IN PARTNERSHIP WITH:
CNRS

Université de Lorraine

Activity Report 2011

Project-Team MAIA

Autonomous intelligent machine

IN COLLABORATION WITH: Laboratoire lorrain de recherche en informatique et ses applications (LORIA)

RESEARCH CENTER
Nancy - Grand Est

THEME
Knowledge and Data Representation and Management

Table of contents

Project-Team MAIA

Keywords: Knowledge, Artificial Intelligence, Real-Time, Robotics, Planning, Reasoning

1. Members

Research Scientists

François Charpillet [Team Leader, Senior Researcher, INRIA, HdR]
Olivier Buffet [Junior Researcher, INRIA]
Alain Dutech [Junior Researcher, INRIA, HdR]
Nazim Fatès [Junior Researcher, INRIA]
Jörg Hoffmann [Senior Researcher, INRIA, HdR]
Bruno Scherrer [Junior Researcher, INRIA]

Faculty Members

Christine Bourjot [Associate Professor, U. Nancy 2]
Vincent Chevrier [Associate Professor, UHP, HdR]
Alexis Scheuer [Associate Professor, UHP]
Olivier Simonin [Associate Professor, UHP, HdR]
Vincent Thomas [Associate Professor, U. Nancy 2]

External Collaborators

Maan El-Badaoui-El-Najjar [Associate Professor, Lille]
Amine Boumaza [Associate Professor, Calais]

Technical Staff

Nicolas Beaufort [Engineer]
Abdallah Dib [Engineer]
Lionel Havet [Engineer]
Marie Tonnelier [Engineer]

PhD Students

Mauricio Araya [scholarship, Chilian governement, U. Nancy 2]
Antoine Bautin [ANR scholarship, UHP]
Olivier Bouré [MENRT scholarship, UHP]
Amandine Dubois [MENRT scholarship, U. Nancy 2]
Arsène Fansi Tchango [CIFRE, Thales, U. Nancy 2]
Tomas Navarrete [scholarship, Mexican government, UHP]
Cédric Rose [CIFRE, Diatelic SA, UHP]
Julien Siebert [ATER, UHP]
Manel Tagorti [ANR project BARQ, UHP]
Mohamed Tlig [EU project InTraDE, UHP]
Jano Yazbeck [EU project InTraDE, UHP]

Post-Doctoral Fellows

Emil Keyder [INRIA scholarship]
Michael Katz [ANR project BARQ]
Arnaud Glad [ATER, INPL]

Administrative Assistant

Céline Simon [TR, INRIA]

2. Overall Objectives

2.1. Overall Objectives

The objective of the MAIA[1] team is to address foundation and engineering aspects of artificial intelligence. In this general framework, the team investigates the design and understanding of intelligent agents[2] which autonomously perceives and acts upon an environment so as to achieve one or several goals. The MAIA group equally addresses the design of a single agent, a team of agents or a large number of agents. This common objective is considered from two perspectives organized around two lines of research:

- The first research activity is about ***sequential decision making***. It has been influenced by Stuart Russell [66] who considers that an agent is rational. In his vision: [For each possible percept sequence, an ideal rational agent should do whatever action is expected to maximize its performance measure] [65]. This view makes Markov decision processes (MDPs) and more generally sequential decision making a good candidate for building the behavior of an agent. It is probably why MDPs have received considerable attention in recent years by the artificial intelligence (AI) community.
- The second activity is about ***understanding and engineering reactive multi-agent systems***. It is influenced by research results from the field of behavioral biology which gives us some of the keys to understand how intelligent and adaptive behaviors appear in natural swarm systems. This encourages us to study principles of emergent behaviors in natural systems and apply them to design artificial intelligent systems. Reactive multi-agent systems are good candidates for building such autonomous and adaptive systems and our work mainly focuses on better understanding how we can soundly build such systems.

3. Scientific Foundations

3.1. Sequential Decision Making

3.1.1. Synopsis and Research Activities

Sequential decision making consists, in a nutshell, in controlling the actions of an agent facing a problem whose solution requires not one but a whole sequence of decisions. This kind of problem occurs in a multitude of forms. For example, important applications addressed in our work include: Robotics, where the agent is a physical entity moving in the real world; Medicine, where the agent can be an analytic device recommending tests and/or treatments; Computer Security, where the agent can be a virtual attacker trying to identify security holes in a given network; and Business Process Management, where the agent can provide an auto-completion facility helping to decide which steps to include into a new or revised process. Our work on such problems is characterized by three main research trends:

(A) *Understanding how, and to what extent, to best model the problems.*

(B) *Developing algorithms solving the problems and understanding their behavior.*

(C) *Applying our results to complex applications.*

[1] MAIA stands for “MAchine Intelligente et Autonome”, that is “Autonomous and Intelligent MAchine”
[2] In the field of artificial intelligence, an “agent” refers to an entity

Before we describe some details of our work, it is instructive to understand the basic forms of problems we are addressing. We characterize problems along the following main dimensions:

(1) Extent of the model: full vs. partial vs. none. This dimension concerns how complete we require the model of the problem – if any – to be. If the model is incomplete, then learning techniques are needed along with the decision making process.

(2) Form of the model: factored vs. enumerative. Enumerative models explicitly list all possible world states and the associated actions etc. Factored models can be exponentially more compact, describing states and actions in terms of their behavior with respect to a set of higher-level variables.

(3) World dynamics: deterministic vs. stochastic. This concerns our initial knowledge of the world the agent is acting in, as well as the dynamics of actions: is the outcome known a priori or are several outcomes possible?

(4) Observability: full vs. partial. This concerns our ability to observe what our actions actually do to the world, i.e., to observe properties of the new world state. Obviously, this is an issue only if the world dynamics are stochastic.

These dimensions are wide-spread in the AI literature. We remark that they are not exhaustive. In parts of our work, we also consider the difference between discrete/continuous problems, and centralized/decentralized problems. The complexity of solving the problem – both in theory and in practice – depends crucially on where the problem resides in this categorization. In many applications, not one but several points in the categorization make sense: simplified versions of the problem can be solved much more effectively and thus serve for the generation of *some* – if possibly sub-optimal – action strategy in a more feasible manner. Of course, the application as such may also come in different facets.

In what follows, we outline the main formal frameworks on which our work is based; while doing so, we highlight in a little more detail our core research questions. We then give a brief summary of how our work fits into the global research context.

3.1.2. Formal Frameworks

3.1.2.1. Deterministic Sequential Decision Making

Sequential decision making with deterministic world dynamics is most commonly known as **planning**, or **classical planning** [59]. Obviously, in such a setting every world state needs to be considered at most once, and thus enumerative models do not make sense (the problem description would have the same size as the space of possibilities to be explored). Planning approaches support factored description languages allowing to model complex problems in a compact way. Approaches to automatically learn such factored models do exist, however most works – and also most of our works on this form of sequential decision making – assume that the model is provided by the user of the planning technology. Formally, a problem instance, commonly referred to as a **planning task**, is a four-tuple $\langle V, A, I, G \rangle$. Here, V is a set of variables; a value assignment to the variables is a world state. A is a set of actions described in terms of two formulas over V: their preconditions and effects. I is the initial state, and G is a goal condition (again a formula over V). A solution, commonly referred to as a **plan**, is a schedule of actions that is applicable to I and achieves G.

Planning is **PSPACE**-complete even under strong restrictions on the formulas allowed in the planning task description. Research thus revolves around the development and understanding of search methods, which explore, in a variety of different ways, the space of possible action schedules. A particularly successful approach is **heuristic search**, where search is guided by information obtained in an automatically designed **relaxation** (simplified version) of the task. We investigate the design of relaxations, the connections between such design and the search space topology, and the construction of effective **planning systems** that exhibit good practical performance across a wide range of different inputs. Other important research lines concern the application of ideas successful in planning to stochastic sequential decision making (see next), and the development of technology supporting the user in model design.

3.1.2.2. Stochastic Sequential Decision Making

Markov Decision Processes (**MDP**) [63] are a natural framework for stochastic sequential decision making. An MDP is a four-tuple $\langle S, A, T, r\rangle$, where S is a set of states, A is a set of actions, $T(s,a,s') = P(s'|s,a)$ is the probability of transitioning to s' given that action a was chosen in state s, and $r(s,a,s')$ is the (possibly stochastic) reward obtained from taking action a in state s, and transitioning to state s'. In this framework, one looks for a **strategy**: a precise way for specifying the sequence of actions that induces, on average, an optimal sum of discounted rewards $\mathrm{E}\left[\sum_{t=0}^{\infty}\gamma^t r_t\right]$. Here, $(r_0, r_1, ...)$ is the infinitely-long (random) sequence of rewards induced by the strategy, and $\gamma \in (0,1)$ is a discount factor putting more weight on rewards obtained earlier. Central to the MDP framework is the Bellman equation, which characterizes the **optimal value function** V^*:

$$\forall s \in S, \quad V^*(s) = \max_{a\in A}\sum_{s'\in S} T(s,a,s')[r(s,a,s') + \gamma V^*(s')].$$

Once the optimal value function is computed, it is straightforward to derive an optimal strategy, which is deterministic and memoryless, i.e., a simple mapping from states to actions. Such a strategy is usually called a **policy**. An **optimal policy** is any policy π^* that is **greedy** with respect to V^*, i.e., which satisfies:

$$\forall s \in S, \quad \pi(s) \in \arg\max_{a\in A}\sum_{s'\in S} T(s,a,s')[r(s,a,s') + \gamma V^*(s')].$$

An important extension of MDPs, known as Partially Observable MDPs (**POMDPs**) allows to account for the fact that the state may not be fully available to the decision maker. While the goal is the same as in an MDP (optimizing the expected sum of discounted rewards), the solution is more intricate. Any POMDP can be seen to be equivalent to an MDP defined on the space of probability distributions on states, called **belief states**. The Bellman-machinery then applies to the belief states. The specific structure of the resulting MDP makes it possible to iteratively approximate the optimal value function – which is convex in the **belief space** – by piecewise linear functions, and to deduce an optimal policy that maps belief states to actions. A further extension, known as a DEC-POMDP, considers $n \geq 2$ agents that need to control the state dynamics in a decentralized way without direct communication.

The MDP model described above is enumerative, and the complexity of computing the optimal value function is **polynomial** in the size of that input. However, in examples of practical size, that complexity is still too high so naïve approaches do not scale. We consider the following situations: (i) when the state space is large, we study approximation techniques from both a theoretical and practical point of view; (ii) when the model is unknown, we study how to learn an optimal policy from samples (this problem is also known as Reinforcement Learning [69]); (iii) in factored models, where MDP models are a strict generalization of classical planning – and are thus at least **PSPACE**-hard to solve – we consider using search heuristics adapted from such (classical) planning.

Solving a POMDP is **PSPACE**-hard even given an enumerative model. In this framework, we are mainly looking for assumptions that could be exploited to reduce the complexity of the problem at hand, for instance when some actions have no effect on the state dynamics (**active sensing**). The decentralized version, DEC-POMDPs, induces a significant increase in complexity (**NEXP**-complete). We tackle the challenging – even for (very) small state spaces – exact computation of finite-horizon optimal solutions through alternative reformulations of the problem. We also aim at proposing advanced heuristics to efficiently address problems with more agents and a longer time horizon.

3.1.3. Project-team positioning

Within INRIA, the most closely related teams are TAO and Sequel. TAO works on evolutionary computation (EC) and statistical machine learning (ML), and their combination. Sequel works on ML, with a theoretical focus combining CS and applied maths. The main difference is that TAO and Sequel consider particular algorithmic frameworks that can, amongst others, be applied to Planning and Reinforcement Learning, whereas we revolve around Planning and Reinforcement Learning as the core problems to be tackled, with whichever framework suitable.

In France, we have recently begun collaborating with the IMS Team of Supélec Metz, notably with O. Pietquin and M. Geist who have a great expertise in approximate techniques for MDPs. We have have links with the MAD team of the BIA unit of the INRA at Toulouse, lead by F. Garcia. They also use MDP related models and are interested in solving large size problems, but they are more driven by applications (mostly agricultural) than we are. In Paris, the Animat Lab, that was a part of the LIP6 and is now attached to the ISIR, has done some interesting works on factored Markov Decision Problems and POMDPs. Like us, their main goal was to tackle problems with large state space.

In Europe, the IDSIA Lab at Lugano (Switzerland) has brought some interesting ideas to the field of MDP (meta-learning, subgoal discovery) but seems now more interested in a *Universal Learner*. In Osnabrück (Germany), the Neuroinformatic group works on efficient reinforcement learning with a specific interest in the application to robotics. For deterministic planning, the most closely related groups are located in Freiburg (Germany), Glasgow (UK), and Barcelona (Spain). We have active collaborations with all of these.

In the rest of the world, the most important groups regarding MDPs can be found at Brown University, Rutgers Univ. (M. Littman), Univ. of Toronto (C. Boutilier), MIT AI Lab (L. Kaelbling, D. Bertsekas, J.Tsitsiklis), Stanford Univ., CMU, Univ. of Alberta (R. Sutton), Univ. of Massachusetts at Amherst (S. Zilberstein, A. Barto), *etc*. A major part of their work is aimed at making Markov Decision Process based tools work on real life problems and, as such, our scientific concerns meet theirs. For deterministic planning, important related groups and collaborators are to be found at NICTA (Canberra, Australia) and at Cornell University (USA).

3.2. Understanding and mastering complex systems

3.2.1. General context

There exist numerous examples of natural and artificial systems where self-organization and emergence occur. Such systems are composed of a set of simple entities interacting in a shared environment and exhibit complex collective behaviors resulting from the interactions of the local (or individual) behaviors of these entities. The properties that they exhibit, for instance robustness, explain why their study has been growing, both in the academic and the industrial field. They are found in a wide panel of fields such as sociology (opinion dynamics in social networks), ecology (population dynamics), economy (financial markets, consumer behaviors), ethology (swarm intelligence, collective motion), cellular biology (cells/organ), computer networks (ad-hoc or P2P networks), etc.

More precisely, the systems we are interested in are characterized by :

- *locality*: Elementary components have only a partial perception of the system's state, similarly, a component can only modify its surrounding environment.
- *individual simplicity*: components have a simple behavior, in most cases it can be modeled by stimulus/response laws or by look-up tables. One way to estimate this simplicity is to count the number of stimulus/response rules for instance.
- *emergence*: It is generally difficult to predict the global behavior of the system from the local individual behaviors. This difficulty of prediction is often observed empirically and in some cases (e.g., cellular automata) one can show that the prediction of the global properties of a system is an undecidable problem. However, observations coming from simulations of the system may help us to find the regularities that occur in the system's behavior (even in a probabilistic meaning). Our interest is to work on problems where a full mathematical analysis seems out of reach and where it is useful to observe the system with large simulations. In return, it is frequent that the properties observed empirically are then studied on an analytical basis. This approach should allow us to understand more clearly where lies the frontier between simulation and analysis.

- *levels of description and observation*: Describing a complex system involves at least two levels: the micro level that regards how a component behaves, and the macro level associated with the collective behavior. Usually, understanding a complex system requires to link the description of a component behavior with the observation of a collective phenomenon: establishing this link may require various levels, which can be obtained only with a careful analysis of the system.

We now describe the type of models that are studied in our group.

3.2.2. Multi-agent models

To represent these complex systems, we made the choice to use reactive multi-agent systems (RMAS). Multi-agent systems are defined by a set of reactive agents, an environment, a set of interactions between agents and a resulting organization. They are characterized by a decentralized control shared among agents: each agent has an internal state, has access to local observations and influences the system through stimulus response rules. Thus, the collective behavior results from individual simplicity and successive actions and interactions of agents through the environment.

Reactive multi-agent systems present several advantages for modeling complex systems

- agents are explicitly represented in the system and have the properties of local action, interaction and observation;
- each agent can be described regardless of the description of the other agents, multi-agent systems allow explicit heterogeneity among agents which is often at the root of collective emergent phenomena;
- Multi-agent systems can be executed through simulation and provide good model to investigate the complex link between global and local phenomena for which analytic studies are hard to perform.

By proposing two different levels of description, the local level of the agents and the global level of the phenomenon, and several execution models, multi-agent systems constitute an interesting tool to study the link between local and global properties.

Despite of a widespread use of multi-agent systems, their framework still needs many improvements to be fully accessible to computer scientists from various backgrounds. For instance, there is no generic model to mathematically define a reactive multi-agent system and to describe its interactions. This situation is in contrast with the field of cellular automata, for instance, and underlines that a unification of multi-agent systems under a general framework is a question that still remains to be tackled. We now list the different challenges that, in part, contribute to such an objective.

3.2.3. Current challenges

Our work is structured around the following challenges that combine both theoretical and experimental approaches.

3.2.3.1. Providing formal frameworks

Currently, there is no agreement on a formal definition of a multi-agent system. Our research aims at translating the concepts from the field of complex systems into the multi-agent systems framework.

One objective of this research is to remove the potential ambiguities that can appear if one describes a system without explicitly formulating each aspect of the simulation framework. As a benefit, the reproduction of experiments is facilitated. Moreover, this approach is intended to gain a better insight of the self-organization properties of the systems.

Another important question consists in monitoring the evolution of complex systems. Our objective is to provide some quantitative characteristics of the system such as local or global stability, robustness, complexity, etc. Describing our models as dynamical systems leads us to use specific tools of this mathematical theory as well as statistical tools.

3.2.3.2. *Controlling complex dynamical system*

Since there is no central control of our systems, one question of interest is to know under which conditions it is possible to guarantee a given property when the system is subject to perturbations. We tackle this issue by designing exogeneous control architectures where control actions are envisaged as perturbations in the system. As a consequence, we seek to develop control mechanism that can change the global behavior of a system without modifying the agent behavior (and not violating the autonomy property).

3.2.3.3. *Designing systems*

The aim is to design individual behaviors and interactions in order to produce a desired collective output. This output can be a collective pattern to reproduce in case of simulation of natural systems. In that case, from individual behaviors and interactions we study if (and how) the collective pattern is produced. We also tackle "inverse problems" (decentralized gathering problem, density classification problem, etc.) which consist in finding individual behaviors in order to solve a given problem.

3.2.4. Project-team positioning

Building a reactive multi-agent system consists in defining a set (generally a large number) of simple and reactive agents within a shared environment (physical or virtual) in which they move, act and interact with each other. Our interest in these systems is that, in spite of their simple definition at the agent level, they produce coherent and coordinated behavior at a global scale. The properties that they may exhibit, such as robustness and adaptivity explain why their study has been growing in the last decade (in the broader context of "complex systems").

Our work on such problems is characterized by five research trends: (A) *Defining a formal framework for describing and studying these systems*, (B) *Developing and understanding reactive multi-agent systems*, (C) *Analysing and proving properties*, (D) *Deploying these systems on typical distributed architectures such as swarms of robots, FPGAs, GPUs and sensor networks*, (E) *Transferring our results in applications.*

Multi-agent System is an active area of research in Artificial Intelligence and Complex Systems. Our research fits well into the international research context, and we have made and are making a variety of significant contributions both in theoretical and practical issues. Concerning multi-agent simulation and formalization, we compete or collaborate in France with S. Hassas in LIESP (Lyon), CERV (Brest), IREMIA (la Réunion), Ibisc (Evry), Lirmm (Montpellier), Irit (Toulouse), A. Drogoul (IRD, Bondy) and abroad with F. Zambonelli (Univ. Modena, Italy) A. Deutsch (Dresden, Germany), D. Van Parunak (Vector research, USA), P. Valkenaers, D. Weyns (Univ. Leuven, Belgium), etc. Regarding our work on swarm robotics we have common objectives with the DISAL[3] EPFL Laboratory, the Bristol Robotics Laboratory, the Distributed Robotics Laboratory at MIT, the team of W. & D. Spears at Wyoming university, the Pheromone Robotics project at HRL Lab.[4], the FlockBots project at GMU[5], the team of G. Théraulaz at CNRS-Toulouse and the teams of J.-L. Deneubourg and M. Dorigo at ULB (Bruxelles).

4. Software

4.1. FF

Participant: Jörg Hoffmann [correspondant].

[3]Distributed Intelligent Systems and Algorithms Laboratory including EPFL Swarm-Intelligent Systems Group (SWIS) founded in 2003 and the Collective Robotics Group (CORO) founded in 2000 at California Institute of Technology USA

[4]HRL, Information and systems sciences Lab (ISSL), Malibu CA, USA (D. Payton)

[5]George Mason University, Eclab, USA (L. Panait, S. Luke)

FF is an automatic planning system, taking as input a high-level description of the planning task in the PDDL language (planning domain definition language), and returning a plan for the task. FF was continuously developed by Jörg Hoffmann over a time span of several years (ca. 1999 – 2006), before joining INRIA. FF has convincingly won the international planning competition in the year 2000, and has been one of the most widely used and cited planning systems (around 1000 citations up to now) ever since then. It still is competitive with the state of the art today. There are several different versions, for deterministic planning with Boolean state variables, for deterministic planning with numeric state variables, for non-deterministic planning with no probabilities (all outcomes are assumed to be equally likely), and finally a version tackling a particular variant of probabilistic planning.

4.2. TorchLight

Participant: Jörg Hoffmann [correspondant].

TorchLight is a system for auomatic domain analysis in planning. It automatically infers properties of the search space surface under a partcular heuristic function, called h^+, that underlies most current state of the art planning systems (including FF). TorchLight examines certain structural properties of the PDDL input, and exploits a number of connections between this structure and the search space surface under h^+. For example, one of its outputs provides an estimate of the fraction of states that lie on local minima.

4.3. AA4MM

Participants: Vincent Chevrier [correspondant], Julien Siebert.

This work is undertaken in a joint Phd Thesis between MAIA and Madynes Team. Laurent Ciarletta (Madynes team, LORIA) is co-advisor of this PhD and correspondant for this software.

AA4MM (Agents and Artefacts for Multi-modeling and Multi-simulation) is a framework for coupling existing and heterogeneous models and simulators in order to model and simulate complex systems. This is the first implementation of the AA4MM meta-model proposed in Julien Siebert's PhD. It is written in Java and relies upon Java Messaging Services (JMS) for its distributed version.

4.4. MASDYNE

Participants: Vincent Chevrier [correspondant], Julien Siebert.

This work is undertaken in a joint Phd Thesis between MAIA and Madynes Team. Laurent Ciarletta (Madynes team, LORIA) is co-advisor of this PhD and correspondant for this software.

Other contributors to this software are: Tom Leclerc, François Klein, Christophe Torin, Marcel Lamenu, Guillaume Favre and Amir Toly.

MASDYNE (Multi-Agent Simulator of DYnamic Networks usErs) is a multi-agent simulator for modeling and simulating users behaviors in mobile ad hoc network. This software is part of joint work with MADYNES team, on modeling and simulation of ubiquitous networks.

5. New Results

5.1. Decision Making

5.1.1. Optimizing Automated Service Discovery

Participant: Jörg Hoffmann.

Michael Stollberg (SAP Research, Germany) and Dieter Fensel (University of Innsbruck, Austria) are external collaborators.

We completed earlier work, done while all authors were employed at the University of Innsbruck, and published it in the International Journal of Semantic Computing [10]. In a nutshell, the work proposes to use first-order logic for annotating web services to accomplish better precisiona nd recall in service discovery; its core contribution is a technique making such discovery more effective – discovery here involves first-order logical reasoning – by designing a caching technique storing known relationships between available services and possible discovery queries.

5.1.2. Overview of Semantic Web Service Technologies

Participant: Jörg Hoffmann.

Stijn Heymans (SemanticBits, USA), Annapaola Marconi (Fondazione Bruno Kessler, Trento, Italy), Joshua Phillips (SemanticBits, USA), and Ingo Weber (University of New South Wales, Sydney, Australia) are external collaborators.

We were invited to write a book chapter about the basic AI technologies underlying semantic Web service discovery and composition. The chapter has been published as part of a book entitled "Handbook of Service Description – USDL and its Methods" in Springer-Verlag [46].

5.1.3. Analyzing Planning Domains to Predict Heuristic Function Quality

Participant: Jörg Hoffmann.

The heuristic search approach to planning (cf. the above) rises and falls with the quality of the heuristic estimates. The dominant method, especially in satisficing (non-optimal) planning, is to approximate a heuristic function called h^+ – this is used in almost every state of the art satisficing planning system. In earlier work, Jörg Hoffmann showed that h^+ has some amazing qualities, in many traditional planning benchmarks, in particular pertaining to the complete absence of local minima. [62] His proofs of this are hand-made, raising the question whether such proofs can be lead automatically by domain analysis techniques. The possible uses of such analysis are manifold, e.g., for automatic configuration of hybrid planners or for giving hints how to improve the domain design. The question has been open since 2002. A serious attempt of Jörg Hoffmann resulted in disappointing results – his analysis method has exponential runtime and succeeds only in two extremely simple benchmark domains. In contrast to this, in our work here we answer the question in the affirmative. We establish connections between certain easily testable syntactical structures, called "causal graphs", and h^+ topology. This results in low-order polynomial time analysis methods, implemented in the Torchlight tool, cf. Section 4.2. Of the 12 domains where Hoffmann proved the absence of local minima, TorchLight gives strong success guarantees in 8 domains. Empirically, its analysis exhibits strong performance in a further 2 of these domains, plus in 4 more domains where local minima may exist but are rare. We show that, in this way, TorchLight can distinguish Hoffmann's "easy" domains from the "hard" ones. By summarizing structural reasons for analysis failure, TorchLight also provides diagnostic output pin-pointing potentially problematic aspects of the domain. A conference paper on this work was published at ICAPS 2011 [25], and nominated for the best paper award there. A journal paper was published in the Journal of AI Research (JAIR) [9].

5.1.4. Relaxing Bisimulation for State Aggregation in the Computation of Lower Bounds

Participant: Jörg Hoffmann.

Raz Nissim (Ben-Gurion University, Beer-Sheva, Israel) and Malte Helmert (University of Freiburg, Germany) are external collaborators.

Like the previous line of work, this addresses planning as heuristic search, specifically the automatic generation of heuristic estimates. This is also the core question investigated in the BARQ project, see below. In preparation of this project, we are conducting this line of research, which explores some of the most basic ideas behind BARQ. The basic technique under consideration was developed in prior work outside INRIA. [61] The heuristic estimates are lower bounds generated from a quotient graph in which sets of states are aggregated into equivalence classes. A major difficulty in designing such classes is that there are exponentially many states. Despite this, our technique allows explicit selection of individual states to aggregate, via an incremental process interleaving it with state space re-construction steps. We have shown previously that, if the aggregation decisions are perfect, then this technique dominates the other known related techniques, and sometimes produces perfect estimates in polynomial time. But how to take these decisions? Little is known about this as yet. In the present work, we start from the notion of a "bisimulation", which is a well-known criterion from model checking implying that the quotient system is behaviorally indistinguishable from the original system – in particular, the cost estimates based on a bisimulation are perfect. However, bisimulations are exponential even in trivial planning benchmarks. We observe that bisimulation can be relaxed without losing any information as far as the cost estimates are concerned. Namely, we can ignore the "content of the messages sent", i.e., the state transition labels. Such relaxed bisimulations are often exponentially smaller than the original ones. We show to what extent such relaxation can be applied also within our incremental construction process. As a result, in several benchmarks we obtain perfect estimates in polynomial time, and we significantly increase the set of benchmark instances that can be solved with this approach. Indeed, the approach obtained a 2nd place in the optimal track of the 2011 International Planning Competition, and was part of the 1st-prize winning portfolio. A conference paper was published at IJCAI 2011 [28], and a journal paper is under preparation for submission to the Journal of the ACM.

5.1.5. Relaxing Bisimulation by Choosing Transition Subsets

Participants: Michael Katz, Jörg Hoffmann.

Malte Helmert (University of Freiburg, Germany) is an external collaborator.

This line of work builds on the previous one by designing new methods for relaxing bisimulations. The key idea is to apply the bisimulation property to only a subset of the transitions in the system under consideration. We showed that one can ignore large subsets of transitions without losing any information, i.e., while still guaranteeing to obtain a perfect heuristic. At the same time, such a relaxed bisimulation makes less distinctions and may thus be exponentially smaller. For practical purposes, we designed several approximate strategies relaxing more, obtaining smaller abstractions at the expense of informtion loss. The techniques are currently being evaluated empirically, and a paper submission is in preparation for ICAPS'12.

5.1.6. Improving h^+ by Taking Into Account (Some) Negative Effects

Participants: Emil Keyder, Jörg Hoffmann.

Patrik Haslum (NICTA, Australia) is an external collaborator.

Like the previous lines, this is on planning as heuristic search. As mentioned above in Section 5.1.3, approximating the h^+ heuristic is the dominant approach to obtain estimates in satisficing (non-optimal) planning. That notwithstanding, h^+ is obtained by ignoring all negative effects, which of course leads to very bad estimates in domains where these domains play a key role, for example puzzle-like domains, e.g. Rubic's cube, where actions intrefere intensively with each other. It has long (for almost 10 years) been an active research issue how to take at least some of the negative effects into account when computing h^+. All attempts, however, remained at rather ad-hoc methods, like, counting the number of violated binary constraints (pairs of facts that cannot be true at the same time) within the relaxed plan underlying the estimate. In the present work, for the first time we provide a well-founded formal approach to the issue. As was suggested in prior work, [60], we design a compiled planning task which introduces constructs allowing h^+ to correctly handle a subset C of fact conjunctions. Whereas this prior work requires a compilation exponential in $|C|$ – and thus allows only to introduce very few conjunctions – in our work we designed a compilation that is linear in $|C|$. We proved that one can always choose C so that h^+ in the compiled task is a perfect heuristic. Of course, in general C might have to be exponentially large to achieve this. We designed practical methods selecting C in a way so that the overhead (the size of C) is kept at bay, while the quality of the heuristic is sufficiently improved to boost search performance. The techniques are currently being evaluated empirically, and a paper submission is in preparation for ICAPS'12.

5.1.7. Accounting for Uncertainty in Penetration Testing

Participants: Olivier Buffet, Jörg Hoffmann.

Carlos Sarraute (Core Security Technologies) is an external collaborator.

Core Security Technologies is an U.S.-American/Argentinian company providing, amongst other things, tools for (semi-)automated security checking of computer networks against outside hacking attacks. For automation of such checks, a module is needed that automatically generates potential attack paths. Since the application domain is highly dynamic, a module allowing to declaratively specify the environment (the network and its configuration) is highly advantageous. For that reason, Core Security Technologies have been looking into using AI Planning techniques for this purpose. After consulting by Jörg Hoffmann (see also Section 6.1.1 below), they are now using a variant of Jörg Hoffmann's FF planner (cf. Section 4.1) in their product. While that solution is satisfactory in many respects, it also has weaknesses. The main weakness is that it does not handle the incomplete knowledge in this domain – figuratively speaking, the attacker is assumed to have perfect information about the network. This results in high costs in terms of runtime and network traffic, for extensive scanning activities prior to planning. We are currently working with Core Security's research department to overcome this issue, by modeling and solving the attack planning problem as a POMDP instead. A workshop paper detailing the POMDP model has been published at SecArt'11 [29]. While such a model yields much higher quality attacks, solving an entire network as a POMDP is not feasible. We have designed a decomposition method making use of network structure and approximations to overcome this problem, by using the POMDP model only to find good-quality attacks on single machines, and propagating the results through the network in an appropriate manner. A conference paper is in preparation for submission to ICAPS'12.

5.1.8. Searching for Information with MDPs

Participants: Mauricio Araya, Olivier Buffet, Vincent Thomas, François Charpillet.

In the context of Mauricio Araya's PhD, we are working on how MDPs —or related models— can search for information. This has led to various research directions that we describe now.

<u>A POMDP Extension with Belief-dependent Rewards</u> — A limitation of Partially Observable Markov Decision Processes (POMDPs) is that they only model problems where the performance criterion depends on the state-action history. This excludes for example scenarios where one wants to maximize the knowledge with respect to some random variables.

To overcome this limitation, we have proposed ρ-POMDPs, an extension of POMDPs in which the reward function depends on the belief state rather than on the state. In this framework, and under the hypothesis that the reward function is convex, we have proved that:

- the value function itself is convex; and
- if the reward function is α-Hölder, then the value function can be approximated arbitrarily well with a piecewise linear and convex function.

These results allow for adapting a number of solution algorithms relying on approximating the value function.

This theoretical work has been first published in an international conference in December 2010, then in [36], where it has received a best paper award.

We are currently pursuing experimental work about the proposed algorithm.

<u>Active Learning of MDP Models</u> — Reinforcement Learning is about learning how to perform a task by trial and error (no model of the system to control being available). Model-based Bayesian RL (BRL) consists in all RL algorithms that maintain a belief (in the Bayesian sense) about the model of the system to control. In fact, this is a way to turn an RL problem into a POMDP—the unknown model becoming an unobservable part of the state—, thus replacing the exploration-exploitation dilemma by the definition of a prior belief over possible models.

A particular BRL task we have been considering is to actively learn the dynamical model itself, i.e., to act so as to improve the knowledge about the transition function. In a way this means solving a ρ-POMDP since the reward depends on a belief, not on a state. To that end, we have proposed several optimization criteria, and derived the corresponding reward functions, making sure that their computational complexity allows for their use in a BRL algorithm. We have also proved that a non-optimistic BRL algorithm—EXPLOIT—could be used in this particular case.

This work, along with experiments, has been published in [36] and [35] (french version).

PAC-BAMDP Algorithms — Exact or approximate solutions to Model-based Bayesian RL are impractical, so that a number of heuristic approaches have been considered, most of them relying on the principle of "optimism in the face of uncertainty". Some of these algorithms have properties that guarantee the quality of their outcome, inspired by the PAC-learning (Probably Approximately Correct) framework. For example, some algorithms provably make in most cases the same decision as would be made if the true model were known (PAC-MDP property).

We have proposed a novel optimistic algorithm, BOUH, that is

- appealing in that it is (i) optimistic *about* the uncertainty in the model and (ii) deterministic (thus easier to study); and
- provably PAC-BAMDP, i.e., makes in most cases the same decision as a perfect BRL algorithm would.

First results about this algorithm are currently under review.

5.1.9. Scheduling for Probabilistic Realtime Systems

Participant: Olivier Buffet.

Maxim Dorin, Luca Santinelli, Liliana Cucu-Grosjean (INRIA, TRIO team), and Rob Davies (U. of York) are external collaborators.

In this collaborative research work (mainly with the TRIO team), we look at the problem of scheduling periodic tasks on a single processor, in the case where each task's period is a (known) random variable. In this setting, some job will necessarily be missed, so that one will try to satisfy some criteria depending on the number of deadline misses.

We have proposed three criteria: (1) satisfying pre-defined deadline miss ratios, (2) minimizing the worst deadline miss ratio, and (3) minimizing the average deadline miss ratio. For each criterion we propose an algorithm that computes a provably optimal fixed priority assignment, i.e., a solution obtained by assigning priorities to tasks and executing jobs by order of priority.

This work has been presented in [26].

We also collaborate on other topics linked to real-time scheduling, as (i) on search algorithms for deterministic, but multiprocessor, problems [38], and (ii) on the problem of which jobs to drop (on-going work).

5.1.10. Adaptive Management with POMDPs

Participant: Olivier Buffet.

Iadine Chadès, Josie Carwardine, Tara G. Martin (CSIRO), Samuel Nicol (U. of Alaska Fairbanks) and Régis Sabbadin (INRA) are external collaborators.

In the field of conservation biology, adaptive management is about managing a system, e.g., performing actions so as to protect some endangered species, while learning how it behaves. This is a typical reinforcement learning task that could for example be addressed through BRL.

Here, we consider that a number of experts provide us with one possible model each, assuming that one of them is the true model. This allows making decisions by solving a mixed observability MDP (MOMDPs), where the hidden part of the state corresponds to the model (in cases where all other variables are fully observable).

We have conducted preliminary studies of this approach, using the scenario of the protection of the Gouldian finch, and focusing on the particular characteristics that could be exploited to more efficiently solve this problem. First results have been presented in [39].

5.1.11. Information Gathering with Sensor Systems

Participant: Olivier Buffet.

Elodie Chanthery, Matthieu Godichaud (LAAS-CNRS) and Marc Contat (EADS) are external collaborators.

The DOPEC project was a DGA PEA (upstream studies project) on the optimization of the use of sensor systems. In collaboration with EADS (project leader) and the LAAS, we have worked on autonomous sequential decision making problems. We were more particularly interested, on the one hand, in multi-agent problems and, on the other hand, in taking uncertainties into account.

The overall architecture that has been developed in the context of this project was presented in a national and an international conference [40], [23].

5.1.12. How do real rats solve non-stationary (PO)MDPs ?

Participant: Alain Dutech.

Etienne Coutureau and Alain Marchand (Centre de Neurosciences Intégratives et Cognitives (CNIC), UMR 5228, Bordeaux) are external collaborators.

For a living entity, using simultaneously various ways for learning models or representations of its environment can be very useful to adapt itself to *non-stationary environments* in a Reinforcement Learning setting. In the rats and in the monkey, two different action control systems lie in specific regions of the prefrontal cortex. Neurobiologists and computer scientists find here a common ground to identify and model these systems and the selection mechanisms between them, selection that could depend on uncertainty or error signals. Using real data collected on rats with or without prefrontal lesions, reinforcement learning models are used and evaluated in order to better understand this behavioral flexibility. MAIA is more particularly involved as a reinforcement learning expert in order to suggest and build models of the various learning mechanisms. In particular, we have used an *on-policy* learning scheme (SARSA) to investigate how well the use of simple or complex representations (with or without memory of the immediate past) can best model the learning behavior of rats in instrumental contingency degradation tasks [7].

This work has led us to investigate in more details the relations between the prefrontal cortex and the basal ganglia and their respective role when rats learn to solve non-stationnary tasks. The research is conducted through the PEPII project IMAVO (see 7.2.7).

5.1.13. Developmental Reinforcement Learning

Participants: Alain Dutech, Olivier Buffet.

Luc Sarzyniec and Joël Legrand (M2R Student of UHP Nancy 1) are external collaborators.

The goal of this work is to investigate how reinforcement learning can benefit from a developmental approach in the field of robotics. Instead of having a robot directly learn a difficult task using appropriate but rich (in the number of dimensions) sensory and motor spaces, we have followed an incremental approach. Both the number of perception and action dimensions increase only when the performance of the learned behavior increases. At the core of the algorithm lies a neuronal approximator used to compute the value function of the current policy of the robot. When the perception or action space grow, neurons or networks, initialized from existing neurons and networks, are added to the control architecture.

Thus far, our research focussed on the approximation architecture used to evaluate the *Q-function*. In simple robotic task, we investigated the use of Multi-Layer Perceptrons, either one approximation for every possible action ([41]) or one unique global approximator with as many outputs as the number of actions (Master Thesis of Joël Legrand). Currently, a *reservoir computing* architecture is under study as depicted in [16].

5.1.14. Classification-based Policy Iteration with a Critic

Participant: Bruno Scherrer.

Victor Gabillon, Alessandro Lazaric and Mohammad Ghavamzadeh (from Sequel INRIA-Lille) are external collaborators.

We study the effect of adding a value function approximation component (critic) to rollout classification-based policy iteration (RCPI) algorithms. The idea is to use the critic to approximate the return after we truncate the rollout trajectories. This allows us to control the bias and variance of the rollout estimates of the action-value function that are strongly related to the length of the rollout trajectories. Therefore, the introduction of a critic can improve the accuracy of the rollout estimates, and as a result, enhance the performance of the RCPI algorithm. We present in [49], [20] a new RCPI algorithm, called *direct policy iteration with critic* (DPI-Critic), and provide its finite-sample analysis when the critic is based on LSTD and BRM methods. We empirically evaluate the performance of DPI-Critic and compare it with DPI and LSPI in two benchmark reinforcement learning problems.

5.1.15. Linear Approximation of Value Functions

Participant: Bruno Scherrer.

Matthieu Geist (Supélec, Metz) is an external collaborator

In the framework of Markov Decision Processes, we consider the problem of learning a linear approximation of the value function of some fixed policy from one trajectory possibly generated by some other policy.

In [30], [42], [51], we describe a systematic approach for adapting on-policy learning least squares algorithms of the literature (LSTD, LSPE, FPKF and GPTD/KTD) to off-policy learning with eligibility traces. This leads to two known algorithms, LSTD(λ)/LSPE(λ) and suggests new extensions of FPKF and GPTD/KTD. We describe their recursive implementation, discuss their convergence properties, and illustrate their behavior experimentally. Overall, our study suggests that the state-of-art LSTD(λ) remains the best least-squares algorithm.

We also consider the task of feature selection. A promising approach consists in combining the Least-Squares Temporal Difference (LSTD) algorithm with ℓ_1-regularization, which has proven to be effective in the supervised learning community. This has been done recently whit the LARS-TD algorithm, which replaces the projection operator of LSTD with an ℓ_1-penalized projection and solves the corresponding fixed-point problem. However, this approach is not guaranteed to be correct in the general off-policy setting. In [21], we take a different route by adding an ℓ_1-penalty term to the projected Bellman residual, which requires weaker assumptions while offering a comparable performance. This comes at the cost of a higher computational complexity if only a part of the regularization path is computed. Nevertheless, our approach ends up to a supervised learning problem, which let envision easy extensions to other penalties.

5.2. Understanding and mastering complex systems

5.2.1. Complex systems : simulation, control and definition

5.2.1.1. Adaptive control of a complex system based on its multi-agent model

Participants: Vincent Chevrier, Tomas Navarrete.

Laurent Ciarletta (Madynes team, LORIA) is an external collaborator.

We are interested in how to build a control mechanism for a complex/dynamic system. Specifically, we want to evaluate the effectiveness of creating a control mechanism based on a multi-agent model of the system[12]. Multi-agent models can be adapted to that purpose since usual approaches using analytical models as basis can be untractable when dealing with such systems; and because if we consider that the available control actions are meant to be applied locally, a multi-agent model is necessary. We are currently working on a case study within the dynamic networks domain, namely the free-riding phenomenon present in peer-to-peer networks.

We propose an architecture that gathers information from the system and uses it to parametrize and tune a set of multi-agent models. The outcome of simulations is used to decide which control actions have to be applied to the system, in order to achieve a predefined control objective. We consider that we do not have complete information to characterize the state of the system and hence would like to focus on the following two issues of the control problem that we have identified:

1. How to build a multi-agent model that represents the evolution of a dynamic network. That is, what to do when the information given by the simulation of the multi-agent is in contradiction with the information gathered from the system
2. How to build an adaptive control mechanism based on the multi-agent model of a dynamic network. That is, how to use the information given by the multi-agent model to achieve the control objective.

The architecture we proposed, is designed as a control loop composed of the following steps: estimate the state of the system and instantiate multi-agent models accordingly, simulate different control actions, choose a control action and apply it. From one cycle to another of the control loop, each step can be tuned (in terms of model parameters, control action selection process, sampling strategy, etc.) to overcome the previously mentioned issues of the control problem.

The architecture is currently specified in terms of a formal notation. We have already implemented the architecture within the context of the free-riding problem where we use the PeerSim simulator as the target system to control.

Within our case study, we have conducted two different sets of experiments to investigate under which conditions our control architecture can achieve its goal and to investigate the efficiency of different sampling methods to estimate the state of the network. We have effectively managed to drive the system to a state where the majority of the peers share, when the initial conditions, without intervention from our architecture, would drive the system to a state where no peer would share.

The elements of the architecture having an impact on the performance of the control obtained have been identified. These are: the initialization of the parameters of the models used to estimate the state of the system, predict the evolution of the system and test the possible control actions, as well as the strategy used to observe the system and the different time horizons to consider within the architecture.

The next steps are to better identify the advantages and limits of the proposed architecture and to widen the problem family in the free riding problem.

5.2.1.2. Multi Modeling and multi-simulation

Participants: Vincent Chevrier, Julien Siebert.

This work is undertaken in a joint Phd Thesis between MAIA and Madynes Team. Laurent Ciarletta (Madynes team, LORIA) is co-advisor of this PhD.

Complex systems generally require to use different points of view (abstraction levels) at the same time on the system in order to capture and to understand all the dynamics and the complexity. Being made of different interacting parts, a model of a complex system also requires simultaneously modeling and simulation (M&S) tools from different scientific fields.

Building a model and a simulation of a complex system from the interaction of the different existing M&S tools present in each scientific field involved, is also a complex task. To represent a complex system, we need to couple several models (multi-modeling) that each represents a part of the whole system. Each model could have been designed by and for a specific scientific domain. Making different models interact raises hard issues on model interoperability (semantic coherence, formalism compatibility). As many simulators exist in the scientific fields involved, a possible approach to make a simulation of a complex system is to reuse and to make interact these existing simulators. Since each simulator has been developed for specific purposes, making them interact (multi-simulation) raises simulation issues (interoperability, synchronization).

The multi-agent paradigm is an homogeneous solution both for multi-modeling and multi-simulation of complex systems. On the one hand, a multi-agent model per se is a multi-model: a multi-agent model is made of interacting agent models and environment models. On the other hand, agent oriented software engineering (AOSE) allows designers to create complex softwares as a set of autonomous, heterogeneous and interacting softwares (i.e. as a multiagent system). Robustness, scalability, openness, modularity and interoperability are some of the properties that AOSE allows to achieve.

This work explores the contribution of multiagent paradigm to the fields of multi-modeling and multi-simulation of complex systems.

The first contribution of this work is to propose an homogeneous multiagent meta-model (called AA4MM[4]) that provides solutions both for multi-modeling and multi-simulation of complex systems by reusing existing and heterogeneous M&S tools . The core idea in AA4MM is to build a society of models, simulators and simulation softwares that solves the core challenges of multimodelling and simulation coupling in an homogeneous perspective. AA4MM has been implemented and used both for proof of concept and for a real case study. A proof of concept has been made by coupling different models together to develop a multi-model of a prey-predator model. This has permitted us to show both conceptual and operational properties of AA4MM such as interaction of heterogeneous models, modularity, interoperability.

This multiagent meta-model has been applied to model complex systems that are ubiquitous networks. Ubiquitous networks are highly dynamic computer networks that are composed of a great number of interacting and sometimes mobile nodes which can join or leave the system, interact together and where the environment plays a significant role either on radio communications or on the behavior of users. Modeling and simulation is the approach to evaluate these technologies or to build new ones.

5.2.1.3. Robustness of Cellular Automata and Reactive Multi-Agent Systems

Participants: Olivier Bouré, Vincent Chevrier, Nazim Fatès.

Our research on emergent collective behaviours focuses on robustness analysis, that is the behavioural resistance to perturbations in collective systems. We progressed in the knowledge of how to tackle this issue in the case of cellular automata (CA) and multi-agent systems (MAS).

We focused on the specific case of a perturbation of the updating scheme in CA, that is, changing the way cells are updated. Using similar ideas to the Influence-Reaction principle developed to resolve conflicts related to simultaneous actions, we created a new type of asynchronism, called beta-synchronism, which aims at disrupting the transmission of information about states between cells. We found out that the different types of asynchronism may induce radical change of behaviour for particular a value range of the synchrony rate [15].

More recently, our interest focused on a bio-inspired discrete dynamical system. Using the formalism of a subclass of cellular automata, lattice-gas CA, we study a model of swarming which displays qualitatively different behaviours under certain experimental conditions. We discussed these observations by relating them to the potential links with certain attributes of the model [48].

We studied a phase transition that occurs in the Greenberg-Hastings CA reaction-diffusion [5].

The density classification problem was taken as a typical framework for studying how decentralised computations can be carried out with simple cells. Although it is known that this problem can not be solved perfectly, we showed that using randomness provides a solution with an arbitrarily high success rate [17]. We also studied how to extend this result to the infinite-space case citerefPapier ???.

We studied the behaviour of the amoebae aggregation model [33] and applied the aggregation scheme on a robotic case (ALICE robots and Khepera III with Romea interactive table).

5.2.1.4. Ant algorithms for multi-agent patrolling

Participants: Olivier Simonin, François Charpillet, Olivier Buffet, Arnaud Glad.

We proposed in 2007 an ant algorithm, called EVAP, to deal with multi-agent patrolling, which is based on the marking and the evaporation of a digital pheromone. During the simulations carried out to measure the performances of EVAP, we identified that the system can self-organize towards stationary cycles (a periodic attractor). These cycles correspond to an Hamiltonian or quasi-Hamiltonian covering of the environment, which is an optimal or quasi-optimal solution to the multi-agent patrolling problem. We then established the mathematical proof that the system can stabilize only in cycles, one per agent, having the same length (cf. publication in ECAI'2008). Moreover, we introduced new heuristics in the agent behavior that improve dramatically the time for convergence, and we proved that under deterministic hypotheses the system always converges to stable cycles (these results have been published in SASO 2009, **AAMAS'10**). Results of 2011 are :

- Defense of Arnaud Glad's PhD. thesis (November 15th) synthesising theoretical and experimental studies of the EVAP algorithm. The writing of a journal article is also in progress.
- EVAP has been adapted to continuous space in the context of the SUSIE project, which consider the surveillance of an area with a set of autonomous aerial robots.

5.2.2. Multi-robot systems : swarm intelligence, cooperation, navigation

5.2.2.1. Multi-robot exploration and mapping

Participants: Olivier Simonin, François Charpillet, Antoine Bautin.

In the context of the ANR Cartomatic project, introduced in Sec. 7.2.3, we study multi-agent models for multi-robot deployment and mapping. This work is in line with the PhD thesis of Antoine Bautin, started in November 2009. New results of 2011 are

- A new frontier assignation algorithm for multi-robot exploration has been proposed. It relies on counting the number of robots towards a frontier rather than considering only distances between robots and frontiers. We measured on benchmarks that the approach outperforms the two classical algorithms *closest frontier* and *Greedy approach*. Results are prensented in [37], [43] and are submitted to ICRA'2012.
- We implemented and experimented the approach with autonomous mobile robots in the context of the ANR Carotte challenge (June 2011, Bourges). Our team "cartomatic" obtained one of the best map of the contest, while deploying several robots.

5.2.2.2. New experimental device: the Interactive Table

Participants: Olivier Simonin, François Charpillet, Nicolas Beaufort.

Olivier Rochel (INRIA research engineer, SED Nancy) is an external collaborator.

During 2010 we developed with the Nancy INRIA SED[6] (Olivier Rochel) a new experimental device dedicated to swarm robotics study. It is composed of two independent components : an interactive table able to display and to compute any active environment and a set of autonomous mobile robots able to read and write information on the environment.

Studies using the Table in 2011 are :

- We revisited the Drogoul & Ferber Foraging model, inspired by ants and also called "robot dockers" as the agents exchange the transported resources when they meet. From this simulated model we examined how it can be implemented with real mobile robots on an interactive environment, by considering that robots drop pheromones as ants. We defined a model extending the docker model with the robots on the Table, and studied its robustness to perception failure/mistakes. This work, done with Thomas Huraux (Master 2 Recherche internship), has been published in ICTAI 2011 Int. Conference [31].
- Several students (from Science Cog. Nancy 2 Master) implemented and explored pheromone-based foraging behaviors and flocking-based navigation models (supervized by François Charpillet and Christine Bourjot).

5.2.2.3. Local control based platooning

Participants: Alexis Scheuer, Olivier Simonin, François Charpillet, Jano Yazbeck.

[6] Service d'Expérimentation et Développement

We consider decentralised control methods to operate autonomous vehicles at close spacings to form a platoon. We study models inspired by the flocking approach, where each vehicle computes its control from its local perceptions. We investigate different decentralised models in order to provide robust and scalable solutions. Open questions concern collision avoidance, stability and multi-platoon navigation.

- *Coupling lateral and longitudinal controls.* A first work [67] focused on longitudinal control, which aims at computing velocities to avoid collision when all the vehicles are moving along a fixed path. When vehicles move in a two dimensional space, a lateral controller is needed to steer the vehicles. While lateral and longitudinal controls can be considered separately, the longitudinal control should be done after the lateral control: while turning, a higher inter-vehicle distance is needed to avoid collisions.

  An innovative approach to improve the quality of lateral control has been proposed during Jano Yazbeck's internship at LORIA (03/10–07/10), entitled "Decentralised local approach for lateral control of platoons" and supervised by A. Scheuer and O. Simonin. This allows to reduce the distance between each vehicle's path and the path of the previous vehicle, by using only embedded sensors such as a laser rangefinder. It relies on memorizing and computing in real time the previous vehicle relative trajectory. This work has been published in 2011 IEEE-RSJ International Conference on Intelligent Robots and Systems (**IROS'2011**) [34].
- *Finding an efficient lateral control.* To obtain an even better lateral control, and to drive each vehicle exactly in the trace of the previous one, we are developing a more efficient lateral control law. This law is defined in order to reduce exponentially the tracking error (which is more or less the distance between each vehicle's path and the path of the previous vehicle). Once again, as for the longitudinal control [67], the formula of the control law is obtained through the proof of its property: necessary conditions are simplified in order to get the final result.

5.2.2.4. Adaptation of autonomous vehicle traffic to perturbations

Participants: Mohamed Tlig, Olivier Simonin, Olivier Buffet.

In the context of the european InTraDE project, one problem is to handle the displacements of numerous IAVs[7] in a seaport. Here we assume a supervisor planning the routes of the vehicles in the port. However, in such a large and complex system, different unexpected events can arise and damage the traffic : failure of a vehicle, human mistake while driving, obstacle on roads, local re-planning, and so on.

We started focusing on a first important sub-problem of space resource sharing among multiple agents: how to ensure the crossing of two opposed flows of vehicles on a road when one of the two paths is blocked by an obstacle, e.g., a disabled vehicle. To overcome this problem, blocked vehicles have to coordinate with vehicles of the other side to share the road and manage delays. The objective is to improve traffic flow and reduce the emergence of traffic jam.

Solving this problem with reactive coordination methods is a major challenge of the PhD thesis of Mohamed Tlig (started in December 2010).

- We started by formalizing the problem and the possible actions of agents (vehicles) following a STRIPS formalism. We adapted this model dedicated to planning to the description of local rules in reactive coordination.
- We then defined and studied in simulation two decision rules that produce two different strategies: the first one alternates between two vehicles from each side of the road, and the second one gives priority to the vehicle with the highest delay. We are preparing a publication of these first results.

5.2.3. Ambient intelligence and Actimetry

5.2.3.1. Robotics and spatial computing : the iTiles - intelligent tiles - model

Participants: Olivier Simonin, François Charpillet, Lionel Havet.

[7] Intelligent Autonomous Vehicle

Olivier Rochel (INRIA research engineer, SED Nancy) is an external collaborator.

In the context of intelligent home and assistant robots, we explore the definition and use of an active floor based on a cellular network approach. We aim at exploring spatial calculus models when considering physical cells augmented with sensors where robots and humans can evolve. Since 2009, we study a model consisting in paving the floor with interconnected *tiles*. Each tile can communicate with its neighbors and can sense the presence of a robot or a human. A first Tile model has been defined and evaluated using a tiles emulator and real mobiles robots (Kheperas III), which validated the interest of the approach. See CAR'2010 publication [68].

In 2011 we designed with the help of INRIA Grenoble SED a prototype of 9 physical tiles embedding a WSN node able of computation and communication with other tiles. From this experimental device we explored several questions :

- How to follow a person walking on such a discrete and sensitive floor ? We proposed a set of distributed algorithms allowing tiles to track a person or a robot (cooperation between neighboring tiles).
- How to make communications between a robot and the tile(s) it occupies ? We developped a set of functions using the wifi communication of the tiles and the robot.
- From the work mentionned in the previous items we propose the definition of a new tile network based on the SensLab technology (wire connections between tiles and wireless communications between tiles and robots/humans). This prototype have been ordered and will be installed in the begining of 2012.

5.2.3.2. Bayesian 3D Human Motion Capture Using Factored Particle Filtering

Participants: Abdallah Dib, Cédric Rose, Amandine Dubois, François Charpillet.

The gait deterioration of elderly people is an important factor in loss of autonomy and it increases the risk of falls. In order to evaluate this risk the MAIA team has been developing since 2003 a markerless human motion capture system that estimates the 3D positions of the body joints over time. The system uses a dynamic Bayesian network and a factored particle filtering algorithm. This year, we have evaluated the impact of using different observation functions for the Bayesian state estimation: chamfer distance, a pixel intersection and finally a pseudo-observation of the subject direction calculated from the previous output of the system. We also compared two methods for the factored generation of the particles. The first one uses a deterministic interval exploration strategy whereas the second one is based on an adaptive diffusion. The capacity of the system to recover after occlusion by obstacles was tested on simulated movements in a virtual scene [57].

An other achievement of the year has been the assessment of the accuracy and precision of this system, especially for measuring the step length of a walking human. This has been realized by Amandine Dubois during her research master [58]. An experiment with young subjects has been designed and realized. Measures of the markerless motion capture system were then compared with real values. These values were obtained through the footprints left by the subjects. Ink swabs placed at the front and rear of the shoes of each subject make it possible to mark a paper strip positioned on the ground. A statistical analysis of the results has been done by Amandine. Thus we were able to determine if the real and measured lengths were significantly different or not.

5.2.3.3. Automatic Evaluation of Vascular Access in Hemodialysis Patients

Participants: Cédric Rose, François Charpillet.

The vascular access that allows to perform the extra-corporeal circulation, is usually a vein of the arm that has been enlarged by a surgical creation of a fistula. The prevention of complications such as stenosis or thrombosis of the vascular access is a key issue in hemodialysis treatment. Many dialysis machines measure ionic dialysance by conductivity measures on the dialysate fluid. Ionic dialysance is an indicator of small molecules transfers through the dialysis membrane. Previous works have shown that the follow-up of the dialysance and the pressures along the extra corporeal circuit can help to detect at an early stage a potential complication on the vascular access. The difficulty of automating the follow-up is the large variability of the measures and the need to detect tendencies. Dynamic Bayesian networks (DBN) allow to formalize expert knowledge as a graphical stochastic model adapted to reasoning under uncertainty. In a DBN the state of the patient and the measurements are represented by interconnected temporal random variables. The relations between those variables are described using probability distributions. The proposed approach [64] is based on a supervised learning of a DBN for classifying the dialysis sessions according to a risk score describing the medical situation (0: no risk, 1: mild risk, 2: severe risk). The training of the system was performed using a dataset labeled by a medical expert. The evaluation of the results was done by performing a double-blind analysis of real data. The result was an 85% agreement rate between the human expert and the automated analysis. The purpose of the system is to assist the human expert by reporting abnormalities. The results show that a score 2 reported by the human is rarely missed by the automated analysis (only 1 case) whereas the opposite is more frequent (8 cases). The final decision to further investigate a case is taken by the human expert.

6. Contracts and Grants with Industry

6.1. Contracts with Industry

6.1.1. Consulting for Core Security Technologies

Participant: Jörg Hoffmann.

Core Security Technologies is an U.S.-American/Argentinian company providing, amongst other things, tools for (semi-)automated security checking of computer networks against outside hacking attacks. For automation of such checks, a module is needed that automatically generates potential attack paths. Since the application domain is highly dynamic, a module allowing to declaratively specify the environment (the network and its configuration) is highly advantageous. For that reason, Core Security Technologies have been looking into using AI Planning techniques for this purpose. After consulting by Jörg Hoffmann (see also below), they are now using a variant of Jörg Hoffmann's FF planner – developed during his PhD work at the University of Freiburg, Germany (cf. Section 4.1) – in their product. Core Security Technologies payed Jörg Hoffmann for the consulting as an Auto-Entrepreneur.

7. Partnerships and Cooperations

7.1. Regional Initiatives

7.1.1. COMAC

Participants: Mauricio Araya, Marie Tonnelier, Vincent Thomas, Olivier Buffet, François Charpillet.

Laurent Bougrain (CORTEX team, LORIA) is an external collaborator.

The COMAC[8] project is part of the Materalia competitive cluster. The main objective of the project is to develop diagnosis tools for the low cost identification of defaults in aeronautics parts made of composite materials.

In collaboration with Laurent Bougrain, one of our objectives is to propose a software toolbox for computer-aided diagnosis in this context. The current project is a system relying on expert knowledge taking the form of a database of labeled images.

[8] COMAC = *contrôle optimisé multi-techniques des aérostructures composites* / optimised multi-technique control of composite aeronautic parts

In the MAIA team, our research effort focuses more precisely on information gathering problems involving active sensors, i.e., an intelligent system which has to select the observations to perform (which sensor, where, at which resolution). Mauricio Araya's undergoing PhD looks precisely at the topic of Active Sensing (Section 5.1.8).

7.1.2. Multi-agent simulation of public transportation

Participant: Vincent Chevrier.

This collaboration with the CUGN (communauté urbaine du grand Nancy - Pole Transport) aims at a better understanding of the functioning of the transportation of the Grand-Nancy. A first part of the work aims at providing an accurate and meaningful understanding of the transportation system. Through student projects we propose different viewpoints to enhance this understanding. After validation of the Pole Transport, some of these viewpoints have been integrated in tools to produce daily report at the Pole Transport.

A second part is dedicated to the explanation of the dynamics of transportation systems. We are developing a multi-agent model of the tramway line which integrates real data (traveling time during stops). We are able to reproduce an equivalent functioning without perturbation. Recently, we showed that if we perturb the system, it responds similarily as the real system. For example, halting a tramway at a stop, figuring a lot of people being waiting and taking it, induces a comparable response in the simulator regarding to the real system.

7.2. National Initiatives

7.2.1. ANR project BARQ

Participants: Jörg Hoffmann, Olivier Buffet, Bruno Scherrer.

This project has been granted by ANR in the "Chaires d'Excellence" program. The project is funded with ca. 400000 EUR and will hire four non-permanent researchers (Doctorants and/or Postdocs). Jörg Hoffmann is the project leader, Olivier Buffet and Bruno Scherrer collaborate. Other collaborators from LORIA are Stephan Merz, Ammar Oulamara, and Martin Quinson. The project also has several international collaborators, in particular Prof. Blai Bonet (Universidad Simon Bolivar, Caracas, Venezuela), Prof. Carmel Domshlak (Technion Haifa, Israel), Prof. Hector Geffner (Universitat Pompeu Fabra, Barcelona, Spain), Dr. Malte Helmert (University of Freiburg, Germany), and Prof. Stephen Smith (CMU, Pittsburgh, USA).

The project unites research from four different areas, namely classical planning, probabilistic planning, model checking, and scheduling. The underlying common theme is the development of new methods for computing lower bounds via state aggregation. Specifically, the basic technique investigated allows explicit selection of states to aggregate, in exponentially large state spaces, via an incremental process interleaving it with state space re-construction steps. The two main research questions to be addressed are how to choose the states to aggregate, and how to effectively obtain, in practical scenarios, anytime methods providing solutions with increasingly tighter performance guarantees.

So far, we have hired Dr. Michael Katz as a PostDoc (for 2 years) working on classical planning, and Manel Tagorti as a PhD student (for 3 years) working on probabilistic planning. The Conseil Regional de Lorraine has accepted to co-finance, for 2011, 50% of the the position of Michael Katz for a period of 1 year. Chao-Wen Perng was funded from BARQ for an internship of 5 months during which she worked on her MSc report, laying some basis for the research direction to be followed by Manel Tagorti.

7.2.2. PEPS project GEST - 2010/2011

Participant: Vincent Chevrier.

This project "Gouvernance Enactive des Systèmes de Transports" (GEST) is the consequence of the work undertaken within the GEST project funded by the IXXI ("Institut Rhône Alpin des Systèmes Complexes").

It involves teams from the LIG (Laboratoire d'informatique de Grenoble) and from the LIESP (Laboratoire d'informatique pour l'entreprise et les systèmes de production), and is associated to the CUGN.

This project aims at a fundamental level at proposing an enactive perspective for the governance issue in case of complex socio-technical systems, and more specifically, in case of public transportation systems. From a more applicative perspective, we seek at specifying a participatory and reflexive simulation system based on a multi-agent model.

This exploratory project is grounded on core ideas coming from the IXXI work. It aims at gathering researchers coming from different domains (social cognition, decision theory, simulation, serious game, etc) in order to clarify interdisciplinary issues.

A workshop was organized in the beginning of January 2011 in Lyon followed by three other meetings.

7.2.3. ANR project CARTOMATIC

Participants: Olivier Simonin, François Charpillet, Antoine Bautin.

This project has been granted by ANR in the Robotics Carotte challenge (CArtographie par ROboT d'un TErritoire) from the *Contenus et Interactions* program. The project is funded with ca. 50000 EUR to purchase the robotics platform. The Maia team was also funded with a PhD fellowship. The Cartomatic consortium is formed by LISA/Angers University (leader), Maia/LORIA and Wany robotics (Montpellier).

This project concerns the mapping of an indoor structured but unknown environment, and the localization of objects, with one or several robots. We aim at studying multi-robot or swarm algorithms to achieve such a challenge, while showing the robustness and the accuracy of the mapping when using cooperation between several autonomous robots. In the work of Antoine Bautin PhD thesis (started in 2009) we have proposed a distributed algorithm XXXXXXXX multi-robot deployment and mapping. (Section 5.2.2.1).

7.2.4. INRIA AEN PAL Personally Assisted Living

Participants: François Charpillet, Olivier Simonin.

The PAL project is an INRIA National Initiative (Action d'Envergure Nationale) involving several teams of the institute (Arobas, Coprin, E-motion, Lagadic, Demar, Maia, Prima, Pulsar and Trio). It is coordinated by David Daney (INRIA Sophia-Antipolis EPI Coprin). The project focuses on the design of smart home and robot assistant. Maia is particularly involved in the People Surveillance work package, by studying and experimenting intelligent environments, funded on smart tiles (cf. Sec. 5.2.3.1) and multi-sensor devices.

As a first action, the consortium organized a Workshop "De l'Autonomie au Domicile" in Nice, on November 29 & 30, 2010. http://www-sop.inria.fr/coprin/pal/workshop/

7.2.5. PEA-DGA SUSIE 2009-12

Participants: François Charpillet, Olivier Simonin, Romain Mauffray.

This project relies on results and questions arising from the SMAART project (2006-08). During this project we adapted the EVAP algorithm 5.2.1.4 to the patrol with UAVs, while providing a generic digital pheromone based patrolling simulator Concerning sharing authority, we proposed an original interface to manipulate groups of UAVs. However, experiments with operators have shown that they succeed in improving the whole system when dealing with the patrolling task.

So, the aim of the SUSIE project is twofold: (i) studying and improving parameters of the EVAP algorithm through the SMAART simulator, (ii) defining new ways to manipulate pheromones fields in order to improve the sharing authority.

7.2.6. INRIA ADT project ROMEA (2009-11)

Participants: Olivier Simonin, François Charpillet, Nicolas Beaufort, Alain Dutech, Olivier Buffet.

ROMEA, for "RObots Mobiles et Environnements Actifs", is a project proposed by the Maia team and funded by INRIA NGE through an ADT "Action de Developpement Technologique". The project deals with the development and the study of intelligent and collective behaviors with Khepera III mobile robots. In particular we develop a new experimental device, called *interactive table for robots*, which provides a graphical active environment where robots can read and write pieces of information (e.g. digital pheromones). During 2010, with O. Rochel (INRIA research engineer) and N. Beaufort (INRIA IJD), we designed such a device which is now used for swarm robotics experiments, see Section 5.2.2.2. Nicolas Beaufort was hired for a second year as an INRIA IJD engineer to develop the required functions on the interactive table and the robots.

7.2.7. CNRS PEPII project "IMAVO" (2011-2012)

Participant: Alain Dutech.

IMAVO, for "Interactions entre Modules pour l'Apprentissage dans un environnement VOlatile", is a PEPII project of the INSB institute of the CNRS. It involves Alain Marchand and Etienne Coutureau from the INCIA Lab of Bordeaux (Behavioral Neurosciences - INSB), Mehdi Khamassi and Benoît Girard from the ISIR Lab of Paris (Robotics and Neurosciences - INS2I), Alain Dutech and Nicolas Rougier from the Loria Lab of Nancy (Computational Neurosciences and Machine Learning - INS2I).

This project investigates *model-based* and *model-free* reinforcement learning approaches for rats learning in volatile environments (*i.e.* context and reward can change during learning). It aims at designing new concept for modularized decision-making systems, allowing a better understanding of the underlying neuro-biological process involved in rats and humans and applications in the field of autonomous robotics.

It is a logical extension of a cooperation between MaIA and the INCIA Lab of Bordeaux [7].

7.3. European Initiatives

7.3.1. Collaborations in European Programs, except FP7

7.3.1.1. Agent Technical Fora

The Agent Technical Fora have been created by AgentLink III. and consist of several working groups called Technical Forum Groups. These groups of researchers and developers share an interest in a specific sub-area of agent and multi-agent technology. Since the end of AgentLink, the Technical Fora have been organized jointly to the EUMAS conference, starting in 2006.

7.3.1.1.1. Technical Forum Group: "Self-Organization"

Participants: Christine Bourjot, Vincent Chevrier, Vincent Thomas.

Vincent Chevrier was promoter of the Technical Forum Group "Self Organization". The aim of the TFG is to work on self-organization in complex distributed systems such as multi-agent systems. The group members have been involved in the writing of a book entitled *Self-Organising Software - From Natural to Artificial Adaptation* where the MAIA team is responsible for two chapters [44], [45].

7.3.1.1.2. Technical Forum Group: Simulation

Participants: Vincent Chevrier, Julien Siebert.

The Simulation Technical Forum meets for the first time this year. It aims at working on the main challenges of agent and multi-agent-based simulation while establishing links between members of the simulation community which could lead to share common research activities and projects.

The promoters of this forum solicited MAIA members to present their point of view on current issues in multiagent simulation.

7.3.1.2. European project INTERREG IVB "InTraDE" (2009-12)

Participants: François Charpillet, Alexis Scheuer, Olivier Simonin, Olivier Buffet.

The InTraDE project (Intelligent Transportation for Dynamic Environments, http://www.intrade-nwe.eu/) is funded by the European North West Region. The project is coordinated by Rochdi Merzouki from University of Science and Technology of Lille (LAGIS lab.). Other partners are the Maia team, Liverpool John Moores University (LOOM), the National Institute for Transport and Logistics in Dublin Institute of Technology, the South East England Development Agency, the AGHO Port of Oostende and the CRITT in Le Havre. In the context of seaports and maritime terminals, the InTraDE project aims to improve the traffic management and space optimization inside confined spaces by developing a clean and safe intelligent transportation system. This transportation system will operate in parallel with virtual simulation software of the automated site, allowing a robust and real-time supervision of the goods handling operation.

The Maia team partner focuses on decentralized approaches to deal with the control of automated vehicle platooning and the adaptation of the traffic. Maia is funded with two PhD fellowships and one engineer. Both PhD thesis started in the end of 2010. The PhD of Jano Yazbeck, supervised by F. Charpillet and A. Scheuer, aims at studying a “Secure and robust immaterial hanging for automated vehicles”. The PhD of Mohamed Tlig, supervised by O. Simonin and O. Buffet, addresses “Reactive coordination for traffic adaptation in large situated multi-agent systems”.

7.3.2. Major European Organizations with which you have followed Collaborations

University of Basel, Departement Informatik (Switzerland)
New lower-bound heuristics for deterministic planning

Partner 2: organisme 2, labo 2 (pays 2)
Sujet 2 (max. 2 lignes)

7.4. International Initiatives

7.4.1. Visits of International Scientists

- Dr. Sylvie Thiébaux, Director of the Canberra Research Laboratory of NICTA (Australia), visited MAIA for 1 week in June 2011.
- Dr. Ingo Weber, Senior Research Associate at the University of New South Wales (Sydney, Australia), visited MAIA for 1 day in August 2011.
- Dr. Iadine Chadès, Research Scientist at CSIRO, Ecosystem Sciences division (Brisbane, Australia), visited MAIA for 1 week in September 2011.
- Dr. Van Parunak, Chief Scientist at Vector Research Center, Ann Arbor, USA, visited MAIA for 2 days at the occasion of Olivier Simonin’s HDR defense. December 2010.

8. Dissemination

8.1. Animation of the scientific community

8.1.1. Conference organization, Program committees, Editorial boards

- Olivier Buffet is a member of the editorial board of both the “revue d’intelligence artificielle” (RIA) and the “Journal of Artificial Intelligence Research” (JAIR).
- Olivier Buffet was a reviewer for the journals: JAIR (Journal of Artificial Intelligence Research), IJAR (International Journal on Approximate Reasoning), RIA (Revue d’Intelligence Artificielle); and for the conferences IJCAI’11 (International Joint Conference on Artificial Intelligence), AAAI’11 (National Conference on Artificial Intelligence), UAI’11 (Uncertainty in Artificial Intelligence), ICRA’12 (International Conference on Robotics and Automation), IROS’11 (International Conference on Intelligent Robots and Systems), JFPDA’11 (Journées Francophones sur la Planification, la Décision et l’Action pour le contrôle de systèmes), Solstice’10 (Summer Solstice International Conference on Discrete Models of Complex Systems).

- Vincent Chevrier was a reviewer for the journals: ACM TAAS (Transactions on Autonomous and Adaptive Systems), JAAMAS (Journal of Autonomous Agents and Multi-Agent System);
- Vincent Chevrier was a member of:
  - the editorial board of Interstices [9], a site to disseminate research work about computer science for French-speaking person,
  - the advisory board of JFSMA (JournÈes Francophones sur les SystËmes Multi-Agents)
  - the program committee of EUMAS 11 (European Workshop on Multi-Agent Systems) , IAT 11 (Intelligent Agent Technology), RFIA12 (Reconnaissance des Formes et Intelligence Artificielle).
- Vincent Chevrier is the moderator of the mailing list of the French spoken community on multi-agent systems.
- Alain Dutech was a reviewer for JFPDA'11 (Journées Francophones de Planification, Décision, Apprentissage) and UAI'11 (Uncertainty in Artificial Intelligence).
- Jörg Hoffmann is an Associate Editor of the Journal of Artificial Intelligence Research (JAIR) and Area Chair for Planning of AI Communications. In 2011, he was a member of the Senior Program Committees of AAAI'11 (the American Conference on Artificial Intelligence), IJCAI'11 (International Joint Conference on Artificial Intelligence), and ICAPS'11 (International Conference on Automated Planning and Scheduling). He was a member of the Program Committees of SOCS'11 (International Symposium on Combinatorial Search) and KI'11 (German Conference on Artificial Intelligence).
- Jörg Hoffmann has been appointed to be Program Chair (together with Prof. Bart Selman, Cornell University, USA) for AAAI'12, the 26th edition of the AAAI series which is one of the two most important international conferences in the area.
- Alexis Scheuer was a reviewer for the journal MENG (Recent Patents on Mechanical Engineering, Bentham Science Publishers), as well as for the last two more important conferences in robotics, IROS'11 (IEEE-RSJ International Conference on Intelligent Robots and Systems) and ICRA'12 (International Conference on Robotics and Automation).
- Olivier Simonin was a program committee member of SASO'2011 (Fifth IEEE International Conference on Self-Adaptive and Self-Organizing Systems), ICINCO'2011 (9th Int. Conf. on Informatics in Control, Automation and Robotics), ICAART'2012 (4th Int. Conf. on Agents and AI.) and JFSMA'2011 (French conference on MAS). He also reviewed papers for IROS'2011 (International Conference on Intelligent Robots and Systems) and ICRA'2012 (International Conference on Robotics and Automation). In 2012 he will be Chair Man of the Contest track of SASO'2012.
- Vincent Thomas was a reviewer for ISA2011 (Intelligent Systems and Agents) and IJCAI2011 (International Joint Conference on Artificial Intelligence).

8.1.2. Specialist Committees (commissions de spécialistes)

- Vincent Chevrier was a member of the "Specialist Committees" in Université Henri Poincaré

8.1.3. PhD and HDR committees

- Olivier Buffet was a member of the following PhD committees:
  - (as a committee member) Arnaud Glad, *Etude de l'auto-organisation dans les algorithmes de patrouille multi-agent fondés sur les phéromones digitales*, 15 Nov. 2011, Univ. Nancy 2;

[9] http://interstices.info

  - (as a committee member) Romain Grasset-Bourdel, *Planification et replanification pour une constellation de satellites agiles d'observation de la Terre*, 30 Nov. 2011, Univ. de Toulouse;
  - (as a committee member) Boris Lesner, *Planification et apprentissage par renforcement avec modèles d'actions compacts*, 8 Dec. 2011, Univ. de Caen.
- Vincent Chevrier was a member of the following PhD and HDR committees:
  - as a referee: Laurent Poligny, de l'ordinateur à la machine: création et régulation du traitement des données, 10 janv. 2011, Univ Evry; Richard Moussa Segmentation multi-agent en imagerie biologique et médicale: application aux IRM 3D, 12 Dec 2011, Univ. Bordeaux 1.
- Jörg Hoffmann was a member of the PhD committee of Etesham Zahoor (10 Nov 2011).

8.1.4. Scientific Diffusion

- Jörg Hoffmann was an Invited Speaker at KI'11, the 34th Annual German Conference on Artificial Intelligence. The title of his (45 min) talk was "Everything You Always Wanted to Know About Planning (But Were Afraid to Ask)". He also published an invited paper on this subject, in the KI'11 proceedings [24].
- Jörg Hoffmann was an Invited Speaker at the 2011 International Summer School on Automated Planning and Scheduling. The title of his (90 min) talk was "Evaluating Planning Algorithms, or: How to Count Sheep?".

8.2. Teaching

Master : " Introduction à la robotique mobile ", 37.5 HETD (A. Scheuer & O. Simonin), M1 Informatique, Univ. Henri Poincaré, Nancy 1, France.

Master : "Apprentissage Numérique", 25 HETD (A. Dutech), M1 Sciences Cognitives, Univ. Nancy 2, France.

Master : "Agent Intelligent", 20 HETD (V. Thomas), M1 Sciences Cognitives, Univ. Nancy 2, France.

Master : "Game Design et Serious Game", 20 HETD (V. Thomas), M2 Sciences Cognitives, Univ. Nancy 2, France.

Master : " Intelligence Artificielle pour la robotique mobile ", 22.5 HETD (A. Scheuer & O. Simonin), M2 RAR (" Reconnaissance, Apprentissage, Raisonnement ", *i.e.*, Recognition, Learning, Thinking), Univ. Henri Poincaré, Nancy 1, France.

Master : "Optimisation et Systèmes Dynamiques Stochastiques", 15 HETD (V. Thomas), M2 Informatique, Univ. Henri Poincaré, Nancy 1, France.

PhD & HdR

PhD in progress : Mauricio Araya-Lòpez, "Contrôle optimal de capteurs actifs", Sept. 2009, F. Charpillet (advisor), O. Buffet, V. Thomas.

PhD in progress : Jeannot Yazbeck, "Secure and robust immaterial hanging for automated vehicles", Oct. 2010, F. Charpillet (advisor), A. Scheuer.

PhD in progress : Antoine Bautin, "Cartographie et recherche d'objets par une flotille de robots mobiles autonomes", Nov. 2009, F. Charpillet (advisor), O. Simonin.

PhD in progress : Arsène Fansi Tchango, "Suivi multi-caméras en environnement partiellement observé", Oct. 2011, A. Dutech (advisor), O. Buffet, V. Thomas.

PhD in progress : Manel Tagorti, "Approximating the Value Function for Heuristic Search in Factored MDPs", Nov. 2011, J. Hoffmann (advisor), B. Scherrer, O. Buffet.

PhD in progress : Mohamed Tlig, "Reactive coordination for traffic adaptation in large situated multi-agent systems", Dec. 2010, O. Simonin (advisor), O. Buffet.

9. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] G. CORONA. *Utilisation de croyances heuristiques pour la planification multi-agent dans le cadre des Dec-POMDP*, Université Henri Poincaré - Nancy I, April 2011, http://hal.inria.fr/tel-00598689/en.

[2] A. GLAD. *Etude de l'auto-organisation dans les algorithmes de patrouille multi-agent fondés sur les phéromones digitales*, Université Nancy II, November 2011, http://hal.inria.fr/tel-00646293/en.

[3] C. ROSE. *Modélisation stochastique pour le raisonnement médical et ses applications à la télémédecine*, Université Henri Poincaré - Nancy I, May 2011, http://hal.inria.fr/tel-00598564/en.

[4] J. SIEBERT. *Approche multi-agent pour la multi-modélisation et le couplage de simulations. Application à l'étude des influences entre le fonctionnement des réseaux ambiants et le comportement de leurs utilisateurs.*, Université Henri Poincaré - Nancy I, September 2011, http://hal.inria.fr/tel-00642034/en.

Articles in International Peer-Reviewed Journal

[5] H. BERRY, N. FATÈS. *Robustness of the Critical Behaviour in the Stochastic Greenberg-Hastings Cellular Automaton Model*, in "International Journal of Unconventional Computing", 2011, vol. 7, n^{o} 1-2, p. 65-85, http://hal.inria.fr/inria-00598379/en.

[6] C. CAPPELLE, M. EL BADAOUI EL NAJJAR, F. CHARPILLET, D. PORMSKI. *Virtual 3D City Model for Navigation in Urban Areas*, in "Journal of Intelligent and Robotic Systems", June 2011, vient d'etre accepté [*DOI :* 10.1007/S10846-011-9594-0], http://hal.inria.fr/hal-00599703/en.

[7] A. DUTECH, E. COUTUREAU, A. MARCHAND. *A reinforcement learning approach to instrumental contingency degradation in rats*, in "Journal of Physiology - Paris", 2011, vol. in Press [*DOI :* 10.1016/J.JPHYSPARIS.2011.07.017], http://hal.inria.fr/hal-00642715/en.

[8] N. FATÈS. *Stochastic Cellular Automata Solve the Density Classification Problem with an Arbitrary Precision - When randomness helps computing*, in "Theory of Computing Systems", 2012, Accepted, long version of the Stacs 2011 proceedings paper, http://hal.inria.fr/inria-00608485/en.

[9] J. HOFFMANN. *Analyzing Search Topology Without Running Any Search: On the Connection Between Causal Graphs and h+*, in "Journal of Artificial Intelligence Research", 2011, vol. 41, p. 155–229 [*DOI :* 10.1613/JAIR.3276], http://hal.inria.fr/inria-00598248/en.

[10] M. STOLLBERG, J. HOFFMANN, D. FENSEL. *A caching technique for optimizing automated service discovery*, in "International Journal of Semantic Computing", 2011, vol. 5, n^{o} 1, p. 1-31 [*DOI :* 10.1142/S1793351X11001146], http://hal.inria.fr/inria-00598250/en.

[11] N. VLASSOPOULOS, N. FATÈS, H. BERRY, B. GIRAU. *Large-scale Simulations on FPGAs: Finding the Asymptotic Critical Threshold of the Greenberg-Hastings Cellular Automata*, in "Journal of Cellular Automata", 2012, http://hal.inria.fr/hal-00644660/en.

Articles in National Peer-Reviewed Journal

[12] T. NAVARRETE GUTIERREZ, J. SIEBERT, L. CIARLETTA, V. CHEVRIER. *Impact des dimensions spatiale et temporelle dans la modélisation d'un phénomène collectif de type free-riding*, in "Revue d'Intelligence Artificielle", 2011, vol. 25, n^o 5, p. 625–651 [*DOI :* 10.3166/RIA.25.625-651], http://hal.inria.fr/hal-00640079/en.

Invited Conferences

[13] N. FATÈS. *Calculer en parallèle avec des systèmes spatialement distribués : automates cellulaires et systèmes multi-agents comme modèle ?*, in "1ère Conférence Internationale sur les Systèmes d'Information et Technologies - ICIST 2011", Tebessa, Algeria, 2011, http://hal.inria.fr/hal-00644119/en.

International Conferences with Proceedings

[14] M. ARAYA-LÓPEZ, O. BUFFET, V. THOMAS, F. CHARPILLET. *Active Learning of MDP Models*, in "European Workshop On Reinforcement Learning", Athène, Greece, September 2011, http://hal.inria.fr/hal-00642909/en.

[15] O. BOURÉ, N. FATÈS, V. CHEVRIER. *Robustness of Cellular Automata in the Light of Asynchronous Information Transmission*, in "10th International Conference on Unconventional Computing", Turku, Finland, C. S. CALUDE, J. KARI, I. PETRE, G. ROZENBERG (editors), Lecture Notes in Computer Science, Springer, 2011, vol. 6714, p. 52-63 [*DOI :* 10.1007/978-3-642-21341-0_11], http://hal.inria.fr/inria-00563904/en.

[16] A. DUTECH. *Dynamic reservoir for developmental reinforcement learning*, in "Workshop on Development and Learning in Artificial Neural Networks (DevLeaNN)", Paris, France, T. KOWALIW, N. BREDECHE, R. DOURSAT (editors), 2011, p. 37-40, http://hal.inria.fr/inria-00633433/en.

[17] N. FATÈS. *Stochastic Cellular Automata Solve the Density Classification Problem with an Arbitrary Precision*, in "STACS 2011", Dortmund, Germany, T. SCHWENTICK, C. DÜRR (editors), Leibniz International Proceedings in Informatics (LIPIcs), Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2011, vol. 9, p. 284–295 [*DOI :* 10.4230/LIPIcs.STACS.2011.284], http://hal.inria.fr/inria-00521415/en.

[18] N. FATÈS. *Stochastic Cellular Automata Solve the Density Classification Problem with an Arbitrary Precision*, in "Symposium on Theoretical Aspects of Computer Science - STACS2011", Dortmund, Germany, 2011, vol. 9, p. 284-295, http://hal.inria.fr/hal-00573619/en.

[19] N. FATÈS, N. VLASSOPOULOS. *A Robust Aggregation Method for Quasi-blind Robots in an Active Environment*, in "ICSI 2011", Cergy, France, 2011, http://hal.inria.fr/inria-00578870/en.

[20] V. GABILLON, A. LAZARIC, M. GHAVAMZADEH, B. SCHERRER. *Classification-based Policy Iteration with a Critic*, in "International Conference on Machine Learning (ICML)", Seattle, United States, ACM, June 2011, p. 1049-1056, http://hal.inria.fr/hal-00644935/en.

[21] M. GEIST, B. SCHERRER. *l1-penalized projected Bellman residual*, in "European Wrokshop on Reinforcement Learning (EWRL 11)", Athens, Greece, 2011, http://hal.inria.fr/hal-00644507/en.

[22] B. GIRAU, N. VLASSOPOULOS. *Tiled cellular automata for area-efficient distributed random number generators*, in "1st International conference on pervasive and embedded computing and communication systems - PECCS 2011", Vilamoura, Portugal, March 2011, http://hal.inria.fr/inria-00585495/en.

[23] M. GODICHAUD, E. CHANTHERY, O. BUFFET, M. CONTAT. *Formalizing and Solving Information Collection Problems with Autonomous Sensor Systems*, in "18th IFAC World congress", Milan, Italy, 2011, http://hal.inria.fr/inria-00593465/en.

[24] J. HOFFMANN. *Everything You Always Wanted to Know About Planning (But Were Afraid to Ask)*, in "Proceedings of the 34th Annual German Conference on Artificial Intelligence (KI'11)", Berlin, Germany, 2011, http://hal.inria.fr/hal-00641851/en.

[25] J. HOFFMANN. *Where Ignoring Delete Lists Works, Part II: Causal Graphs*, in "21st International Conference on Automated Planning and Scheduling", Freiburg, Germany, 2011, http://hal.inria.fr/inria-00578653/en.

[26] D. MAXIM, O. BUFFET, L. SANTINELLI, L. CUCU-GROSJEAN, R. DAVIS. *On the Optimality of Priority Assignment for Probabilistic Real-Time Systems*, in "19th International Conference on Real-Time and Network Systems - RTNS 2011", Nantes, France, September 2011, http://hal.inria.fr/inria-00614480/en.

[27] D. MAXIM, O. BUFFET, L. SANTINELLI, L. CUCU-GROSJEAN, R. DAVIS. *On the Optimality of Priority Assignment for Probabilistic Real-Time Systems*, in "the 19th International Conference on Real-Time and Network Systems (RTNS2011)", Nantes, France, September 2011, http://hal.inria.fr/hal-00646579/en.

[28] R. NISSIM, J. HOFFMANN, M. HELMERT. *Computing Perfect Heuristics in Polynomial Time: On Bisimulation and Merge-and-Shrink Abstraction in Optimal Planning*, in "22nd International Joint Conference on Artificial Intelligence (IJCAI'11)", Barcelona, Spain, July 2011, http://hal.inria.fr/inria-00592438/en.

[29] C. SARRAUTE, O. BUFFET, J. HOFFMANN. *Penetration Testing == POMDP Solving?*, in "Workshop on Intelligent Security", Barcelona, Spain, July 2011, http://hal.inria.fr/inria-00607403/en.

[30] B. SCHERRER, M. GEIST. *Recursive Least-Squares Learning with Eligibility Traces*, in "European Wrokshop on Reinforcement Learning (EWRL 11)", Athens, Greece, 2011, http://hal.inria.fr/hal-00644511/en.

[31] O. SIMONIN, T. HURAUX, F. CHARPILLET. *Interactive Surface for Bio-inspired Robotics, Re-examining Foraging Models*, in "23rd IEEE International Conference on Tools with Artificial Intelligence (ICTAI)", Boca Raton, United States, IEEE, November 2011, http://hal.inria.fr/inria-00617155/en.

[32] O. SIMONIN, A. LANOIX, A. SCHEUER, F. CHARPILLET. *Specifying in B the Influence/Reaction Model to Study Situated MAS: Application to vehicles platooning*, in "V2CS First International workshop on Verification and Validation of multi-agent models for complex systems", Paris, France, November 2011, http://hal.inria.fr/hal-00646288/en.

[33] N. VLASSOPOULOS, N. FATÈS. *Clustering Behavior of a Bio-inspired Decentralized Aggregation Scheme*, in "European Conference on Artificial Life - ECAL 2011", Paris, France, MIT Press, 2011, p. 836-837, http://hal.inria.fr/inria-00617250/en.

[34] J. YAZBECK, A. SCHEUER, O. SIMONIN, F. CHARPILLET. *Improving near-to-near lateral control of platoons without communication*, in "IEEE-RSJ International Conference on Intelligent Robots and Systems (IROS)", San Francisco, United States, September 2011, p. 4103-4108, http://hal.inria.fr/inria-00603726/en.

National Conferences with Proceeding

[35] M. ARAYA-LÓPEZ, O. BUFFET, V. THOMAS, F. CHARPILLET. *Apprentissage actif de modèle de MDP*, in "Journées Francophones de Planification, Décision et Apprentissage pour la conduite de systèmes", Rouen, France, June 2011, http://hal.inria.fr/hal-00642913/en.

[36] M. ARAYA-LÓPEZ, O. BUFFET, V. THOMAS, F. CHARPILLET. *Une extension des POMDP avec des récompenses dépendant de l'état de croyance*, in "Conférence francophone sur l'Apprentissage automatique", Chambéry, France, May 2011, http://hal.inria.fr/hal-00642910/en.

[37] A. BAUTIN, O. SIMONIN, F. CHARPILLET. *Stratégie d'exploration multi-robot fondée sur les champs de potentiels artificiels*, in "Journées Francophones sur les Systèmes Multi-Agents (JFSMA)", Valenciennes, France, October 2011, http://hal.inria.fr/inria-00598608/en.

[38] O. BUFFET, L. CUCU-GROSJEAN. *Recherche systématique pour l'ordonnancement temps réel global multi-processeur*, in "12e congrès annuel de la Société française de Recherche Opérationnelle et d'Aide à la Décision", Saint Etienne, France, 2011, http://hal.inria.fr/inria-00584251/en.

[39] I. CHADÈS, J. CARWARDINE, T. MARTIN, S. NICOL, O. BUFFET. *Les POMDP: une solution pour modéliser des problèmes de gestion adaptative en biologie de la conservation*, in "Journées Francophones de Planification, Décision et Apprentissage pour la conduite de systèmes", Rouen, France, June 2011, http://hal.inria.fr/hal-00642915/en.

[40] M. GODICHAUD, E. CHANTHERY, O. BUFFET, M. CONTAT. *Formalisation et résolution de problèmes d'acquisition d'informations par des systèmes autonomes*, in "ROADEF (12e congrès annuel de la Société française de Recherche Opérationnelle et d'Aide à la Décision)", Saint-Etienne, France, 2011, http://hal.inria.fr/inria-00593463/en.

[41] L. SARZYNIEC, O. BUFFET, A. DUTECH. *Apprentissage par Renforcement Développemental en Robotique Autonome*, in "Conférence d'Apprentissage - CAP 2011", Chambéry, France, 2011, http://hal.inria.fr/inria-00633426/en.

[42] B. SCHERRER, M. GEIST. *Moindres carrés récursifs pour l'évaluation off-policy d'une politique avec traces d'éligibilité*, in "Journées Francophones de Planification, Décision et Apprentissage pour la conduite de systèmes", Rouen, France, June 2011, http://hal.inria.fr/hal-00644874/en.

Conferences without Proceedings

[43] A. BAUTIN, O. SIMONIN, F. CHARPILLET. *Towards a communication free coordination for multi-robot exploration*, in "6th National Conference on Control Architectures of Robots", Grenoble, France, INRIA Grenoble Rhône-Alpes, May 2011, 8 p., http://hal.inria.fr/inria-00599605/en.

Scientific Books (or Scientific Book chapters)

[44] C. BOURJOT, D. DESOR, V. CHEVRIER. *Stigmergy*, in "Self-organising Software: From Natural to Artificial Adaptation", G. DI MARZO SERUGENDO, M.-P. GLEIZES, A. KARAGEORGOS (editors), Natural Computing Series, Springer, 2011, p. 123-138, http://hal.inria.fr/hal-00641515/en.

[45] V. CHEVRIER, C. BOURJOT, V. THOMAS. *Region Detection in Images*, in "Self-organising Software", G. DI MARZO SERUGENDO, M.-P. GLEIZES, A. KARAGEORGOS (editors), Natural Computing Series: From Natural to Artificial Adaptation, Springer, 2011, p. 425-446, http://hal.inria.fr/hal-00641525/en.

[46] S. HEYMANS, J. HOFFMANN, A. MARCONI, J. PHLIPPS, I. WEBER. *Semantic Web Services Fundamentals*, in "Handbook of Service Description – USDL and its Methods", A. BARROS, D. OBERLE (editors), Springer-Verlag, 2011, http://hal.inria.fr/hal-00642315/en.

Research Reports

[47] N. BEAUFORT, F. CHARPILLET, O. ROCHEL, O. SIMONIN. *Interactive table to study interactions between swarm of robots and active environments*, March 2011, http://hal.inria.fr/inria-00575650/en.

[48] O. BOURÉ, N. FATÈS, V. CHEVRIER. *Observation of novel patterns in a stressed lattice-gas model of swarming*, October 2011, http://hal.inria.fr/inria-00634771/en.

[49] V. GABILLON, A. LAZARIC, M. GHAVAMZADEH, B. SCHERRER. *Classification-based Policy Iteration with a Critic*, May 2011, http://hal.inria.fr/hal-00590972/en.

[50] J. HOFFMANN. *Where Ignoring Delete Lists Works, Part II: Causal Graphs*, INRIA, January 2011, n^{o} RR-7505, http://hal.inria.fr/inria-00552255/en.

[51] B. SCHERRER, M. GEIST. *Recursive Least-Squares Off-policy Learning with Eligibility Traces*, 2011, http://hal.inria.fr/hal-00644516/en.

[52] B. SCHERRER. *Performance Bounds for Lambda Policy Iteration and Application to the Game of Tetris*, INRIA, 2011, n^{o} RR-6348, http://hal.inria.fr/inria-00185271/en.

[53] O. SIMONIN, F. CHARPILLET, E. THIERRY. *Collective Construction of Numerical Potential Fields for the Foraging Problem*, December 2011, http://hal.inria.fr/inria-00546325/en.

[54] J. YAZBECK, A. SCHEUER, O. SIMONIN, F. CHARPILLET. *Improving near-to-near lateral control of platoons without communication*, April 2011, http://hal.inria.fr/inria-00584560/en.

Other Publications

[55] A. BUSIC, N. FATÈS, J. MAIRESSE, I. MARCOVICI. *Density classification on infinite lattices and trees*, 2011, http://hal.inria.fr/hal-00643468/en.

[56] O. SIMONIN, T. HURAUX, F. CHARPILLET. *Comportements collectifs bio-inspirés : vers des robots-fourmis*, September 2011, PROgrammez.com, http://hal.inria.fr/hal-00646310/en.

References in notes

[57] A. DIB, C. ROSE, F. CHARPILLET. *Bayesian 3D Human Motion Capture Using Factored Particle Filtering*, in "22th International Conference on Tools with Artificial Intelligence - ICTAI 2010", Arras, France, IEEE, October 2010, http://hal.inria.fr/inria-00546925/en.

[58] A. DUBOIS. *Les technologies de l'information et de la communication pour le maintien à domicile des personnes âgées.*, December 2010, http://hal.inria.fr/inria-00551748/en/.

[59] M. GHALLAB, D. NAU, P. TRAVERSO. *Automated Planning: Theory and Practice*, Morgan Kaufmann, 2004.

[60] P. HASLUM, J. SLANEY, S. THIEBAUX. *Incremental Lower Bounds for Additive Cost Planning Problems*, in "Proceedings of the 3rd Workshop on Heuristics for Domain-Independent Planning (HDIP'11)", 2011.

[61] M. HELMERT, P. HASLUM, J. HOFFMANN. *Flexible Abstraction Heuristics for Optimal Sequential Planning*, in "Proceedings of the 17th International Conference on Automated Planning and Scheduling (ICAPS'07)", 2007, p. 176–183.

[62] J. HOFFMANN. *Where 'Ignoring Delete Lists' Works: Local Search Topology in Planning Benchmarks*, in "Journal of AI Research (JAIR)", 2005, vol. 24, p. 685-758.

[63] M. PUTERMAN. *Markov Decision Processes*, Wiley, New York, 1994.

[64] C. ROSE, B. BÉNÉ, F. CHARPILLET, J. CHANLIAU. *Automatic Evaluation of Vascular Access in Hemodialysis Patients*, in "XXXVII European Society for Artificial Organs Congress (ESAO 2010)", Skopje, Macédoine, L'Ex-République Yougoslave De, September 2010, http://hal.inria.fr/inria-00546732/en.

[65] S. RUSSELL, P. NORVIG. *Artificial Intelligence: A Modern Approach*, 2nd edition, Prentice-Hall, Englewood Cliffs, NJ, 2003.

[66] S. RUSSELL. *Rationality and Intelligence*, in "Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI)", 1995, Invited paper (Computers and Thought Award).

[67] A. SCHEUER, O. SIMONIN, F. CHARPILLET. *Safe Longitudinal Platoons of Vehicles without Communication*, in "Proceedings of the IEEE International Conference on Robotics and Automation - ICRA'09", Kobe (JP), May 2009, p. 70–75, http://www.loria.fr/~scheuer/Platoon.

[68] O. SIMONIN, F. CHARPILLET. *Indirect cooperation between mobile robots through an active environment*, in "5th National Conference on "Control Architecture of Robots"", Douai, France, May 2010, p. 71-80, http://hal.inria.fr/inria-00546388/en.

[69] R. SUTTON, A. BARTO. *Reinforcement Learning, An introduction*, BradFord Book. The MIT Press, 1998.