high throughput technology and lower sequencing or mass spectrometry costs, brings in potential difficulties. Not only our algorithms need to process increasing volume of information, but user demand could evolve towards being able to compare a wider range of genomes while taking multiple sources of information. We clearly see here that we will need to acquire competencies in data mining in a way or another.

This being said, scalability cannot be tackled by solely improving our algorithms – since the problems we handle are part of problem classes forbidding exact or exhaustive solutions. Developing interactive working environment taking user feedback into account is a direction we wish to explore further [Ou1] – by using or developing semi-supervised strategies, collaborating with teams in statistical learning (the INRIA team GRAVITÉ has had the occasion to experiment with these approaches with the help of a summer intern in 2008 [Au7, Ca19]).

Including the user in the loop is a priority, and is even more obvious when addressing simulation issues – this perspective is more than mandatory when tackling multiscale models of the cell. We sincerely feel we can bring a contribution to this overall goal of the international bioinformatics community [Ca80]: to build a model that would cover the whole cell behavior, from biochemical reactions up to tissues (!), allowing users to change the scale at which they observe the system.
5.2.3 Algorithmics and analysis of metabolic and metabolomic networks

We wish to reinforce our work on the study metabolic networks or networks of interacting biological entities based on proximities (derived from data semantics capturing biological meaning). These networks are most often modeled as graphs (where nodes correspond to entities such as metabolites and edges correspond to or enzymes, reactions or more generally to interactions). These graphs are then used as artefacts supporting interactive exploration and analysis of biological data, where they are typically embedded in a visualization environment easing exploration and hypothesis building and validation [Rs22, Rs8, Ca21, Au26, Ca63].

We need to improve our techniques for comparing biochemical networks [Au46, Rs6, Au10, Au20]. These structured comparisons improve over comparisons of flat sets of genes by imposing coherency constraints on the sets of genes that are present in a genome, since the genes must comprise a working whole that realizes the biochemical functions of the cell. Furthermore, these comparisons provide an important tool for inference of biochemical networks for new species, which is one way of moving from a ‘static’ genome to a dynamic functional model.

Graphs can moreover be used to study metabolic flows (or other types of networks). Modeling these flows using Petri nets, for instance, allow the use or experimentation with either continuous or stochastic, deterministic or not, modeling approaches.

The goals of this research is to identify biologically relevant structural properties of these networks, and develop a framework allowing the integration of heterogeneous data sources. We should aim at the integration of data at all levels, turning these networks into central modeling objects, where graphs are central conceptual and modeling tools.

This work has partly been explored, bringing data integration close to the study of biological networks and relying on existing collaborations with biologists (A. Blanchard et P. Sirand-Pugnet – UMR1090, INRA, Bordeaux 2) and a computer scientist (Matt DeJongh, Hope College, Michigan, USA).

These networks can also be studied as such, where modularity becomes a central notion guiding our research [Rs34, Ca34]. Most definitions of a “module” found in the literature either rest on topology induced from the mathematical objects themselves, or on functional aspects carried by data attributes. Efforts must be spent to reconcile these two points of view, correlating structural patterns and expression data for instance, and propose a definition that would appear meaningful from a functional perspective.

5.2.4 Dynamic systems and models

The increasing volume of data is not the sole challenge we are facing. We ought to take into account its intrinsic dynamic dimension. Our first results on genome reconstruction [Cs9] or simulation of cell aging [?] open promising avenues in this direction. We have also considered this issue in other contexts [Ca27, Au34].

The dynamic aspects of data goes with users asking for systems that can take dynamic aspects of data into account and deliver it through its visual representations and interactive features. This brings it load of questions and problems to solve:

- Studying the dynamics of biological processes is a central issue in the modeling and simulation of biological systems. The temporal dimension cannot however be addressed through naive approaches. Indeed, the time scale of these processes forbids any realistic simulation – simulating the $10^{14}$ biochemical reactions involved in *E. Coli* would require $4^{1024}$ seconds or 12 years on a single processor machine. These simulations must mandatorily rely on efficient algorithms, but more importantly on multiscale models themselves built on top of solid formalism.

- On the visualization side, animated drawing algorithms can certainly support the evolution and topological or attributes changes in biological networks. However, we plan to look at alternative and multiple views, letting a user directly monitor a simulation according to its output and directly impact its graphical representation. Why not also let a user experiment with clustering algorithms or (semi)supervised classification algorithms coupled with more
classical views of biological networks?

This being said, we plan to pursue our efforts exploiting the computing power of graphics hardware together with the possibilities offered by parallel implementations of the latest graphics cards. We already have begun this work on graph drawing side improving on force-directed algorithms [Ca6], and on the comparison and analysis of strongly structured biological data.

5.2.5 Multiscale tree graphs models for plants and self-similarity

Plant architecture has established as a central notion in botanics since the last thirty years. This notion captures two fundamental aspects of plants: the geometry of its component on the one hand, and their topology (their types and their adjacencies) on the other hand. A number of researchers have worked on the characterization of these two types of information for various plant species and developing conditions. However, a general methodology quantifying these fundamental aspects and closely linked to plant architecture is still lacking:

- characterizing plant shape irregularity (often greatly present, such as its overall leave distribution for instance)
- the similarity of its ramification at all levels

These two notions both rely on a multiscale description of the plant structure, referring to techniques and concepts from fractal theory. Our goals are to develop quantification methods of these two fractal properties, geometric and topological. These notions are essential to a better understanding and modeling of plant architecture. This research was recently boosted through a Ph.D. thesis.

We also plan to look at this work from a graph drawing perspectives, exploiting the results on plant architecture as compression techniques for graphs helping to improve the time complexity of drawing algorithms.

5.2.6 Text algorithmics

Sequences, just as dictionaries or text documents used by biologists can be seen as large-scale database equipped with elaborate data structures combined with astute search algorithms. Suffix trees are used in all these contexts to build index and improve searching, although they were originally designed to search genome sequences. We plan to extend suffix trees to search tree graphs of arc-annotated sequences to build index for RNA secondary structures. These new highly compact representations (DAGS, graphs, …) will undoubtedly open new directions to tackle the comparison and analysis of biological structures.

5.3 Evolution of our team

MABioVis deploys efforts at building a cross-disciplinary culture with all its collaborators as well as internally. Incidentally, our team is involved in a number of research grants at a national or international level implementing this multi-disciplinary strategy, just as in local projects such SIMBALS or the INRIA teams MAGNOME and GRAVITÉ.

Our priorities are to numerically reinforce the team to allow for a better supervision of Ph.D.’s and non-permanent staff (post-doc fellow, software engineers). The rate at which staff has increased must remain in the coming years. Grants and projects come with an administrative load that puts extra pressure on our members. The amount of energy devoted to control and reporting procedures penalizes our efforts to build a common vision and conduct scientific team activities, which turn out to be critical as our staff spreads over the whole Bordeaux campus (LaBRI/Université Bordeaux 1, CBib/Université Bordeaux 2, INRIA). We wish to establish ourselves as a reference task force in bioinformatics and visualization, serving the dissemination of results from our team, as well as from collaborators and foreign visitors. An important fraction of our members already contribute to the various committees and institutions we depend upon.
The Formal Methods group is well positioned for the next four year period. It is backed up with solid personal and scientific potential. It has world-class competence in a number of established, as well as emerging and attractive fields of research. The links created between its three themes in the last years are a precious asset. Institutionalised scientific and industrial collaborations are balanced and diversified. The working environment offered by the laboratory and campus is very good.

The present document provides an analysis of strengths and weaknesses of the group, followed by the discussion of opportunities and risks. For each issue a short analysis and a direction of development is proposed. This discussion is accompanied by a presentation of a number of scientific directions and goals.

6.1 Self Assessment

6.1.1 Forces.

The main strength of the group comes from the quality of its members and their scientific contributions (attested among others by IUF nominations, and high international visibility). The unusually wide spectrum of interests of the group allows it to follow a large variety of developments. Later in this document several research themes of high impact are presented. The group has the world-class expertise required to make important contributions in the outlined directions.

Substantial software projects also contribute to the strength of the team. The industrial base for the modeling and verification tool AltaRica is large and solid. The challenge is to capitalize this achievement. The technology developed in the syntax analyser GRAIL is at the forefront of development in this area. VISIDIA is a well established system that aims at an ambitious goal of formally proving correctness of distributed algorithms. A verification system FASTER is another well established system that is constantly growing, in large part due to theoretical results of the team. One should also mention a number of smaller but still important projects: AutoWrite, VisAA, XLFG.

In recent years the group has welcomed a considerable number of researchers from CNRS and INRIA. The arrivals of H. Gimbert (CR CNRS) and S. Salvati (CR INRIA) in 2007 has given a new impetus to the research on stochastic games and formal grammars, respectively. The recruitment of M. Rao in 2008 (CR CNRS) has allowed to solidify and expand the competences in graph decompositions. The arrival of D. Kuske in 2009 (DR CNRS, external hire) will bring a very much needed reinforcement in the domains of algorithms on infinite structures and of the theory of distributed systems. The group’s ambition is to continue to be an attractive place on the French and European research scene.

The group is very visible nationally and internationally as attested by the number of ANR and other projects, international collaborations, invited talks at conferences, organization of conferences, participation in programme committees. All these activities help spread the group’s results and boost its visibility. The national and international projects also guarantee that our work develops in directions that are relevant for large communities internationally. The clerical effort that is involved in these projects, and even more in the participation in program committees and the organisation of conferences has however a price, and our objective is to maintain a healthy balance between actual research work and other activities.

Local collaborations are an important part of research life of the group. The three scientific seminars form its backbone, and their quality will be maintained. Collaborations with other groups are well balanced and will develop naturally. The group is active in the construction of the Pole TI of the University, that may become one of the prominent axes of local cooperations in the coming years. The group’s participation in the development of a common laboratory with Thales is, and will continue to be significant.
6.1.2 Weaknesses.

One weak point of the group is the lack of rank A researchers in some of its areas of activity. Even though the group manages several important software projects, it has no professor with a strong expertise in conducting software projects. While at present this is compensated in various ways, in the future it may turn out that projects like AltaRica will need a dedicated rank A person. The theme IL has only one rank A, and it is in need of reinforcement.

Finding good PhD students is another challenge. The majority of the best local students choose professional curricula. In consequence, the group has difficulty to find enough well prepared candidates for the relatively theoretical and difficult subjects it proposes. This is one of the reasons why it actively seeks internship students from ENS and IITs (in India). The advantage of these two institutions is that one can have a quite precise judgement about the quality of a candidate. This is very important as an internship demands a substantial investment of time on the part of academic tutors. Of course other sources are not neglected, as witnessed by several theses in “co-tutelle”. In general, the number of habilitations and thesis in the group is growing, with the adequate proportion of thesis CIFRE.

A potential weakness of the group is the increasing involvement in administrative work, stemming from a more and more complex French university system. We hope that the LRU and the associated possibilities of balancing research, teaching and administration (including "délégations"), will benefit especially the younger members of the group.

6.1.3 Opportunities.

The opportunities for the incoming four year period are numerous and exciting. Most notably, several research subjects where the group has very strong competences are gaining in interest and importance. As an example, one can mention: graph decompositions, the efficient verification of MSOL properties, the algebraic and algorithmic theory of trees, infinite state systems, distributed systems, synthesis, categorial grammars in linguistics and links with proof theory. Later in the text these themes are treated in some detail.

Another very positive factor is the environment of the group: the laboratory and the campus. The group contributes to but also profits from the high visibility of the laboratory and the expertise of other groups. Scientific excellence is one of the important priorities of the University. The development of the “pôle TT” of Bordeaux and the common laboratory with Thales is singularly compatible with development of the group. Moreover, the INRIA Sud-Ouest center offers important resources, and know-how in project management.

International collaborations offer many noteworthy opportunities. The privileged contacts of the group with India should be continued and intensified. In 2009 the group will welcome its first student from IIT coming to do his PhD in Bordeaux. Cooperation with China has been initiated, bilateral cooperation projects are planned, and a first-rate post-doc from LIAMA (Beijing) will come for the year 2009/2010. Naturally, more “standard” French, European and American collaborations will not be neglected. The members of the group will be encouraged to organize such cooperations: ANR projects of various kinds, bilateral projects, European projects.

Several of the group’s research themes have been turned into long lasting collaborations with the industry. AltaRica is one of the prominent items on the list of cooperations in the project of the Thales scientific laboratory on the Bordeaux campus. Issues linked with low level code (security, obstruction, analysis) were developed in the CRYSCOE project and will continue, among others, in the BINCOA project. Other ANR projects as SPaCIFY and RIMEL involve strong industrial cooperations.

6.1.4 Risks.

For a group of this size the question of scientific coherence and interactions between different threads of development is always valid. A large number of different research themes is an important asset, but it can be also a handicap if care is not taken. The group plans to keep its current
configuration, that will be of course influenced by recruitments, departures and shifts of research interests. These should be anticipated and accompanied. For example, it will be important to handle the changes due to the retirement of B. Courcelle sometime during the next four years. The arrivals of D. Kuske, F. Mazoit, and M. Rao are important, but not final, steps in this transformation.

The group has one INRIA project: SIGNES. Its integration with the group is exemplary. The question of the creation of another INRIA project issued from the themes MV or LGL is always present. While there would be numerous advantages in having a second project in the group, it is not clear how to create one without unbalancing its scientific coherence. One should also note that the present form of functioning of INRIA projects is not well suited for an important part of the research done in the group.

6.2 Principal research directions

Structure of finite and infinite graphs There is a number of important open problems concerning “graph structure” such as graph decompositions, duality properties, automatic representations. These questions have algorithmic motivations, and the results contribute to the development of efficient algorithms, at least in terms of complexity theory. Generality is a big advantage of this approach: instead of concentrating on an algorithm for a particular property, one considers at once all properties expressible in some logic, most often first-order or monadic-second order logic, and their variants. The following are some examples of problems of high priority for the group: (i) decompositions and rank width for countable graphs, (ii) extension of clique width to relational structures. (iii) notions of automatic structures based on automata on objects such as ordinal or scattered words.

Another important objective is to extend to all relational structures the result of Courcelle and Oum on Seese’s conjecture. Although going from graphs to relational structures may seem routine, this problem already resisted several attempts, and new insights are surely necessary.

Logic and algorithms Polynomial time algorithms based on monadic second-order logic and graph decompositions of bounded width, as discussed above, are not implementable directly due to the huge associated constants. The size of these constants is linked with the size of the constructed finite automata, that in turn depends mainly on the quantifier depth of formulas. Two ways of overcoming this obstacle can be considered: either make the logic richer to limit quantification depth, or avoid the use of the general theorem and construct the algorithms directly. The first approach requires experimental work, the collaboration of an engineer-programmer over several years would be very helpful. Both approaches are developed in the group and will offer research questions for several years and several thesis. This work is urgent as the interest and competition are strong.

Automatic graphs and automatic structures are another field where the same problems arise. These are structures represented by languages of word or tree automata, and deciding their properties reduces to the construction of efficient algorithms on automata. We plan to get more insight into the question why the first-order theory of some automatic structures is algorithmically simple, while for others it is non-elementary.

The construction of labelling schemes is a related subject with big potential where logic and the structural theory of graphs meet. The first task is to study in terms of logical definability the fundamental concepts used for constructing labelling schemes. One should also study methods of composing labelling schemes for “composed” classes of graphs: for example planar compositions of graphs of bounded clique width, given that it is known how to deal with planar graphs and with graphs of bounded clique width separately. All the algorithmic results already obtained need to be extended to the dynamic case when the given graph can be “slightly” modified.
Algorithmic problems in monoids and groups  Algorithmic problems in group and monoid theory often involve discrete structures such as words, trees and diagrams, and the techniques of formal language theory and combinatorics on words can be used in that context. The group focuses on the combinatorial theory of free groups, and on the solvability of systems of equations.

We will for instance attack the problem of mixed orbits, to decide whether a given finitely generated subgroup of a free group $F$ contains an element of the orbit of a given word under the automorphism group of $F$. Of interest also is the possibility of an efficient solution for the problem of deciding whether two subgroups are in the same automorphic orbit (the problem is known to be solvable in exponential time). Finally, we will extend our recent results on the distribution of subgroups of free groups, to determine the frequency of certain properties of subgroups.

The solvability of systems of equations in a given group is closely related with the decidability of first-order logic in that group. Solvability is known to be decidable in free and partially commutative groups and monoids, and first-order theory is known to be decidable in free groups (Myasnikov, Kharlampovich). We will continue to explore the possibility of transfer theorems to extend these results: to show that the solvability of systems of equations in a group $G$, can be used to solve them in a more complex group constructed from $G$. We hope to unify and generalize recent results of this nature, on HNN extensions and amalgamated products. An ambitious goal would be to establish the decidability of first order theory in virtually free groups.

Algebra and automata  Automata theory is not only one of the corner stones of computer science, but also a discipline that witnessed a very substantial progress in the last years.

Recent results on profinite topologies and applications of Stone’s and Priestley’s duality theories provided a framework shift in the algebraic approach to the theory of recognizable languages. In particular, profinite monoids give a powerful tool for the solution of algorithmic problems on regular languages – such as interpolation problems: to separate, if possible, two given regular languages by a formula in some given logic. Such problems arise in several contexts, yet they are surprisingly poorly understood.

The other substantial recent development is an algebraic approach to the classification of recognizable tree languages. This subject has gained recent interest with questions on the expressiveness of standard navigational and query formalisms used in XML. The group will continue to develop an algebraic approach to tree languages, constructing new tools and characterisations of new languages, hopefully arriving at the landmark goal that is a decidable characterisation of the first-order logic.

Higher-order pushdown automata  Higher-order pushdown automata have also attracted more interest in recent years. This is due, among other factors, to the better understanding of their connections with such seemingly unrelated objects as higher-order programs, or with the Cauca hierarchy of graphs. One major open problem, the order hierarchy problem, consists in deciding the minimal order of a stack that is sufficient to accept a given language. Another problem asks whether the newly introduced operation of collapse (also called: panic) is really necessary.

One way to learn more about these automata is to study the sequences of numbers they can generate. Let us recall that the study of sequences generated by finite automata already has a long history (cf. the recent book by Allouche and Shallit). It is only natural to develop a similar theory for higher-order pushdown automata. A milestone open problem is to determine whether the equality of sequences generated by such automata is decidable.

Analysis of infinite state systems  The automatic verification of infinite state systems is at present limited by problems due to the explosion of the size of symbolic state representations. A long-term goal is to develop verification methods whose precision level can be modified on-the-fly, allowing an effortless passage between exact and approximated analysis.

One direction is to extend the applicability of acceleration methods to the computation of fixpoint functions in abstract domains. Another, is to extend techniques of abstraction by predicates, to predicates with parameters. Finding an effective construction of this kind poses many problems,
coming from the fact that the obtained abstract system is infinite. As mentioned above, the analysis
of infinite state systems requires very efficient symbolic representations. The development of a
mixed representation based on both NDD and Presburger arithmetic will be pursued, following
promising results obtained with the TAPAS suite.

**Distributed systems** Distributed systems are omnipresent: modern microprocessors, distrib-
uted databases, web services. The goal is to develop models of concurrency as well as methods
allowing to synthesise, or control, distributed systems as well as to verify them automatically. The
group has strong expertise in this area, in particular for “closed” models based on communication
by buffers or synchronization. More recently the “open” models where such systems communicate
with an environment have also been studied.

The first goal is to develop a better understanding of models of concurrency such as event
structures. At present, a systematic study of the links between distributed synthesis and the
decidability of certain logics over event structures seems unavoidable. The second objective is
to work on concrete frameworks for distributed synthesis. In particular, to extend the theory
of control proposed by Ramadge and Wonham from words to Mazurkiewicz traces. Moreover,
the group investigates a model based on graph rewriting systems to study and prove distributed
algorithms. This work has been structured as a project called VISIDIA that will be detailed in
Section 8.1.

**Analysis of low level code** The issue of security of byte code is important in several highly
critical applications: protection against timing attacks in cryptography, timing constraints in
control programs. The group develops a software architecture dedicated to software protection, be
it under reverse engineering or side-channel attacks. The principle is to execute the critical parts
of a program on a remote tamper-resistant trusted device, for example a smart card. A system
under development will cut a code into small pieces, and transmit these pieces of code to the
trusted device one by one at run-time. This task requires the development of a particular, byte
code oriented, kind of static analysis. In a broader perspective the objective is to develop methods
and tools for a byte code oriented variant of static analysis. At this level it is unavoidable to
consider issues such as low level memory, or dynamic jumps. This poses completely new challenges
that are not addressed by the techniques of analysis used for high level programs.

**AltaRica** AltaRica is a formalism developed in LaBRI for modelling critical systems. It is based
on a high level language for describing modular and hierarchical systems. An obvious objective
of development is to be able to handle even bigger models. In the period of reference, the system
was completely rewritten in order to use symbolic BDD representation of state space. The next
challenge is to improve the constraint solving engine. With the same objective in mind, tools
for the hierarchical analysis of systems will be developed. Specific features of AltaRica make it
impossible to simply use solutions proposed in the literature. A stochastic extension of the system
is also planned. This poses interesting scientific challenges and answers demands from the industry.
Finally, adding some synthesis modules can be very interesting, if only to synthesise automatically
less relevant parts of the system.

**Towards a unified view of formal grammar for natural language syntax** Two kinds of
formalisms may be used for describing natural language syntax: rewriting systems (of strings or
of trees) and categorial grammars based on proof systems. They offer complementary advantages:
for the first kind, numerous efficient algorithms exists, while the second kind of analyses can be
automatically turned into semantic representations. The team already obtained results connecting
the two kinds of grammars; these results often make use of de Groote’s abstract categorial gram-
mars that rely on the encoding of trees within simply typed lambda calculus. We are willing to
establish more general results linking the two kinds of formalisms. Three outcomes are expected:
algorithms for computing semantic representation associated with analyses in rich syntactic for-
malisms like TAG or MG; a description of a grammar as a set of constraints as done in Monadic
Second Order logic or in model theoretic syntax; general parsing algorithms making use of Datalog query optimisation along the lines identified by Kanazawa.

**Acquisition algorithms for syntax and parts of semantics** In order to make the parsers developed by the team work, we need grammars, and we therefore must develop automated acquisition from annotated corpora. This process often leads to grammars with a huge number of categories or rules per word. We then consider only a fixed number of most likely sequences of categories. The success of such a technique of analysis is purely empirical; the ratio of correct analysis, or the percentage of words with a correct category assignment etc. For other data types (genome, XML trees) one often uses the techniques of learning in the limit (Gold, Angluin, Valliant, . . .) where convergence can be proved for some classes of grammars, in particular for restricted classes of categorial grammars (several people of the team have already explored such issues). A promising direction is the combination of those two types of acquisition for being at the same efficient and convergent. One could define the categories of grammatical words (propositions, pronouns, . . .) and only infer the categories of words from the open classes like nouns, adjectives, verbs or adverbs. Notice that by now one can also think of acquiring some aspects of semantics like argument structure. This is appealing for a group studying the relations between syntax and semantics. From a practical point of view, acquisition was first used for Dutch data and is now developed for French, using the Le Monde corpus and the regional and historical corpus form the médiathèque of Pau.

**New direction for the modelling of natural language semantics** Up to now, the automated semantic analysis of natural language is far behind its syntactic analysis. We have already defined algorithms extracting meanings, viewed as logical formulae, out of parsed sentences. We intend to extend this process to larger units (discourse and dialog), and to deepen the analysis by refining the logical formulae. In particular, those associated with the words (lexical semantics). Both research directions will be developed within a single framework, logic (classical, intuitionistic, linear, according to the issues) in such a way that words, sentences, paragraphs all have the same kind of structure. This logical approach (also studied by Asher, Kamp, Pustejovsky among others) can easily compose units into bigger ones.

**Grail’s future** This parser has two important characteristics: it produces semantic representations, and it works fine on large corpora with an excellent precision, especially with respect to the number of correct complete analyses of sentences. Grail is presently being developed for French as well. Grail allows us to test immediately most of the theoretical advances of the team: richer semantics (some experiments already started with lambda-DRT), automated acquisition of syntax and also of semantics. Thus we can test our results on real linguistic data. The study of the French historical and regional annotated corpus of the médiathèque of Pau, which requires an acquisition process, syntactic, and semantic, analyses is a good challenge for Grail.
Team SATANAS

7.1 Self Assessment

Strong points
- The Team has a good scientific visibility and has on-going partnerships with many world-renowned scientists in the domain.
- The Team follows a successful scientific approach in validating its theoretical results by integrating them in distributed software libraries effectively used at the national and international levels. Here again, some of these pieces of software have gathered a strong international visibility.
- The Team constantly takes care of transferring its results to real life towards the industrial world of Research and Development, hence the large number of partnerships and contracts involving industrials. The use of software tools as a mean to go beyond a hard threshold in solving complex applicative scientific challenges is a central objective in the scientific policy of the team.
- The Team obtained results through the combining of everyone’s skills and such results largely exceed the sum of what each team member would obtain individually.
- The Team has a renowned and acknowledged "view" in the HPC domain, such that several of its members a regularly invited to give talks (Conferences, Ministries, CEA, GENCI) or to carry out expertise missions.
- The Team is dynamic and attractive (many new members recruited over the last 4 years) and is able to reorientate or open its research axis when necessary. Besides, the Team acknowledges with some satisfaction the positive future of its Ph.D students, most of them now holding a position in their initial competence domain.
- The Team actively worked at setting up a Parallel Computing scholarship path common to the Computer Science Department of ENSEIRB (3rd year’s PRCD track) and University Bordeaux 1 (Master SR) to train young people with the new challenges opened by numerical simulations and large scale computing.

Weaknesses
- The Team must improve the number of its international journal publications. Most publications currently target international symposiums.
- The number of HDR should increase in the following years.
- The Team should reinforce its direct implication in the normalizing consortiums of some languages and tools for parallelism (for instance the MPI Forum or the OpenMP ARB).
- The work in progress may suffer from some fragility related to the difficulty of finding and enrolling new development engineers in the highly technical topics involved.

Opportunities
- The Team’s current research topics are perfectly on par with the directions followed by the scientific community which puts strong efforts on numerical simulation and high-performance computing, at the level of hardware platforms (peta and exaflops computers), at the computer use level (from the runtime systems to the algorithmic libraries and languages), and at the application level with academic or industrial partner (CEA, EDF, TOTAL, partners of the AESE pole). One may also cite the national initiatives related to GENCI and the international initiatives in Europe (PRACE, FP7) or in the USA (IESP).

Risks
- The volatility of HPC technologies and the strong evolution (both in terms of scalability and of technological improvements) make it difficult to clearly envision the upcoming challenges beyond a five-year horizon.
– This volatility requires to work in tight relationship with hardware vendors that hold the technology and the keys of their use (for instance NVIDIA in the case of GPGPU).
– Hiring highly skilled people in the domain is difficult due to the lack of courses on the topics of high performance and parallel computing in French universities and schools. The situation is even worse in the case of HPC for large scientific challenges that require a cross-domain training in both high performance computing and applied mathematics.
– The lack of HDR in the Team results in a low proportion of senior researchers to lead junior researchers. This point should be corrected quickly.

7.2 Project and prospects

Runtime

New Execution models

Structure to Conquer The complexity of contemporary parallel architectures is such that in order to exploit them efficiently it becomes necessary not only to generate massive parallelism to feed the numerous available cores, but also to generate structured parallelism to optimize the mapping of application tasks onto the hierarchical topology of the machine.

In that context, the OpenMP language stood out to be a nice support for automatically extracting the parallelism structure out of an application. Many tracks remain to explore in terms of compiling OpenMP applications for multicore architectures. Among those tracks, we intend to study more specifically the principle of gathering information at compile time as a means to better characterize those pieces of data shared between threads. We also intend to study the gathering of machines’ hardware counters to dynamically refine the location of processings and data, the memory management issues on NUMA architectures and on the expressiveness of scheduling strategies on that kind of machines.

Scheduling on Hierarchical Architectures The rise of specialized accelerators such as the GPGPUs or the upcoming coprocessors of the Intel Larrabee processor introduce a dreadful heterogeneity at the heart of parallel machines in terms of programmability, memory management, degrees of parallelism or computation granularity. For now, the multiple research efforts in the domain mainly focus on offloading some parts of a program onto the accelerators, that are only seen as co-processors.

We believe that the real challenge is to consider all the processing units as a whole and to implement task scheduling strategies inside runtime system in such way as to allow the dynamic computation of the optimal task assignment. Our most promising approach is to design runtime systems able to schedule graphs of sub-divisible tasks with multiple implementations and to optimize the data moves and redistributions using cost models and directives from the application or extracted by a compiler.

Our exploring works towards that direction (on the starPU runtime system) have already yield interesting results. These works will continue in the context of european project PEPPHER (strep FP7 ICT-2009.3.6.a).

Parallel Communication Engines The increase of the number of cores inside cluster nodes opens new problems with communications and I/O. Contention may appear when multiple cores concurrently call the underlying library. It is therefore important to decrease the traffic of requests by applying global optimizations at the node scale. Moreover, some processing occurring inside the libraries could itself be parallelized.

We plan to focus on request schedulers able to use several computing cores simultaneously to implement aggressive optimizations for both communications and I/O. These works will be conducted in the context of a bilateral collaboration with University of Tokyo (PAI Sakura).
Programming Model Composability

Hybrid Programming on Multicore Clusters  The progressive deepening of computing machine hierarchies breaks the "flat" programming models (such as pure MPI programs) that do not offer abstractions to take into account the underlying architecture topology. This is why hybrid approaches are now proposed – most often associating MPI and OpenMP – in order to better match the parallelism of the application with the characteristics of the hardware.

Runtime systems that would allow running such hybrid applications in a flexible manner remain to be designed (in particular, one should be able to choose the MPI/OpenMP ratio in accordance with the number of cores). There is also a similar lack of tools suited for profiling and debugging. This study will be conducted within the context of Associate Laboratory INRIA-UIUC.

Towards a True Composability of Parallel Libraries  An increasing number of parallel applications will be composite, in that they will use libraries parallelized with different tools – computation kernels typically (BLAS, FFT). Despite the necessity to efficiently support that kind of applications, it is currently very difficult to have several codes using differing multithread engines (OpenMP, Intel TBB, Pthreads) coexist in a same execution context.

For such a coexistence to be possible, one must ensure that the implementations of those engines all are built on the same runtime system. For that coexistence to be efficient, one must also address several issue : How to schedule the multiple tasks on the machine cores so that the combined parallelism works well? How to dynamically choose the granularity of computations? How to extend these approaches towards exploiting heterogeneous architectures (use of CUBLAS + MKL, for instance)?

Improving Interaction with Hardware

Virtualize to Better Optimize  The relationship of a runtime system, which is by definition specialized, with the underlying operating system is often a compromise between performance and portability. This is particularly obvious in the case of thread scheduling and memory management where truly optimal performances would require a total control of the runtime system.

The hardware improvements of recent architectures in terms of virtualization support made it possible to use the technology of virtual machines for HPC, in particular to more easily implement fault tolerance. We plan to use such technology with the intent to customize the operating system and to smooth the runtime systems boundaries. This work has already started with a CEA/DAM collaboration.

Bacchus

In the context of the INRIA team Bacchus, our works aim at developing parallel algorithms and tools for the solving of very large linear and non-linear systems modeling application problems described by systems of equations of an essentially hyperbolic nature: fluid mechanics with Navier-Stokes equations, resistive or non-resistive magnetohydrodynamics, linear or non-linear wave propagation, etc.

Parallel dynamic repartitioning and remeshing  The efficient execution of instationary numerical simulations basing on high-order numerical schemes requires to compute accurately the evolutions of the physical phenomena taking place in some portions of the mesh, while not wasting compute power. Since the modeled phenomena evolve with time and in space, meshes have to be dynamically adapted so as to refine areas where physical quantities evolve quickly (high gradients) and coarsen uninteresting areas.

Meshing and dynamic remeshing are arduous problems, which are now adequately solved sequentially. Basing on the expertise we acquired in parallel graph algorithms, we wish to devise and implement a parallel middleware for the handling of distributed meshes, that would relieve
users from the burden of dynamically remeshing and redistributing their meshes. In order to be-
nefit from the services offered by this middleware, they would only have to provide a sequential
dynamic remeshing routine suitable for their problem. This middleware will base on parallel repart-
titioning routines which minimize data redistribution. Our goal is to run efficiently instationary
simulations of several hundred million nodes on more than a thousand processors.

**Sparse linear solvers** We want to continue our previous work around the resolution of large
sparse linear systems. We will focus on the following topics:

**Adaptation to new architectures** The new parallel clusters are more and more often
using multi-core architectures. Our recent work has enable to develop an MPI/thread version of
**PaSTiX**. we intend to further develop this approach by also studying the use of GPUs for our
linear solvers.

**Numerical pretraitement** One of our objectives is to develop a parallel algorithm for
constructing a permutation of the unknowns that maximizes both the symmetry of the matrix
and the weight of the diagonal.

**«Out-of-Core» methods** Such features are developed to push forward the memory limit
of sparse linear solvers. New problems linked to the scheduling and the management of the com-
putational tasks may arise (processors may be slowed down by I/O operations).

**Multilevel and Hybrid methods** In our previous works, we have developed some robust
hybrid direct/iterative solvers. These methods, implemented in the **HIPS** software, use a domain
decomposition of the matrix adjacency graph. The interface between domains is decomposed into
levels. We have developed some parallel algorithms that allow to build recursively an incomplete
factorization of the Schur complements corresponding to each level. In each level one can choose
to use an **ILUT** factorization or a dense block factorization. We want to go further on this topic :
in particular, we want to use the level decomposition to obtain the optimal scalability that can be
achieved by multigrids method under certain conditions.

**Incomplete block factorization** In previous works, we have studied and developed an
incomplete block factorization based on the level-of-fill (**ILU(k)**). This kind of factorization keeps
matrix entries based on the matrix adjacency graph. Yet, we want to improve the efficiency of this
approach and study the impact of the unknown reordering and of the numerical pre-treatments
that can be done on the initial system. We also want to study block **ILUT** factorization. In this
case, the criterion to keep an entry in the matrix is based on its magnitude (absolute value for a
real). To our knowledge this problem has not yet received much attention from the community.

**Application domains** The applications we foresee include several different physical models
such as MagnetoHydroDynamics (in collaboration with the PUMAS team headed by H. Guillard
as well as with a group in CEA Cadarache headed by G. Huysmans) on the numerical simulation
of the ELM instabilities of **ITER**. Mesh refinement technique is particularly suited to the ELM
problem where the perturbations become very localized. In the context of high performance parallel
computing, we have to study the main difficulties encountered in these methods such as : dynamic
load-balancing, management of the hierarchy of the grids which themselves are distributed, criteria
for refinement/unrefinement, management of the versatility of the parallel solvers.

In the context of non linear algebra solvers, our objective is the development of some new tools
adapted to iterative solution strategies driven by numerical schemes for essentially hyperbolic
problems. The idea here is to reuse as much as possible the preconditioner from the last Newton
step(s) to devise the preconditioner of the current step. As a first step, we want to develop a first
prototype based on **FluidBox** from BACCHUS team and **JOREK** from CEA-Cadarache.
HiePACS

Considering the context described in the introduction, the objectives of the HiPACS project (High-End Parallel Algorithms for Challenging Numerical Simulations) is to develop a set of algorithmic and applicative tools aiming at the efficient resolution of high-end simulations in the context of multi-scale applications. Our approach is multidisciplinary consisting in both applied mathematics (related to the numerical behavior of our algorithms) and computer science (around high-performance computing to design robust and scalable tools targeting next generation platforms). To be more precise, our target platform is composed by hundreds of thousands of computational cores with a hierarchical layout requiring multi-level schemes to be efficiently exploited. Note that for this objective to be achieved, the cost of data accesses cannot be neglected.

HiePACS is composed by INRIA researchers, members of the PRES of Bordeaux previously in the ScAlApplix group, and researchers from the ALGO team from CERFACS led by I. Duff. In its INRIA project-team configuration in common with PRES of Bordeaux and LaBRI, HiePACS will be a research group of the joint lab INRIA – CERFACS (which will be created at the end of the year). This will allow us to have strong relationships with industrial partners from CERFACS. The research topics studied by the HiePACS are briefly described below.

**HPC for next-generation platforms** This transversal topic concerns all the methodological aspects required to be able to scale on future high-performance computers which will be massively multi-processors where each processor has a large number of cores. Thus, this topic is centered around the analysis and the design of algorithms exploiting parallelism in a hierarchical and multi-level way that corresponds to the memory hierarchy of the underlying platform. To be more precise, problems like the scheduling of a large set of tasks on a platform while taking into account memory affinities between tasks and the data they deal with will be considered. In addition, problems related to the design of more adapted data structures to manage parallelism or the management of communication and high-performance I/O operations will be considered. This last point is crucial since our target applications will deal with huge amounts of data. Finally, we plan to introduce in our studies some new concepts like fault tolerance (which will be a major issue when the number of resources will increase). It is important to note, that all the research directions presented above will lead to collaborations with the RUNTIME part of the team.

**High-performance algorithms**

**Solvers for large linear systems** Concerning the resolution of large sparse linear systems, we will mainly focus on Krylov methods which have a large potential for scalability and a good capacity to be adapted to next generation computers. We will work on hybrid methods combining iterative and direct methods using Schur complements based on an algebraic decomposition like the additive Schwarz method. A prototype has been developed in the context the Solstice ANR project. It has successfully and efficiently solved a problem of 70 Millions of unknowns on near to 2000 processors of Blue Geen system. A highly scalable version using GPU will be developed in the context of a Phd thesis and an engineer position founded by INRIA to produce both algorithmic results and a robust software which will be available to the community.

Another hierarchical method combining a direct solver for computation on coarse grained grids and the multigrid method for finer grained levels is studied in the context of a PhD thesis with CEA-CESTA. The objective is to be able to solve the Maxwell equation on irregular meshes (the size of the target problems will be of the order of the billion).

Finally, other studies using Krylov methods and related to the resolution of multiple right-hand side linear systems or eigenvalues solvers will continue to be addressed. This work will be done in collaboration with the NACHOS and MAGIQUE3D INRIA project-teams.

**N-body problems ans fast Multipole methods** Concerning Fast multipole methods, we will continue our previous work by improving calculation efficiency on forthcoming processors.
(Graphics Processing Units, massively multicore,...) and by adapting our method for new potential arising from material physics such as in dislocation framework. Moreover we plan to use this approach for problem arising from boundary elements method to perform matrix-vector product in a linear time in the framework of Krylov subspace methods.

**Efficient Algorithmics for Code Coupling in Complex Simulations** The complex simulations we focus on are typically multiscale or multiphysics simulations, based on the coupling of two parallel codes, with very different natures. The performance of the coupled codes depends on how the data are well distributed on the processors. Generally, the data distributions of each code are built independently from each other to obtain the best load-balancing. But once the codes are coupled, the naive use of these decompositions can lead to important imbalance in particular, when we have an overlap zone between the different models. Therefore, the modeling of the coupling itself is crucial to improve the performance and to ensure a good scalability of the coupled codes. The goal here is to find the best data distribution for the whole coupled code and not only for each standalone code. The main idea is to use an hypergraph model as the one provided by the ZOLTAN toolkit and to take into account more information in the coupling than the classical one used by graph partitionner. Recent works on hypergraph partitioning with fixed vertices have demonstrated their effectiveness for dynamic load balancing of adaptative simulations. As this load balancing problem is quite close to the redistribution one, we expect to provide new redistribution algorithm using similar strategies. These techniques should be applied in material physics to improve the load-balancing of a crack propagation simulation, based on the coupling of a molecular dynamic code and an elasticity code. Moreover, we hope to apply our results in the context of M-by-N computational steering (coupling parallel simulation with parallel visualization tools), where the simulation distribution is fixed while the visualization distribution can be freely adapted.

**Application domains** Many applications in material physics need to couple several models like quantum mechanic and molecular mechanic models, or molecular and mesoscopic or continuum models. These couplings allow scientists to treat larger solids or molecules in their environment. Many of macroscopic phenomena in science depend on phenomena at smaller scales. Full simulations at the finest level are not computationally feasible in the whole material. Most of the time, the finest level is only necessary where the phenomenon of interest occurs; for example in a crack propagation simulation, far from the tip, we have a macroscopic behavior of the material and then we can use a coarser model. The idea is to limit the more expensive level simulation to a subset of the domain and to combine it with a macroscopic level. This implies that atomistic simulations must be speeded up by several orders of magnitude.

At the beginning, we will focus on two applications; the first one concerns the computation of optical spectra of molecules or solids in their environment. In the second application, we will develop faster algorithms to obtain a better understanding of the metal plasticity, phenomenon governing by dislocation behavior. Moreover, we will focus on the improvement of the algorithms and the methods to build faster and more accurate simulations on modern massively parallel architectures.

Finally, our research works about high performance linear algebra will continue in the context of industrial collaborations for specific problems; this point concerns BRGM, TOTAL and EDF.
8.1 Project and prospects for VISIDIA

8.1.1 Self assessment

The VISIDIA project aims to develop a set of tools and a unified environment for encoding, studying and teaching various aspects of distributed algorithms. Members of this project belong to distinct teams of the laboratory and have in common the investigation of a model for distributed algorithms.

**Strengths**: The project is based on a model defined as graph rewriting, which is simple and intuitive, and which presents a good level of abstraction for a large family of distributed algorithms. One of the strengths of the project is the synergy between various skills of the laboratory to develop a unified environment from modeling distributed algorithms by rewriting graphs, to their correctness and until software implementation. We obtained new results in the field of distributed algorithms and developed new tools for teaching distributed algorithms and distributed systems. The project is visible in the laboratory, and we have collaborations with other teams (as the team LSR). National collaborations are developed not only with researchers in distributed algorithms but also with other research groups interested in this model. Examples include the joint work with the LORIA, formalized by the ANR RIMEL. At international level, we have established collaborations with South Africa (university of Pretoria), Germany (university of Oldenburg) and we developed advanced contacts with the Laboratory LIAMA (China).

**Weaknesses**: We are not so many assigned to the VISIDIA project regarding the amount of work. This project has an anchor in theoretical teams and has little capacity for developing software tools. In fact, producing a useful software that can be distributed requires the assignment of a permanent engineer to this project.

**Opportunities**: Computer technology relies more and more on information that is complex and massively distributed over networks (like the Internet). The distributed algorithms that operate on these networks are difficult to design, especially when, with only information about the neighbours, processes must take part in a global behaviour. To cope with this complexity of distributed algorithms, it is necessary to build tools that help to understand their fundamental mechanisms. This research project spans from the most theoretic question to the most pragmatic ones. Thus, it brings together experts in a variety of fundamental domains as well as more applied domains. This gives us the opportunity to work with many other research groups. For instance, we are regularly approached by research groups in formal proofs for collaborations and projects. We also believe that the tools developed in the VISIDIA project will be used for academic purposes.

**Risks**: The concept of internal project at the laboratory is not yet well defined. Traditional structuring of the laboratory into teams and themes may let the project in a marginal situation, and thus the project may not be a priority.

8.1.2 Project and prospects

Within the VISIDIA project, we continue to enrich and investigate the model of graph relabeling systems for expressing, studying and implementing distributed algorithms. This encoding has yielded new results in the field distributed algorithms and a modular implementation of a large class of distributed algorithms. We will continue to develop the following topics:

**Formal proofs of distributed algorithms**

We are interested recently in using the encoding of distributed algorithms by rewriting graphs for developing a complete environment for certifying distributed algorithms. Proving that a distributed algorithm conforms to a specification could be a long and a complex task, and may contain
many calculations that could be prone to errors or omissions if it is made by hand. Based on
the theoretical work developed mainly around the project VISIDIA, we formalize the behavior of
various classes of local computations. The theorems proved by using this formalization should al-
low not only to facilitate proofs of specific algorithms, but also to show simulation results between
various classes of algorithms. We have formalized in Coq the class of algorithms of type LC0. An
algorithm of this class is encoded by rules that rewrite labels of a pair of adjacent vertices.

The progress of this work has been distributed as a Coq contribution. The developments on
the class LC0 show the feasibility of our approach.

Future work includes several directions:

- Extension of the Coq formalization to other classes of synchronizations (including LC1, LC2)
  and investigation of main paradigms of distributed algorithms such as termination detection.
- Definition of procedures facilitating the (semi-)automated proofs of algorithms to make them
  useful by non specialists in interactive demonstrations.
- Integration of our work in the Why tool.
- Interface development with VISIDIA (Coq-VISIDIA, RODIN-VISIDIA). This environment
  confidence-implementation-visualization of algorithms will assist in the development of new
  algorithms, and will serve to teaching.
- Technical proofs of self-stabilizing algorithms, composition of proved algorithms

To carry out these activities, we will collaborate with other French teams: Cedric Lab. (auto-
mation of proofs, self-stabilizing algorithms), INRIA Project-Team Proval (integration with Why),
INRIA Sophia Antipolis, ENS-Lyon (self-stabilizing algorithms, composition). We also plan to de-
vvelop proof techniques for probabilistic algorithms. These latters are also developed in the project
VISIDIA and will be explained in the next section.

Implementation of randomized algorithms

A randomized algorithm is an algorithm where we use randomization: nodes in the system
generate a random value and according to this value, make decisions or do some actions.

The implementation of a distributed algorithm described by a set of rewriting rules needs the
design and the analysis of randomized algorithms that can synchronize two neighbouring vertices
or a vertex with all its neighbours. These algorithms are implemented in VISIDIA as a building
block for any distributed algorithm which needs this kind of synchronization.

In the last years, we studied a set of distributed algorithms which solves classical problems
from the graph theory. In particular, we obtained the following results:

- given a graph $G = (V,E)$ which models a network of processors, an independent set is a subset $I$
of $V$ such that, for any $u,v \in I$, $\{u,v\} \not\in E$. An independent set $I$ is maximal if for any $v \in V\setminus I$,
  there exists $u \in I$ such that $\{u,v\} \in E$. We design a distribute randomized algorithm which
  uses messages of size 1 bit to compute a maximal independent set for arbitrary graphs. This
  algorithm runs in $O(\log n)$ on average and with high probability. That is, it is an optimal
  algorithm for the MIS problem.
- given a graph $G = (V,E)$ which models a network of processors, a proper colouring of $G$
is a function $c$ which gives a colour $c(v)$ to each vertex $v$ such that if $\{u,v\} \in E$, then
  $c(u) \neq c(v)$. We also design and analyze an optimal distributed randomized algorithm which
  achieves a proper colouring in $O(\log n)$ on average and with high probability. This algorithm
  uses messages of size 1 bit.

VISIDIA allows the user to visualize and test distributed algorithms. When we use distribute randomized algorithms, the running time is a random variable (r.v.). VISIDIA allows the user to run the algorithm many times and to display the observed values of this r.v. Hence, one can observe the graphs and can make some conjectures on the r.v.’s distribution. It is then clear that VISIDIA can be used to teach distributed randomized algorithms, but also to analyze the empirical values of their running time.
Local computations and dynamic networks

The distributed algorithms studied in the project are supposed to run on static networks. Extending these algorithms to dynamic networks, or showing their limitations became necessary due to the emerging practical applications. Dynamic networks may correspond to a deletion / addition of vertices or edges. The mobile agent approach provides a way to separate the computation from the topology of the network, and therefore allows to develop algorithms resistant to changes of network topology. The objective of this work is to develop techniques specific to distributed algorithms on dynamic graphs, both theoretical and practical aspects. We will study how to combine transformations of components of graphs and rewritings of labels to encode distributed algorithms in dynamic environments.

Another perspective of this work deals with the development of population protocols using local computations. These protocols are useful for mobile sensor networks, which have no fixed underlying topology. The sensors move and communicate with their neighbors (through wireless medium). Angluin et al. have been interested in this type of networks and studied its computational power. In theoretical terms, one could study the distributed algorithms by considering various mobility models. We have already begun to experience some algorithms using the platform VISIDIA. Our goal is to extend our work on distributed algorithms defined by local computations to such networks.

8.2 Project and prospects for SIMBALS

8.2.1 Self assessment

Drawing from both skills of the Image and Sound and the MaBioVis teams, a couple years ago, the Simbals project gave rise in the LaBRI to a new research theme for the analysis of symbolic music. This collaboration between two teams, apparently unrelated, is a strong asset that has supported the development of innovative and powerful methods, not only in the context of computer music but also in bioinformatics. Based on existing collaborations, this double membership has also allowed building a new collaboration network at both national and international levels.

As it is often the case for emerging research topics, only a few number of permanent staffs is involved in Simbals project (only 3 associate professors for 1 postdoc, 2 PhD candidates, and a couple of master students). The low group size is more accentuated by the imbalance between the permanent members belonging to the MaBioVis team and those belonging to the Image and Sound team. Moreover, the multidisciplinary nature of the project strongly orients the research activity toward the applications. Thus, project members develop prototypical software (recognized by the MIR community through different prices) to justify theoretical developments. These software are mostly supported and maintained by associate professors whose development is not the core business. Although a software developer from LaBRI-Transfer has been integrated to the project during the last 6 months, there is still a significant lack for the dynamic of our project.

Many companies involved in the distribution of digital music aim at developing new methods for navigating into musical databases. The industrial context is thus very favorable for the transfer of technology. In the future, new collaborations, coming either from industrial groups or research labs, should therefore support the development and the visibility of the project.

In the mean time, as we have stated previously, many collaborations, mainly on the algorithmic aspects of the project, have been carried out. These collaborations (King’s College, University of Calgary) and new ones should be reinforced, expanded or emerged, particularly in the domain of computational music. There will help alleviate the problems of under-staffing either through the exchange of PhD students or post-doctoral fellows or with the mobility of permanent staff.

Finally, the dynamic, the multidisciplinarity and the impact in terms of transfer of technology to the industry, allow us to consider in the future this emerging theme as a project supported by INRIA.

The under-staffing, if is sustained, may prevent the development of efficient collaborations. It does not allow an optimal management of students or any other non-permanent staff of the team.
Furthermore, research in Music Information Retrieval domains becoming extremely competitive, in order to develop methods that will be both innovative and related to the applications, musical knowledge must be integrated to these methods. Therefore, if the gap between the number of staff members coming from the MaBioVis Team and Sound theme is maintained, we could eventually lose our leadership in the MIR community.

8.2.2 Project and prospects

The number of digital music documents available on the Internet is highly increasing. However the only available tools are based on browsing from text queries (name of songs, performers). New methods for retrieving or classifying musical pieces have to be proposed. One of the key issues is then the development of algorithms for estimating the musical similarity between audio data.

Measuring similarity between sequences is a well-known problem in computer science having applications in many fields such as computational biology or text processing. However, musical sequences are characterized by specific properties and then imply to take into account areas such as sound perception and music theory in the development of efficient and accurate algorithms.

Since 2006, a new cross-domain project in the LaBRI propose to develop systems for estimating musical similarity by considering several properties, mainly from music theory, like for example melody, harmony, key, rhythm, etc. This multidisciplinary research team gathers both Image and Sound and MaBioVis team staff members.

First results on symbolic melodic similarity relying on the adaptation of sequence alignment algorithms have been already obtained. They will be used as a basis to investigate new research way following two main lines: music information analysis from audio signals and application to large databases.

Modeling Musical Information

First results concern symbolic music (representation of notes as a musical score). Considering audio databases requires the development of new algorithms for analyzing musical properties from audio signals, taking into account harmony, rhythm or timbre.

Methods for analyzing tonal information, like key or chord progression, are under development. These researches rely on the estimation of accurate features describing tonal context and on a system based on musical rules, and propose an analysis based on dynamic programming. Moreover, a first tree model of tonal information has been recently proposed. This approach must now be experimented and adapted to audio signals. Rhythm and timbre are two other main musical characteristics. In order to be able to analyze and compare these information, specific features, similar to chromas for tonal information, have to be proposed. These works, in particular concerning metric information, will be done in collaboration with the SCRIIME.

Analysis of the Musical Structure

Multidimensional sequences alignment During the last few years, we have introduced several methods to compare two musical pieces based on string alignment algorithms. These methods use a symbolic representation of the melody as a sequence of notes (pitch/duration). Taking into account other music characteristics such as harmony, rhythm, tone, ...would certainly improve results but it will increase the complexity of the representations of melodies. Indeed, the comparison of such sequences of set of symbols leads to a n-dimensional sequences alignment problem. Following models developed in bio-informatics, we propose a new model of a musical piece made of n sequences, each one modeling one kind of information. Several methods have been already developed to compare multidimensional biological sequences. As for the one dimension alignment, we will have to adapt these algorithms to fit musical context.

Indexing Most of western music are made of repeated patterns. Thus, a musical piece can be described as repetitions of structures (for example verses, chorus), repetitions of chords or
repetitions of notes. The main issue is the occurrences are not exact: for example a chorus may be transposed, accelerated, etc. Recently, we proposed an algorithm to extract repeated factors in a musical piece. These repeated factors must be errorless; a challenge is to improve this algorithm to allow errors in the repeats. To do so, we could use the distances we have previously developed.

**Self similarity** Formally, a structure is so-called self-similar when small parts of it are similar (or qualitatively the same) to the whole structure. In musical context, since musical piece have a finite size, we will study approximate self-similarity. We proposed a first method to evaluate this self-similarity by comparing a musical piece against itself. The first developed tools must be tuned and novel original approaches must be designed in order to detect efficiently the different repetitions of a musical piece. Thereby, the self-similarity analysis would be a useful tool to automatically detect patterns in musical pieces.

We can consider various applications based on the extracted patterns. The main one concerns the indexation of musical databases. Once the repetitions are automatically extracted, a musical piece could be compressed by keeping only one occurrence for each repetition. Then, we will be able to generate accurate summary of the piece (like thumb-nailing). For music database navigation, a structural analysis of the pieces will allow exhibiting structural comparisons. As the structure of a piece is strongly linked to its genre, this will lead to genre based or sound based browsers.

**Application to large databases** For a few years, browsing large databases has become a classical operation. Based on dynamic programming, local alignment methods investigate all the possible mappings between two musical pieces and return the optimal local alignment as result. This exhaustive analysis is very costly from a computational point of view. The exponential growth of databases makes the use of exact alignment algorithms impossible. Alternatives have been proposed in Bioinformatics, by applying heuristics or by developing faster implementations.

The first approach relies on the indexing of musical pieces. It has already been applied with success in Bioinformatics and leads to tools like BLAST, one of the most used and efficient tools of the Bioinformatics community. The main idea of this approach consists in indexing pieces of the database in order to be able to quickly retrieve exact occurrences of fragments of the query. Once the potential candidates have been identified, an alignment algorithm allows refining results. To our knowledge, no indexing structure dedicated to polyphonic music has been proposed. These new comparison methods may rely on methods developed for music indexing.

In parallel, the quality of the results obtained by applying exact comparison algorithms lead us to develop new faster implementation. Indeed, we propose to consider the huge computation power of graphic cards available in public domain in order to develop high performance applications for Query-by-Humming or plagiarism detection for example.

**Applications**

Applications from our cross-domain project are numerous. We thus aim at developing musical content-based search engines. Musical pieces could then be retrieved from a symbolic query, based on the similarity with respect of the melody, the harmony, the rhythm, the timbre or the structure. Current works on structure inference or detection of repetitions in musical pieces may lead to propose automatic systems to summarize musical pieces.

Another kind of applications concerns the development of musical video games and of some new teaching tools for musicians. The quality of the feedback, the rightness of the evaluation of a musical performance is an essential aspect for a gamer or a student. We aim at applying our similarity systems, using the advantages of the alignment algorithms, to compare the musical performances to models issued from either a score, the performance of an expert, or the same performer in the past. A global score may be computed, and the differences could also be precisely located. Adapting the feedback to the technical level of the performer (correctness of the notes for a beginner, of the tuning for a more advanced student) may improve the efficiency of the teaching process in identifying the main technical lacks.
Including some works on musical synthesis may enhance active listening systems. A listener could then be active on metric or harmonic parameters, changing the chord progression, the tonality or the rhythm of a musical piece.

8.3 Project and prospects for RHOBAN : PERSONAL ROBOTICS

These last few years, the research in artificial intelligence, learning machines, or shape recognition has produced some significant concepts and results. However, we are far from being able to design machines able to adapt themselves to our physical environment, and also to our social environment. Indeed, machines are far from reaching the level of flexibility, robustness and reactivity of the human being, or even of the little child. Our challenge consists in designing robots being able to discover their environment and to adapt themselves, to learn new knowledges in an unknown environment like human being.

Rhoban project focus on motor aspects for complex robots like humanoids. We already have a prototype, humanoid, able to stabilize during moves like walking, etc. This development follows the direction of “personal robotic”, which aims at introducing robots in the every day life, for instance, helping elderly, or else entertainment robots.

The project has been initiated in a private framework. O. Ly, which launched it, has been laureate at the 11th “Concours national d’aide à la création d’entreprise de technologies innovantes”, (OSEO, MESR). At the moment, the project is in development in a framework grouping the LaBRI, the IMS and the INRIA.