INRIA

Project-Team Maia

Autonomous intelligent machine

Nancy - Grand Est

Theme : Knowledge and Data Representation and Management

Activity Report

2010

Table of contents

1. Team

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]
Joerg 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, Secondment at CNRS, HdR]
Vincent Thomas [Associate Professor, U. Nancy 2]
Maan El-Badaoui-El-Najjar [External collaborator, Assistant professor, Lille]
Rodolphe Charrier [Prag U. Nancy 2]

Technical Staff

Nicolas Beaufort [Engineer]
Abdallah Dib [Engineer]
Amandine Dubois [Engineer]
Romain Mauffray [Engineer]
Marie Tonnelier [Engineer]

PhD Students

Mauricio Araya [scholarship, Chilian governement, U. Nancy 2]
Antoine Bautin [ANR scholarship, UHP]
Olivier Bouré [MENRT scholarship, UHP]
Gabriel Corona [CORDI Scholarship]
Arnaud Glad [ATER, UHP]
François Klein [ATER, INPL]
Tomas Navarrete [scholarship, Mexican government, UHP]
Cédric Rose [CIFRE, Diatelic SA]
Mohamed Tlig [InTrade, UHP]
Jano Yazbeck [InTrade, UHP]
Julien Siebert [scholarship, INRIA-Région Lorraine (1/9/10), UHP]
Cherif Smaili [ATER, Nancy 2]
Christophe Thiery [MENRT scholarship, UHP]

Post-Doctoral Fellow

Nikolaos Vlassopoulos [ARC Amybia INRIA]

Administrative Assistant

Céline Simon [TR, INRIA]

2. Overall Objectives

2.1. Introduction

MAIA[1] research belongs to the field of *artificial intelligence*: our goal is to model, design and simulate computer based entities (agents) that are able to sense their environment, interpret it, and act on it with

---

[1] MAIA stands for "MAchine Intelligente et Autonome", that is "Autonomous and Intelligent MAchine"

autonomy. We mainly work on three research themes: 1) sequential decision making, 2) complex systems and 3) embodied and spatially Distributed Intelligence.

2.2. Highlights

- The BARQ project proposed by Joerg Hoffman has been granted by ANR in the *Chaires d'Excellence program*
- Joerg Hoffmann was Conference Co-Chair of the 20th International Conference on Automated Planning and Scheduling (ICAPS-10).
- Olivier Buffet, François Charpillet, Alain Dutech, Bruno Scherrer, Daniel Szer and Simon Le Gloannec have participated in the writing of an introductory book on MDPs in Artificial Intelligence. After a french version in June 2008, an English version has been published in 2010 [3]. Maia team members were involved as authors in five chapters [9, 12, 8, 16, 19].
- Alain Dutech and Olivier Simonin have defended their HDR
- Both the robotic and smaart room platforms have been acquired
- Structuration of the EPI MAIA group around three interplaying groups has been acted

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.*

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** [75]. 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**) [83] 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 $E\left[\sum_{t=0}^{\infty} \gamma^t r_t\right]$. 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 [89]); (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. *Contemporary Related Work*

Sequential Decision Making is an active area of research in Artificial Intelligence, and has always been (e.g., planning research started in the 1960s). Our research fits well into the international research context, and we have made and are making a variety of significant contributions both in theoretical issues – e.g. on heuristic function topology, structure of value functions in active sensing, convergence of hybrids between value iteration and policy iteration – and in practical issues – e.g. in medicine together with the local start-up company "Diatélic", computer security with the American/Argentinian company "Core Security Technologies", and Business Process Management with SAP. A comprehensive description of related work would, of course, go beyond the scope of this section. The following is an incomplete list of teams and researchers working on related issues; many of these teams/researchers are, or have been in the past, collaborators of ourselves.

- **France:** BIA (INRA, Toulouse), DCSD (ONERA, Toulouse), Prof. Malik Ghallab (LAAS, Toulouse), IMS (Supélec, Metz), MAD (GREYC, Caen), Perception & Mouvement (ISIR, Paris), SAMMI (FEMTO-ST, Besançon), Sequel (INRIA Lille), TAO (INRIA Orsay).

- **Europe:** Prof. Hector Geffner (Barcelona, Spain), IST Lab (Lisbon, Portugal), Dr. Alexander Koller (Saarbrücken, Germany), Prof. Kim Larsen (Aalborg, Denmark), Prof. Lee McCluskey (Huddersfield, UK), Prof. Bernhard Nebel and Dr. Malte Helmert (Freiburg, Germany), Prof. Andreas Podelski (Freiburg, Germany), Robot Learning Lab (Max Planck Institute for Biological Cybernetics, Tübingen, Germany), SAP Research (Karlsruhe, Germany), Strathclyde Planning Group (Glasgow, UK).

- **Rest of the world:** Prof. Blai Bonet (Caracas, Venezuela), Prof. Ronen Brafman (Ben-Gurion University, Israel), Prof. Carmel Domshlak (Technion Haifa, Israel), Institute for Computational Sustainability (Cornell, Ithaca, USA), LIDS (MIT, Boston, USA), Prof. Paolo Traverso (Trento, Italy), RL[3] (Rutgers, USA), Dr. Sylvie Thiébaux (ANU-NICTA, Canberra, Australia), Reasoning and Learning lab (Mc Gill, Montreal, Canada), RLAI (Alberta, Canada), Prof. Stephen Smith (CMU, Pittsburgh, USA), Dr. Ingo Weber (Sydney, Australia).

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 exogenous 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. *Related works*

Among the principal approaches to study complex systems, we have chosen the line of reactive multi-agent systems which emphasizes the notions of interactions and self-organization.

This choice is reflected in the numerous collaborations that we have undertaken with researchers of this field as well as in the kinds of research groups we associate and work such as the Agent Technical Fora in Europe (more precisely the fora on self-organization and on simulation); or the research group "Colline" (under the aegis of GDR I3 and the AFIA) since its creation.

The Maia team is also involved in different editions of conferences and workshops on multi-agent and/or complex systems (as program committee members or advisory board members) such as JFSMA in France, EUMAS in Europe or MA4CS (Multi-Agent for modeling Complex systems) a satellite workshop of ECCS (the European Conference on Complex Systems), or the AAMAS (Autonomous Agents and Multi-Agent Systems) conference.

Since the study of complex systems involves interdisciplinary research, it is out of scope to provide an extensive view of the domain in this report. We can however mention some of the people or laboratories that work on similar domains of interest as the Maia team.

- In France: S. Hassas in LIESP (Lyon), CERV (Brest), IREMIA (la Réunion), Ibisc (Evry), Lirmm (Montpellier), Irit (Toulouse), A. Drogoul (IRD, Bondy), etc.

- Abroad: F. Zambonelli (Univ Modena, Italy) A. Deutsch (Dresden, Germany), D. Van Parunak (Vector research, USA), P. Valkenaers, D. Weyns (Univ. Leuven, Belgium), etc.

3.3. Embodied and Spatially Distributed Intelligence

3.3.1. *General context*

This section concerns the study and the design of autonomous and intelligent artificial systems, which are based on their interaction with the physical world. In this context, we are interested in defining models allowing the emergence of intelligent behaviors/functions without any *a priori* model. For this purpose, we investigate two approaches. On one side, we consider learning techniques and analytical models to acquire or provide new behaviors. On the other side, we investigate swarm intelligence and bio-inspired behaviors to define self-organized systems. In both cases, we aim at defining robust systems which are able to autonomously behave in unknown and dynamic environments.

As physical interaction plays an important role in this research line, experimentation with hybrid systems mixing both physical and computational parts takes an increasing part of the scientific development we envision. Mobile robots and smart rooms provide a natural way to study the interplay between software and embodied systems. Therefore we have acquired and developed this year mobile robot platforms and a smart room equipped with sensor and actuators networks.

More specifically, we address the following challenges/questions:

- Learning in cognitive robots, especially studying the influence of physical interaction in the learning process.

- Self-organization in multi-robot systems : how to define self-organized robots, i.e. how to ensure a global function to emerge from local interactions ? We investigate such an issue by studying bio-inspired mechanisms such as pheromone-based communication, signal diffusion and collective behaviors.

- How to compute inside or via the environment ? In the context of ambient intelligence, we are interested in defining discrete and physically distributed environments, which allow to define new calculus models dedicated to spatial computing and to the interaction with robots, communicating agents and humans.

- Active sensing : mastering the inherent complexity of massively spatially distributed devices (sensors, actuators).

3.3.2. *Current Challenges*

3.3.2.1. *Learning in cognitive robots*

This work is based on the concept of embodied cognition that says that intelligent behaviors, in an living or artificial agent, will emerge from its low level interactions with its environment. This requires, at least, adaptive and learning skills for the agent, and these concepts are investigated within the framework of reinforcement learning by the MaIA team. Thus, our goal in investigating "*cognitive robots*" is more about exploring learning and adaptive algorithms than building efficient robotic behaviors, at least for the very future.

Reinforcement Learning, while very appealing, must overcome several difficulties:

- **learning in continuous environments**. Most reinforcement learning algorithms rely on discrete state and action spaces. While the framework of reinforcement learning stays valid in continuous state and actions spaces, the formulation of learning algorithms is an open and active question in that context (see, for example, our new results in 5.1.13 and 5.1.14 for large state space).

- **costly (and harmful) interactions**. Every interaction between a robot and its environment consumes time and energy and, sometimes, can be fatal to the artificial agent. Ideally though, every interaction should be explored a large number of times by the agent for learning an optimal behavior. One has thus to define an alternate learning scheme making efficient use of the interactions so as to learn with as few interactions as possible.

- **Rich and only partially informative environment**. Even with the crude sensors available today, the raw quantity of information than can be perceived by a robot can be huge (take, for example, a simple camera: it offers a raw perception space of more than 100 000 dimensions). This raw perception can and must be processed, avoiding two pitfalls: very specialized perceptions (and a task too easy to learn without adaptation to other tasks) and over-simplified perceptions (making the task to learn very difficult, possibly leading to a partial knowledge of the environment).

Our approach to robotic reinforcement learning follows two ideas. On the one hand, we take inspiration from developmental robotics [78] to explore how to combine incremental perception and action spaces with shaping to enhance the efficiency of reinforcement learning. On the other hand, we combine artificial neural networks and the reinforcement learning framework of direct policy gradient ascent to develop new learning rules local to neurons that lead to promising neuro-controllers in continuous action spaces.

3.3.2.2. *Swarm robotics*

The main question we are interested in is how to build robust and adaptive robots by inspiration of the mechanisms and the behaviors present in natural swarm systems. In the MaIA team we focus on two bio-inspired models, which are flocking [84] (for multi-robot navigation) and digital stigmergy [71] – or environment marking – (for multi-robot spatial organization and cooperation). The study of these two swarm intelligence models is organized in three sub-challenges:

- **Self-organization**. Since the 90's, implementing biological mechanisms in real robots is a way to define intelligent and collective behaviors. In this context, our main objective is to study and to define local behaviors allowing a global self-organization of the system. Self-organization can be defined as the ability to adapt its global behavior to environmental changes, without requiring any external or internal leader decisions. For this purpose, we investigate two approaches, relying first on flocking techniques (direct coordination), and second on robots' interaction with an active environment (indirect cooperation via the environment). We aim at defining robots able to carry out tasks in *unknown and/or dynamic environments*, where robustness to perturbations is probably the main criterion to evaluate the performance.

- **Robotics and spatial computing**. Considering the interaction between a set of autonomous robots and an active environment is a current issue of ambient intelligence and swarm robotics [79]. This concept relies on the biological paradigm of stigmergy. It consists in marking or modifying the environment to guide the system organization and to provide an indirect means of communication. In the context of autonomous robots, we consider the possibility of using such an active environment,

by addressing the following questions: how to design/program such an active environment? what are the links between swarm robotics and spatial or amorphous computing? how to connect the robots' environment and such spatial models?

- **New experimental devices**. Addressing the study of bio-inspired behaviors with robots requires having adapted experimental devices. Indeed, bio-inspired robotics is yet an unconventional approach in robotics, so it requires to define specific models, robots, environments and tools for its study.

3.3.3. Related Works

Our project of swarm robotics study and design has common objectives with the DISAL[2] 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.[3], the FlockBots project at GMU[4], the team of G. Théraulaz at CNRS-Toulouse and the teams of J.L. Deneubourg and M. Dorigo at ULB.

4. Software

4.1. Software

4.1.1. ISeeML

ISeeML is an Integrated Smooth, Efficient and Easy-to-use Motion Library. It offers a simple way to compute continuous-curvature paths for car-like robots, using line segments, circular arcs and pieces of clothoids [85], [86]: C++ simple constructors are provided for these paths, as well as for more classical paths (*i.e.* Dubins'paths [74]).

ISeeML has been registered this year to the APP (French Agency for Programs Protection), and will be distributed under a free license for academics soon (last transactions between INRIA's and Nancy University's administrations are still underway).

- *Availability*: ISeeML's web pages (http://ISeeML.loria.fr) contain a documentation (generated using Doxygen, http://www.doxygen.org) detailing the library's classes, either in HTML or in PDF [64], a Java demonstration of a graphical example, and a form to apply for a license.
- *Contributors*: Alexis Scheuer
- *Contacts*: Alexis.Scheuer@loria.fr

4.1.2. AA4MM

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).

- *Availability*: Download at (http://www.loria.fr/~siebertj)
- *Contributors*: Julien Siebert
- *Contacts*: julien.siebert@loria.fr

---

[2]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

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

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

4.1.3. MASDYNE

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.

- *Contributors*: Julien Siebert, Tom Leclerc, Christophe Torin, Marcel Lamenu, Guillaume Favre et Amir Toly.
- *Contacts*: julien.siebert@loria.fr

4.1.4. FiatLux

FiatLux is a cellular automata simulator that allows the user to experiment with various models and to perturb them. These perturbations can be of two types. On the one hand, perturbations of dynamics change the type of updating, for example from a deterministic parallel updating to an asynchronous random updating. On the other hand, the user may perturb the topology of the grid by removing links between cells randomly.

FiatLux may be run in an interactive mode with a Graphical User Interface or in a batch mode for longer experiments. The interactive mode is suited for small size universes whereas the batch may be used for experiments involving several thousands of cells. The software uses two external libraries for the random generator and the real-time observations of variables ; it is also fitted with output procedures that writes in Gnuplot, Tex, HTML formats.

In 2010, the software has evolved with the development of Lattice Gas models (LGCA) and Interacting Particle Systems (IPS).

- *Availability*: Download it at http://webloria.loria.fr/~fates/fiatlux.html
- *Contributors*: Nazim Fatès
- *Contact*: Nazim.Fates@loria.fr

4.1.5. FPG

FPG is a probabilistic planner addressing its problem as a reinforcement learning one. Its principle is to optimize a controller's parameters (such as a neural network whose input is a state and whose output is a decision) using a domain simulator. It relies on the libPG library.

Although the first version was meant to deal with concurrent temporal probabilistic planning, current development effort is focusing on FPG-ipc, the version used for the international planning competition.

- *Availability*: http://fpg.loria.fr/
- *Contributors*: Olivier Buffet, Douglas Aberdeen, Joerg Hoffmann
- *Contact*: olivier.buffet@loria.fr

4.1.6. SUSIE Simulator

In the framework of the 'Susie' PEA-DGA project (7.2.5) we develop in collaboration with ENST Brest a simulator to study multi-agent patrolling algorithms funded on digital stigmergy (e.g. the EVAP model 5.2.11). This simulator provides an implementation of EVAP with an OpenGL interface to use graphical capabilities and it is connected to the LightSim AUV (flying robots) simulator developed by ENST Brest.

- *Availability*: http://recherche.telecom-bretagne.eu/susie/
- *Contributors*: Olivier Simonin, François Charpillet, Romain Mauffray
- *Contact*: Olivier.Simonin@loria.fr

5. New Results

5.1. Decision Making

5.1.1. *Decentralized Partially Observable Markov Decision Processes (DEC-POMDPs)*
Participants: François Charpillet, Gabriel Corona, Alain Dutech.

Decentralized Partially Observable Markov Decision Processes define a framework for solving sequential decision problems in which several agents act synchronously in a distributed way. The problem to solve can be expressed as follow : find for each agent a local policy such that the joint policy maximizes the long term accumulated rewards of the system. State transitions are determined stochastically by the current state of the process and the set of actions, one for each agent. The rewards are associated to the global state of the process and each local policy is executable autonomously by each agent i.e it is expressed in term of local observations and actions of the considered agent. Several approaches have been proposed to solve optimally Dec-POMDPs. Maia group has been very active in this area proposing bottom-up dynamic approach [90], top down search [91] and mathematical programming [70] [5] . The computation of an optimal joint policy is **NEXP**-complete in a finite horizon. Thus the exact computation is not affordable except for very small problems (especially, with a small number of observations) and for small horizons. To overcome this issue, some very efficient approximated algorithms have been proposed so far: MBDP (Memory Bounded Dynamic Programming), IMBDP (Improved MBDP), MBDP-OC (MBDP with Observation Compression), PBIP (Point Based Incremental Pruning). They use only a bounded number of partial plans at each planning step. By bounding the memory, they are able to solve problems with dramatically longer horizons than previous algorithms. By avoiding the exhaustive backup, some of them can solve problems with a higher number of observations. This year, we propose such a new point based approach, PSMBDP (Policy Search Memory Bounded Dynamic Programming) [23] [12] which uses a heuristic estimation of the prior probability distribution of reachable beliefs and formulates the choice of a bounded number of partial plans as a combinatorial optimization problem. Using this heuristics information, PSMBDP find solutions of better quality than state-of-the-art

This year, we also finalized the dissemination of our important work on using mathematical programming for solving Dec-POMDPs in an exact way. Building on the work of the past few years, we communicated our findings (new algorithms that can be faster than existing ones, knowledge about the structure of the solutions to Dec-POMDPs) in an extensive and detailed way in a important and widely read journal (JAIR) [5].

5.1.2. *Mixed Observability MDPs*
Participants: Mauricio Araya, Vincent Thomas, Olivier Buffet, François Charpillet.

Partially Observable Markov Decision Processes (POMDPs) represent a class of sequential decision making problems which are made especially difficult because of the limited knowledge of the current state of the system. The accessible observations only make it possible to have beliefs about the real state of the system to be controlled. This typically requires working with $|S|$-dimensional functions (S being the finite set of states). One of our objectives is to find structure (/specific features) in POMDP problems that could be exploited to ease the search for a good (if not optimal) policy.

We observed in particular that, in a number of POMDPs, the state s is not completely hidden. It can typically be decomposed into a visible part s_v and a hidden part s_h ($S = S_v \times S_h$). This makes it possible to work with $|S_v|\,|S_h|$-dimensional functions instead of one $|S|$-dimensional function. Even if it is difficult to theoretically prove that the former leads to more efficient algorithms than the latter, we have shown experimentally that there is a dramatic speed up (and a similar benefit regarding memory consumption).

The same idea has been proposed recently by Ong *et al.* [80] under the name of Mixed Observability MDPs (MOMDPs). Our work—published in [14]—is complementary, especially because it gives a new insight on why this approach is efficient.

5.1.3. A POMDP Extension with Belief-dependent Rewards
Participants: Mauricio Araya, Olivier Buffet, Vincent Thomas, François Charpillet.

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 published in [13], the proofs appearing in a companion tech-report [57].

5.1.4. Utility Functions vs Preference Relations in Recommender Systems
Participant: Olivier Buffet.

Anne Boyer, Armelle Brun and Ahmad Hamad (Kiwi team, Loria) are external collaborators.

Classical approaches for recommender systems need ratings (*i.e.*, utilities) to represent user preferences and deduce unknown preferences of other users. In this work, we focused on an original way to represent preferences under the form of preference relations. This approach can be viewed as a qualitative representation of preferences at the opposite of the usual quantitative representation. The only information known is whether a user prefers one item over an other item. "How much" this item is preferred is not known. This approach has the advantage to not require users to rate items in a small rating scale of integer values, which may be a difficult task and the resulting ratings may highly depend on the user, his mood, the preceding items he has rated, etc. We have proposed to adapt classical measures that exploit utilities so as to exploit preference relations, such as the similarity measure between users. First experiments have been conducted on a well-known user data set that represents user utilities. We have transformed this data set under the form of preference relations. First results have shown that this approach leads to comparable performance with the classical approach. This work has been published in the proceedings of national and international conferences [42], [17], [16].

5.1.5. Multiprocessor Realtime Scheduling under Uncertainty
Participant: Olivier Buffet.

L. Cucu-Grosjean (TRIO team, LORIA) is an external collaborator.

Many embedded systems (e.g. in cars or planes) have to treat repetitive tasks (with different periods) using several processors. However, existing work focuses on distributing jobs on the processors under the assumption that their execution time is fixed, which requires considering the worst-case execution time.

We have first considered the problem of scheduling jobs over processors in this worst-case deterministic scenario. We have proposed improved systematic search algorithms that dramatically speed up the search of a valid schedule (or the proof that a problem is unfeasible) [58].

More recently we have started considering a specific problem linked to the schedulability of uniprocessor systems when the execution time is variable. The open question we are looking at is the impact of dropping some jobs on schedulability, as this may render a schedule feasible [19].

5.1.6. Reward Shaping in the Factored Policy-Gradient (FPG) Planner
Participants: Olivier Buffet, Joerg Hoffmann.

FPG addresses probabilistic planning. A key issue in such planning is to exploit the problem's structure to make large instances solvable. Most approaches are based on algorithms which explore — at least partially — the state space, so that their complexity is usually linked to the number of states. Our approach is very different in that it is exploring a space of parametrized controllers. By choosing factored controllers (one sub-controller per action) with state variables as inputs, we strongly reduce the complexity problem.

Most work on FPG was done in previous years. In 2010 (partly in 2009), we developed a new method for reward shaping, where the problem's reward function is modified so as to encourage progress towards goal states. The reward shaping is non-intrusive in that it does not change the optimal solution to the problem; it is essential to success in problems with large search spaces, where reaching the goal by a pure random walk is very unlikely. The shaping is based on progress estimates which are derived from the structure of the problem, specifically from a straightforward relaxation transforming the problem into a deterministic planning problem. In a pre-process, our technique automatically detects landmarks — variable values that every successful path must at some point traverse — as well as pairwise constraints on the order in which that will happen. The progress estimator is based on reasoning about how many landmarks yet need to be achieved. A paper is published at an ICAPS 2010 workshop [20].

5.1.7. *Semantic Validation of Business Processes*
Participant: Joerg Hoffmann.

Ingo Weber (University of New South Wales, Australia) and Jan Mendling (Freie Universitaet Berlin, Germany) are external collaborators.

Business process control the flow of activities within and between enterprises. The verification of control-flow soundness is well understood as an important step before deploying business process models. However, the control flow does not capture what the process activities actually do when they are executed. Semantic annotations offer the opportunity to take this into account. Inspired by semantic Web service approaches such as OWL-S and WSMO, we consider process models in which the individual activities are annotated with logical preconditions and effects, specified relative to an ontology that axiomatizes the underlying business domain. Verification then addresses the overall process behavior, arising from the interaction between control-flow and behavior of individual activities. To this end, we combine notions from the workflow community with notions from the AI actions and change literature. We introduce a formal execution semantics for annotated business processes. We point out four verification tasks that arise, concerning precondition/effect conflicts, reachability, and executability. We examine the borderline between classes of processes that can, or cannot, be verified in polynomial time. In particular, we identify a non-trivial class of processes that can be verified in polynomial time by a fixpoint algorithm which we design for that purpose. We show that this class of processes is maximal in the sense that, when generalizing it in any of the most relevant directions, the validation tasks become computationally hard. Most of this work was performed in 2009, before Joerg Hoffmann joined MAIA. The work was completed in 2010, and a journal paper has been published in 2010 in the Journal on Distributed and Parallel Databases [11].

5.1.8. *Composition of Business Processes at SAP*
Participant: Joerg Hoffmann.

Ingo Weber (University of New South Wales, Australia) and Frank Michael Kraft (SAP, Germany) are external collaborators.

The behavior of certain software artefacts can be naturally expressed, at an appropriate level of abstractions, in terms of their effect on state variable values. At SAP, a set of 2700 system transactions, which underlie the execution of business processes, have been modeled in this way (the uncertainty lies in the fact that many transactions may have different outcomes depending on details that are abstracted away at the level of the model). Our work leverages on this model by providing a formalization in a planning language, and an adaptation of Joerg Hoffmann's planning tool FF. The resulting technology fully automatically composes useful business process fragments, requiring as input only the "goal", i.e., a specification of which variables should assume which values. This corresponds well to the background and language of the targeted group of

end-users (managers); in a prototype developed at SAP, the goal specification is given using simple drop-down menus. Most of the work was performed in 2009. In 2010, we completed it, in particular running large-scale experiments. A conference paper is published at AAAI 2010 [33]. A long version has been completed and is currently under review for the Computational Intelligence journal.

5.1.9. *Using Planners for Natural Language Sentence Generation*
Participant: Joerg Hoffmann.

Alexander Koller (Saarland University, Saarbruecken, Germany) is an external collaborator.

A relevant problem in many applications is to communicate, in natural language, with the human user. For this purpose, one sub-problem that needs to be solved is to compose a natural language sentence that expresses a given (logical) meaning. Alexander Koller is an expert in this area, and has since a few years been experimenting with the use of planning system to accomplish this task. His hitherto results were disappointing: off-the-shelf planners were not able to generate sentences in problem instances of meaningful size and content. In our work together, we identified that the causes of the computational problems were mostly easy to fix. We produced a new version of Joerg Hoffmann's planning tool FF. This new version shows a dramatic speed-up over previous planners, in this domain, and for the first time makes such an application a realistic possibility. A conference paper on this work is published at ICAPS 2010 [32].

5.1.10. *Improving Local Search for Resource-Constrained Planning*
Participant: Joerg Hoffmann.

Hootan Nakhost and Martin Mueller (University of Alberta, Edmonton, Canada) are external collaborators.

Most state of the art planning systems are based on heuristic search, where the heuristic estimates of remaining cost are derived from a relaxed problem, i.e., a simplified version of the actual problem. The simplification, in most of the leading systems, consists in ignoring so-called "negative effects". This basically means to act as if every action changes the world only in adding new facts, never removing old ones. For example, if a car drives and uses fuel, then its new fuel level is lower however it still retains the old level as well. This relaxation has the obvious problem of not being able to take into account the use of consumable resources – like fuel. Indeed, if one runs current planners on resource-constrained problems and scales the tightness of the constraint, decreasing the available amount of the resource, then one finds that planners excel when resources are plenty and fail completely when they are very tight. In our work, we address this problem via designing new search methods, based not on exhaustive search but on frequent probing. This enables us to look at a larger fraction of the search space, and to recover more easily from bad decisions. In the resource-constrained situation, this results in a speed-up of several orders of magnitude. A first paper on this work is published at SOCS 2010 [35]. Our ongoing work completes this with additional experiments, and we plan to submit a paper to IJCAI 2011.

5.1.11. *Analyzing Planning Domains to Predict Heuristic Function Quality*
Participant: Joerg Hoffmann.

The heuristic search approach to planning (cf. the above) rises and falls with the quality of the heuristic estimates. The dominant method is to approximate a heuristic function called h^+ – this is used in almost every state of the art planning system. In earlier work, Joerg Hoffmann showed that h^+ has some amazing qualities, in many traditional planning benchmarks, in particular pertaining to the complete absence of local minima. Hoffmann's 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 Joerg 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 a tool we call TorchLight. 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 is currently under review for ICAPS 2011, and a long paper version is already prepared for submission to the Journal of AI Research (JAIR).

5.1.12. *Relaxing Bisimulation for State Aggregation in the Computation of Lower Bounds*
Participant: Joerg Hoffmann.

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

Like the previous two lines 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 by Joerg Hoffmann, Malte Helmert, and Patrik Haslum (NICTA, Canberra, Australia) in 2007 [76]. 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. A conference paper is in preparation for submission to IJCAI 2011.

5.1.13. *Least-Squares λ Policy Iteration: Bias-Variance Trade-off in Control Problems*
Participants: Bruno Scherrer, Christophe Thiéry.

In the context of large space MDPs with linear value function approximation, we introduce a new approximate version of λ-Policy Iteration [72], a method that generalizes Value Iteration and Policy Iteration with a parameter $\lambda \in (0, 1)$. Our approach generalizes LSPI [77] which makes efficient use of training samples compared to classical temporal differences methods. The motivation of our work is to exploit the λ parameter within the least-squares context, and without having to generate new samples at each iteration or to know a model of the MDP. We have derived a performance bound that shows the soundness of the algorithm. We have also shown empirically on a simple chain problem and on the Tetris game that this λ parameter acts as a bias-variance trade-off that may improve the convergence and the performance of the policy obtained. This work was accepted and presented in the national workshop JFPDA [46] and at the ICML 2010 conference [40]. The performance bound we mention is available as a technical report [63].

5.1.14. *Should one compute the Temporal Difference fix point or minimize the Bellman Residual ? The unified oblique projection view*
Participant: Bruno Scherrer.

In another paper published at ICML 2010 [37], we investigate projection methods, for evaluating a linear approximation of the value function of a policy in a Markov Decision Process context. We consider two popular approaches, the one-step Temporal Difference fix-point computation (TD(0)) and the Bellman Residual (BR) minimization. We describe examples, where each method outperforms the other. We highlight a simple relation between the objective function they minimize, and show that while BR enjoys a performance guarantee, TD(0) does not in general. We then propose a unified view in terms of oblique projections of the Bellman equation, which substantially simplifies and extends the characterization of [88] and the recent analysis of [94]. Eventually, we describe some simulations that suggest that if the TD(0) solution is usually slightly better than the BR solution, its inherent numerical instability makes it very bad in some cases, and thus worse on average.

5.1.15. *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), and Olivier Pietquin (Supélec Metz (IMS)) 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 [27].

5.2. Understanding and mastering complex systems

5.2.1. *Planning as a Method for Leading Convergence Proofs of Cellular Automata*
Participants: Joerg Hoffmann, Nazim Fatès.

Hector Palacios (Universidad Simon Bolivar, Caracas, Venezuela) is an external collaborator.

An interesting question for several types of cellular automata is that of which behaviors lead to a stable system state where no more changes can be made. In particular, a question of interest is whether or not random updates will eventually converge to such a stable state. For sequential updating, this question is equivalent to the question whether one can make the system converge by explicitly *choosing* the cells to update. Posed in this form, convergence corresponds to a planning problem – the "decision" to be taken is which cell to update. For leading a convergence proof, one needs to prove that good decisions can be taken no matter what the starting state is. Such proving is highly cumbersome, and often impossible, to do by hand due to the multitude of different start configurations – even when considering only a small system with 16 cells and 2 possible values for each cell, one needs to consider 2^{16} possibilities. In our work, we explore the idea to leave the enumeration of these possibilities up to the computer – more precisely, up to a planner, which has the advantage that it works for a very large set of updating rules (namely for all those that can be translated into a planning language). The overall proof method, then, is to identify how a stable state can be decomposed into fixed-size local parts, and let the planner enumerate the cases for these local parts. For the human user, leading the proof thus comes down to designing the decomposition, and specifying the local parts. The latter can be done in a graphical way. So far, we have led a first convergence proof in this fashion, and published a conference paper about this at ECAI 2010 [31]. Our ongoing work is to integrate this technology, in a convenient form, into FiatLux.

5.2.2. *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. 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.

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.

A first set of experiments [43] demonstrates that individual dimensions (such as time and space) can have a significant impact on the simulation results thus justifying the use of multi-agent models. A second set of experiments shows us that the control architecture can indeed achieve the defined control objective but still under some restrictions.

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.

5.2.3. *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) that provides solutions both for multi-modeling and multi-simulation of complex systems by reusing existing and heterogeneous M&S tools [39], [38]. 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 [34]. 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.

5.2.4. *A bio-inspired model for aggregating quasi-blind agents*

Participants: Nazim Fatès, Nikolaos Vlassopoulos.

Hugues Berry (INRIA Lyon) and Bernard Girau (CORTEX team, LORIA) are external collaborators.

The decentralised gathering problem consists in grouping in a compact cluster agents that are initially randomly scattered. We studied a bio-inspired algorithm that allows to group agents that have limited abilities [8] (AMYBIA project). The agents and their environment are described with a stochastic model inspired by the aggregation of the *Dictyostelium* cellular slime mold. The environment is an active lattice, whose cells transmit information according to a reaction-diffusion mechanism. The agents are virtual amoebae; they trigger excitations randomly and move by following reaction-diffusion waves.

Despite its simplicity at the local scale, this model exhibits interesting properties of self-organisation and is efficient for gathering agents [69]. Moreover, observations show that the system is robust to various perturbations, such as the presence of obstacles on the lattice or noise in the movements of the agents.

In this context, we studied the phase transitions that appear in the stochastic version of the reaction-diffusion cellular automata (Greenberg-Hastings model) [28]. Interestingly, we found out an asymptotic property for the value of the critical setting of these models. As, the simulations required to compute the asymptotic value were demanding, we developed an FPGA-based platform that allowed us to speed up the simulations. So far, the results confirm our hypothesis that the value of asymptotic threshold is not trivial, contrarily to what one would expect with a simple naive analysis [41]. It is currently an open problem to explain the discrepancy between the theoretical expected value and the empirical measure.

5.2.5. *Density classification problem*

Participant: Nazim Fatès.

The density classification problem is a consensus problem which illustrates the difficulty of synthetising a behaviour in the field of complex systems. It consists in using a binary cellular automaton (CA) to decide whether an initial configuration contains more 0s or 1s. This problem is known for having no exact solution in the case of deterministic one-dimensional CA. Stochastic cellular automata have been studied as an alternative for solving the problem.

We investigated how a particular class of stochastic CA rules can be described as a "blend" of deterministic CA rules. We analysed two previously studied rules and presented a new rule that solves the problem with an arbitrary precision [68]. Using analytical calculations and numerical simulations, we estimated how the quality of classification and the average time of classification varies as a function of the blending parameter. From a practical point of view, simulations show that for a great range of values, the new rule exhibits a quality of classification never attained so far.

This research is currently continued within a collaboration with Jean Mairesse (CNRS, LIAFA, Paris 7) and co-workers.

5.2.6. *Study of robustness of Reactive Multi-Agent Systems*

Participants: Olivier Bouré, Vincent Chevrier, Nazim Fatès.

As part as our research on emergent collective behaviours, we have taken interest in studies related to robustness analysis – behavioural resistance to noise perturbations – in collective systems, such as Cellular Automata (CA) and Multi-Agent Systems (MAS), and have acquired an overview of the existing ways to tackle this type of problem. As a first step, we decided to focus on a particular perturbation: asynchronism, which aims at randomizing the way cells are updated in a CA.

We study the asynchronism in cellular automata from a multi-agent point of view: we apply the Influence-Reaction principle for MAS to CA and uncovered a new type of asynchronism (that we called perceptive asynchronism) that focuses on desynchronizing the communication between cells, instead of the update scheme. We derived a formalism for describing asynchronism in both ways, and found out that they have qualitatively similar effects on the Game of Life, although their behaviour drifts for low synchrony rate.

We are currently interested in studying the impact of perceptive asynchronism on other systems, as well as other types of perturbations, in order to bring us closer to a unified definition of robustness for both CA and MAS.

5.2.7. *Modeling Reactive Multi-Agent Systems as discrete dynamical systems*

Participants: Vincent Chevrier, Nazim Fatès.

It is to date an open question to know how the updating methods affect the evolution of a multi-agent system. This question has been tackled for various complex systems such as cellular automata, Boolean networks, neural networks but little is known for multi-agent systems, especially for the models with a complex behaviour which emerges from simple local rules. Finding adequate descriptions of multi-agent systems is a central issue for modelling collective dynamics and thus to explicitly design choices in updating schemes.

We have proposed a mathematical description of multi-agent systems as discrete dynamical systems [22]. The ground of our proposition is the influence-reaction method of Ferber and Müller. The key idea is that agents should never act directly on other components of the system (agents or environment) but release influences which are then combined to update the state of the system. We propose a method which decomposes the definition of a multi-agent system into six parts: (1) the basic sets, (2) the perception of the agents, (3) the influences, (4) the updating of the agents' internal state, (5) the updating of the environment, (6) the updating of the position and observable states of the agents.

We have applied our method on the multi-turmite model [29], also known as the multiple Langton's ants model. All the agents are updated simultaneously and the variation of the updating scheme consists only in choosing different strategies for solving the conflicts produced when two or more agents want to go on the same location. We show that for the *same* formulation of the agents' behaviour, and the same initial conditions, the use of different updating schemes may lead to qualitatively different evolutions of the system. As a positive spin-off of this study, we exhibit new phenomena of the multi-turmite model such as deadlocks or gliders.

The robustness of the multi-turmite model was tested and it was shown that a wide range of phenomena could be observed. In particular, we observed that some patterns are very sensitive to changes in the updating scheme while there exist some patterns that resist well to such perturbations [66].

5.2.8. Building multi-agent systems by the use interactions and RL techniques
Participant: Vincent Thomas.

Mahuna Akplogan participated during his internship.

This approach concerns both decision making and complex systems research themes. It focuses on the DEC-POMDP model, proposed by S. Zilberstein in 2000. DEC-POMDP was one of the first models to formally describe distributed decision problems, but works have proved that building the optimal policies of agents in this context is in practice intractable (NEXP complexity).

Our work tries to introduce explicit interactions among agents in this model in order to structure collective problems and reduce the complexity of building agents behaviours. The core idea of this approach is to mix multi-agent concepts and reinforcement algorithms to find new ways to organize a society of agents to solve a goal.

We are currently pursuing this approach through the concept of social actions as a way to represent actions and interactions in a similar manner. The objective is to stay close to the DEC-POMDP formalism in order to compare the results obtained with those obtained with other decision making approaches but also to propose new coordination mechanisms on which the agents can reason. By doing that, we try in the long run to propose a new framework to model reactive multi-agent systems but also to propose decentralized algorithms to control the agents' behaviors so that they can organize themselves to solve collective problems. A new journal article concerning this approach has been published end of 2009 [93]

5.2.9. Simulating complex fish behaviors
Participants: Christine Bourjot, Olivier Buffet, Vincent Thomas.

Jean-Noel Gardeur (URAFPA, INRA Nancy) and Pascal Fontaine (URAFPA, INRA Nancy) are external collaborators.

One approach to have a better understanding of complex systems is to analyze and observe existing ones. We have naturally focused on living systems which present the properties of complex systems. They are composed of several elements (a population, several organs) organized so as to collectively solve a problem. For this reason, we have begun a collaboration with biologists form URAFPA two years ago concerning the modeling of fishes feeding and reproduction collective behaviors.

The feeding behaviors of fishes is not yet understood. When several fishes are put together in a tank and are regularly fed, a differentiation emerges leading to two groups of fishes: one constituted by big fishes and another one by small fishes which have difficulties to feed themselves whereas the food was abundant. The URAFPA team has proposed the mobifish model to understand this phenomenon. As Hamelin, a model we have developed [92], it is based on a coupling between dominance relationships and another behavior (here the feeding behavior), suggesting that this interaction structure appears in several collective and complex biological phenomena and is worth investigating. By studying how this mechanism is working, we want to propose in the long run new coordination mechanisms that can help to understand and build complex systems.

The other collective behaviors that we have focused on is the reproduction behavior of fishes. The domestication of a fish species is very complex since a lot of experiments about this species must be conducted to gather enough information. The URAFPA team has proposed an original approach consisting in modeling and understanding a specific species and trying to adapt this model to lesser known species. We are currently trying to model the physiology of individuals through interacting components requiring a holistic view of the system.

5.2.10. Collective construction of artificial potential fields (APF)
Participants: Olivier Simonin, François Charpillet.

In the context of path-planning, we rewrote the classical Artificial Potential Field (APF) computation proposed by Barraquand & Latombe in a distributed and asynchronous version based on a collective construction by reactive agents. We proved that this model builds an optimal APF while dealing with the collective foraging problem (research and transport of resources by a set of autonomous agents/robots). In 2009/10 we extended simulations and measures by introducing dynamic environments (moving obstacles). Then we have shown that our approach is more efficient in static environments than the classical ant algorithm, and needs to be extended with a behavioral heuristic to compete with it in dynamic environments. Details are given in the article "Collective Construction of Numerical Potential Fields for the Foraging Problem" which has been accepted for publication in ACM TAAS Transaction on Autonomous and Adaptive Systems (to appear in 2011).

5.2.11. 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). Results of 2010 are :

- Study of the EVAP behavior when considering different execution models (agents scheduling, non determinism, asynchronism). This work has been published in AAMAS'2010 [30].
- Coding of EVAP with the iTiles model and experiments with real robots (see [30]).
- Writing of Arnaud Glad's PhD. thesis, that he should defend in the beginning of 2011. The writing of a journal article is also in progress.

5.2.12. Logistic Ants
Participants: Rodolphe Charrier, Christine Bourjot, François Charpillet.

The "Logistic" Multi-Agent System (LMAS) is used to explore swarm phenomena with a reactive multi-agent system. It takes inspiration from the Coupled Map Lattices (CML), which constitute a basic model in the complex systems sciences. CML may be considered as cellular automata with continuous cellular states and nonlinear transition functions, in particular the logistic maps. In fact, the logistic map is used in the Logistic MAS as an internal decision generator for the agents, providing them with regular to chaotic decisions.

Then two basic mechanisms may be used with this internal logistic map: one is to couple an agent with others (its neighbors) so as to obtain a synchronized group, the other consists in changing the control parameter of the logistic map, caused by perceptions from the environment, to get a regular or chaotic individual behavior.

Actually the LMAS model provides a basic tool to explore bio-inspired complex systems like ant colonies or bird flocks, with the same theoretical framework, which is not the case with swarm models at the moment. In particular, we have carried on the design of a new metaheuristic for combinatorial optimization problems, by proposing the logistic ant model which derives from the Logistic MAS [21].

5.3. Embodied and Spatially Distributed Intelligence

5.3.1. Learning in cognitive robots

5.3.1.1. Direct policy-gradient approach in noisy neuro-controller
Participant: Alain Dutech.

Emmanuel Daucé (Institut des sciences du mouvement, UMR 6233, Marseille) is an external collaborator.

In this work, we have developed a new neuronal learning rule making use of a global scalar reinforcement signal to adapt the weights of a one-layer neuro-controller in order to solve a highly non-linear task. The artificial neurons we use are linear neurons with a gaussian noise and they can be seen as stochastic decentralized elements of a global policy, policy that is improved by a gradient ascent in policy space. One key characteristic of our work is that the local learning rule, very similar to the classical Widrow-Hoff delta rule for the perceptron, has a low computational cost. Thus, the input signal can be topographically recoded into a very high dimensional space that is then fed to the neuro-controller allowing it to find a solution even to highly non-linear problems.

Our neuro-controller has been validated on a control task involving a simulated 4-segments robotic arm and on a gaze-control task [24], [25]. Other preliminary results show that this approach compares well to state of the art methods like Natural Actor Critic [82].

5.3.1.2. Developmental reinforcement learning for autonomous robotic
Participants: Olivier Buffet, Alain Dutech.

Luc Sarzyniec (M2R Student of UHP Nancy 1) is an external collaborator.

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.

Some preliminary experiments have shown that, on very simple tasks, the incremental approach is both faster and reaches better performances than an approach tackling directly the problem with the full range of perceptions and actions [62].

5.3.2. Swarm robotics

5.3.2.1. 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 2010 are

- Study of a new multi-robot exploration algorithm funded on the WaveFront technique introduced by Barraquand and Latombre[5]. The idea of our approach consists in computing a potential field from the frontiers between unexplored and explored areas and to go down the gradient (a first publication is under submission). We now explore decentralized multi-fields computation to optimize the frontiers allocation among the robots.

- Since the end of 2010, we dispose of a new mobile robot platform composed of six mini-Pekee developed by the Wany robotics partner. During the year we performed experiments with Khepera III robots equiped with laser rangefinder available in the Maia team.

5.3.2.2. New experimental device: the Interactive Table
Participants: Olivier Simonin, François Charpillet, Nicolas Beaufort, Romain Mauffray.

---

[5]J. Barraquand and B. Langlois and J.C. Latombe, Numerical Potential Field Techniques for Robot Path Planning, IEEE Trans. on Systems, Man and Cybernetics 2(22) 224-241, 1992

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.

Design, positioning and first experiment :

- The table, equipped with a PC, can compute and display any active environment (we tested for instance digital pheromone diffusion and evaporation). The table is also fitted with infrared cameras to track robots, objects and human motion on the surface. We use Khepera III robots which have been extended by two new embedded electronic boards. The first one provides seven colorimetric sensors and three infrared transmitters under the robot frame, the second board is placed at the top of the robot to support an Arduino electronic platform and other peripheral devices. These last components have been developed mainly by Nicolas Beaufort, who is IJD[7] engineer in the INRIA ROMEA project (Sec. 7.2.6) funding these researches.

- This experimental device provides an original alternative to existing ones proposed by T. Kazama[8] and G. Theraulaz[9] that use image projection over the robots. These last propositions are subject to light variation and shadows, and do not allow accurate interactions between robots and the environment.

- *Study and experiment of the EVAP model (Sec. 5.2.11)*. As a first study, we examined the expression of the EVAP algorithm on a continuous environment such as the table, and defined the corresponding gradient descent on robots. Robots drop digital pheromones under their frame, which is diffused and evaporated by the table. As predicted by the theoretical model, robots carry out a multi-agent patrolling of the environment. We now investigate the long term property of self-organization of the robots (spatial organization in stable cycles).

5.3.2.3. Testing the decentralised gathering scheme on the Interactive Table
Participants: Nazim Fatès, Nikolaos Vlassopoulos.

The purpose of our study is to apply the virtual amoebae aggregation scheme (see Sec. 5.2) to the Khepera robots, using the Romea (see above) interactive table in order to simulate the active environment where the agents reside. We endeavor to quantify and study the robustness of the model on this multi-robot system in order to analyse how the various types of perturbations on the agent movement and agent interaction influence the dynamics of the system. This introduces several interesting challenges, for instance the interaction of a discrete-time based environment, that emulates the behavior of a reaction-diffusion cellular automaton, with the robot agents that operate in continuous time. We are also dealing with problems such as collision avoidance and "dead-lock" avoidance, where a robot can be "trapped" by the presence of other robots in its surroundings.

The initial study plan in this direction can be decomposed into several stages. The first stage is to study the aggregation time, which is probably the most important metric, and compare it with the respective experimental results from the particle and cellular automata models. The second stage is to find an optimal set of parameters so as to study the effect of the various types of perturbation on the aggregation time of the robots. Later stages include developing a behavioral analytical model, as well as investigating how the aggregation scheme could be tested in noisy environments.

5.3.3. Platooning and multi-robot navigation

5.3.3.1. Local control based platooning
Participants: Alexis Scheuer, Olivier Simonin, François Charpillet, Jano Yazbeck, Sophie Jacquin.

---

[6]Service d'Expérimentation et Développement
[7]Ingénieur Jeune Diplomé
[8]T. Kazama and K. Sugawara and T. Watanabe, Traffic-like movement on a trail of interacting robots with virtual pheromone, Proceedings of the 3rd International Symposium on Autonomous Minirobots for Research and Edutainment (AMiRE 2005), 383-388, 2006.
[9]S. Garnier and F. Tache and M. Combe and A. Grimal and G. Theraulaz, Alice in Pheromone Land: An Experimental Setup for the Study of Ant-like Robots, Proc. 2007 IEEE Swarm Intelligence Symposium, 37-44, 2007.

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.

- *Safe decentralised control reducing inter-vehicle distances.* We studied a more realistic model of the platooning task, introducing a delay between perceptions and actions and noise/errors in perceptions and actions. Within this modelling, Daviet and Parent controllers [73] show their limitations. We thus proposed a high-level controller, which transforms any controller into a safe controller, i.e. avoiding collision, and we proved this property. This work has been published in the 2009 IEEE International Conference on Robotics and Automation (ICRA'09) [87]. An extended version of this article, with detailed proof, is available as INRIA research report and will be submitted to an international journal.

- *Coupling lateral and longitudinal controls.* The work presented in the previous paragraph focuses on longitudinal control: 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. A lateral control only based on individual sensors has been studied and the current longitudinal control has been adapted to curvilinear distances. Several strategies have been studied:

  – The work of Jano Yazbeck (for 5 months), intern student at LORIA, is entitled "Decentralized local approach for lateral control of platoons" (03/10–07/10), and has been supervised by A. Scheuer and O. Simonin. J. Yazbeck examined a lateral control in order to drive each vehicle exactly in the trace of the previous one, by using only embedded sensors such as a laser rangefinder [65]. An article about this work has been submitted to 2011 IEEE International Conference on Robotics and Automation.

  – The work of Sophie Jacquin (for 2 months), intern student at LORIA, is entitled « Planification de chemins à courbure C^n » ("Planning paths whose curvature is of differentiability class C^n") and has been supervised by A. Scheuer. Its main idea is to improve the paths used by ISeeML (*cf* § 4.1.1) in order to increase their smoothness [61]. These paths can be used to define a new lateral control of better quality (higher smoothness).

5.3.3.2. Reinforcement learning for controlling unicycle robots in leader-follower formation
Participants: François Charpillet, Nicole El Zoghby.

The objective of this work was to evaluate a new approach for designing controllers for robots evolving in leader follower formation, i.e., an approach in which the control law defining the behavior of each vehicle is automatically learnt using a simple reinforcement learning algorithm (Q-learning) [60]. A simulator based on a kinematic model defining the dynamics of each robot in the platoon permits through an exploration and exploitation process, to obtain an optimal policy for each vehicle in the platoon. Although Q-learning is based on a discrete representation of states and actions, we demonstrated that this technique makes it possible to build efficient controllers when compared to traditional approaches. A comparison with other longitudinal platooning algorithms has been performed and it demonstrates the relevance of our approach.

5.3.3.3. Bayesian methods for localizing vehicles in leader-follower formation
Participants: Maan Badaoui, Cherif Smaili, François Charpillet.

This year concluded a piece of work around multi-sensor data fusion and map matching for localizing a set of vehicles evolving in leader-follower formation. The originality of our approach has been to propose a unified framework based on Bayesian networks, generalizing two of the most successful families of model commonly used for this purpose: linear Kalman filters and Hidden Markov Models. Cherif Smaili defended his PhD thesis this year which synthesizes all work done during the last three years [3].

5.3.3.4. Decentralized platoons crossing : a reactive coordination approach
Participants: Olivier Simonin, François Charpillet.

In 2009/2010, during the internship of Sébastien Albouze (Mines de Nancy, initiation a la recherche internship) we studied a model for non-stop crossing of two orthogonal decentralized platoons. We have shown that a local rule is able to ensure the global synchronization of the two platoons to perform their crossing. At the crossing area, we introduced a local control on ghost vehicles. Ghost vehicles are the projection of the opposite platoon on the current one (see details in [56]). A publication detailing this model is in preparation.

5.3.4. Ambient intelligence and Actimetry

5.3.4.1. Robotics and spatial computing : the iTiles - intelligent tiles - model
Participants: Olivier Simonin, François Charpillet, Romain Mauffray.

In order to extend robots' abilities/functions and to equip homes with intelligent environments, we explore the definition of distributed and physical calculus models. Since 2009, we study a model consisting in paving the floor with "communicating" and autonomous *tiles*. Each tile is defined to ensure communication with its neighbors, and to allow a possible supported robot to read and write information. As a consequence tiles can be exploited to extend robots' perceptions and communications, and to physically implement bio-inspired algorithms. A first Tile model has been defined and evaluated using a simulator, then we developed a tiles' emulator and performed some experiments with real mobiles robots (Kheperas III), which validated the interest and the efficiency of the approach. The iTiles model was presented in a paper published in ICAART'09 [81].

Work progress of 2010 is :

- We started the development of a physical and electronic prototype of the iTiles. This development is made in cooperation with our colleagues from the SED of INRIA Grenoble. We use the technology and the software developed in the context of the SensTools INRIA project (D-NET INRIA team).

- In Nancy, we develop the iTiles model inside the CPER Informatique Située project (for which O. Simonin is in charge). The project explores ambient intelligence models and technologies for Smart Home design. In 2011, we will install one hundred iTiles on the floor of the new LORIA platform, called Intelligent Apartment.

- We published the first algorithms and experimental results, using mobile robots and the iTiles emulator, in CAR'2010 conference [44].

5.3.4.2. 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 [26].

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 [59]. 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.

5.3.4.3. 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 [36] 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: Joerg 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. They contacted us in summer 2010 since they needed expertise for helping to use and configure the FF planning tool, which Joerg Hoffmann had developed during his PhD work at the University of Freiburg, Germany. Joerg Hoffmann has started in late 2010 to provide them with consulting, in this context. Core Security Technologies is paying for this consulting to Joerg Hoffmann as an Auto-Entrepreneur (a demande de cumul has been duly submitted and confirmed). A long-term research collaboration with Core Security Technologies, to develop this approach further, is planned. We are currently negotiating this with their Research Department.

6.1.2. *DOPEC*
Participants: Olivier Buffet, François Charpillet.

The DOPEC project is 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 work on autonomous sequential decision making problems. We are more particularly interested, on the one hand, in multi-agent problems and, on the other hand, in taking uncertainties into account.

7. Other Grants and Activities

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[10] 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.

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 5.1.2).

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. We are currently working on the study of the influence of delays at stops on the global performance.

7.2. National Initiatives

7.2.1. ANR project BARQ
Participants: Joerg 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). Joerg Hoffmann will be the project leader, Olivier Buffet and Bruno Scherrer will collaborate. Other collaborators from LORIA will be 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

---

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

to aggregate, and how to effectively obtain, in practical scenarios, anytime methods providing solutions with increasingly tighter performance guarantees. To answer the former, we will relax known state aggregation criteria (such as bisimulation) from probabilistic planning and model checking, where traditionally state aggregation techniques are aimed at incurring no information loss at all (and thus often exhaust computational resources). To answer the latter, we will investigate the tight interleaving of lower bounding and upper bounding methods, specifically using abstraction refinement to obtain incrementally better lower bounds.

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 is organized in the beginning of January 2011 in Lyon.

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. Antoine Bautin has recently started a PhD on the topic of multi-robot deployment and mapping. (Section 5.3.2.1).

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. sect. 5.3.4.1) 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 5.2.11 to the patrol with UAVs, while providing a generic digital pheromone based patrolling simulator (see 4.1.6). 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 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 5.3.2.2. 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.

7.3. European Initiatives

7.3.1. 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. 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. Currently, the group members are 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.

7.3.1.2. 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.

7.3.2. 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 recently. 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".

8. Dissemination

8.1. Animation of the scientific community

8.1.1. *Journal and Conference reviewing*

- Olivier Buffet was a reviewer for the journals: AIJ (Artificial Intelligence Journal), AMAI (Annals of Mathematics and Artificial Intelligence), IJAR (International Journal on Approximate Reasoning), JAIR (Journal of Artificial Intelligence Research), and RIA (*Revue d'Intelligence Artificielle*); for the conferences: AAAI-10, AAMAS-10, ECAI-10, ICAPS-10, JFPDA-10, RFIA-10, SASO-10; and for the ICAPS-10 workshop on planning and scheduling under uncertainty.

- Alain Dutech was a reviewer for the journals ACM-TSLP (Transaction on Speech and Language Processing) and RIA (*Revue d'Intelligence Artificielle*); for the conferences JFPDA-10, RFIA-10 and ICML-10.

- Vincent Chevrier was a reviewer for the special issue 2010 on multi-agent systems of the "revue d'intelligence artificielle" (RIA).

- François Charpillet was a reviewer for the journals: JAAMAS (Journal of Artificial Intelligence Research), and RIA (*Revue d'Intelligence Artificielle*); for the conferences: ECAI-10, ICAART-10, IROS-10, COGNITIVE 2010, IEEE-Healthcom2010, MSDM-10, JFSMA-10, JFPDA-10.

- Nazim Fatès was a reviewer for the journals: ACS (Advances in Complex Systems) and JTB (Journal of Theoretical Biology) and for the conferences : AUTOMATA'10, SOLSTICE'10, ANTS'10, ACRI'10, CAAA/HPCS'10. He contributed to the MathSciNet reviews edited by the American Mathematical Society.

- Alexis Scheuer was a reviewer for IEEE TRO (Transactions on Robotics) and for Elsevier's international journal SIMPAT (Simulation Modelling Practice and Theory).

- Olivier Simonin was a reviewer for the journals: ACM TAAS (Transactions on Autonomous and Adaptive Systems), JAAMAS (Journal of Autonomous Agents and Multi-Agent System), Journal of Field Robotics (impact factor 2.7), The Computer Journal (Oxford Journals) and Mechatronics (International Journal); for the international conferences: SASO'2010, EuroCon-2011, ICINCO (2010, 2011); for the AT2AI-7 workshop (From Agent Theory to Agent Implementation); and for the national conferences JFSMA'10 and HUMOUS'10 (Humans Operating Unmanned Systems).

- Vincent Thomas was a reviewer for the IADIS Intelligent Systems and Agents 2010 (ISA 2010) Conference.

- Bruno Scherrer was a reviewer for NIPS'2010, ICML'2010 and ECML'2010.

8.1.2. Conference organization, Program committees, Editorial boards

- Olivier Buffet was an organizer and program committee member of the "ICAPS-10 workshop on planning and scheduling under uncertainty", and is a member of the editorial board of both the "revue d'intelligence artificielle" (RIA) and the "Journal of Artificial Intelligence Research" (JAIR).

- François Charpillet is a member of the editorial board of the "revue d'intelligence artificielle" (RIA) and program committee member of ECAI-10, ICAART-10, COGNITIVE 2010, IEEE-Healthcom2010, MSDM-10, JFSMA-10, JFPDA-10

- Nazim Fatès was organiser, conference chair and program committee member of Automata'10 and Solstice'10, INRIA Nancy – LORIA, 14-18 June 2010. He was member of the program committee of ANTS'10, ACRI'10, CAAA/HPCS'10.

- Joerg Hoffmann was Conference Co-Chair of the 20th International Conference on Automated Planning and Scheduling (ICAPS-10).

- Joerg Hoffmann is an Associate Editor of the "Journal of Artificial Intelligence Research" (JAIR).

- Joerg Hoffmann is Area Chair for Planning of "AI Communications".

- Joerg Hoffmann was a member of the following conference program committee:

  – Senior Program Committee of the 24th National Conference of the American Association for Artificial Intelligence (AAAI-10).

  – Senior Program Committee of the AI and the Web Track (AIW-10) at the 24th National Conference of the American Association for Artificial Intelligence (AAAI'10).

- Vincent Chevrier was a member of:

  – the editorial board of Interstices[11], a site to disseminate research work about computer science for French-speaking person,

  – the advisory board of JFSMA (Journées Francophones sur les Systèmes Multi-Agents)

  – the program committee of AAMAS 10 (the International Conference on Autonomous Agents and Multiagent Systems) ; Eumas10 (European workshop on multi-agent systems) and of JFSMA10.

- Vincent Chevrier is the moderator of the mailing list of the French spoken community on multi-agent systems.

- François Charpillet is a member of the editorial board of the "revue d'intelligence artificielle" (RIA) and program committee member of ECAI-10, ICAART-10, COGNITIVE 2010, IEEE-Healthcom2010, MSDM-10, JFSMA-10, JFPDA-10

- Olivier Simonin is a program committee member of SASO'2010 and ICINCO'2010,'2011 international conferences, of the AT2AI-7 workshop and of JFSMA'10 and HUMOUS'10 french conferences.

- Vincent Thomas was a member of the program committee of the IADIS Intelligent Systems and Agents 2010 (ISA 2010) Conference.

---

[11]http://interstices.info

8.1.3. PhD and HDR committees

- Olivier Buffet was a member of the following PhD Committee:
    – (as a committee member) Jacques Bibai, *Segmentation et évolution pour la planification : le système Divide-And-Evolve*, Univ. Paris-Sud 11, October 8, 2010.

- Joerg Hoffmann is a member of the PhD committee of Emil Keyder (Universitat Pompeu Fabra Barcelona, Spain), December 2010.

- François Charpillet was a member of the following PhD Committees:
    – (as a reviewer) Yoann Maurel, *Ceylan : un canevas pour la création de gestionnaires autonomiques extensibles et dynamiques*, Université Joseph Fourier, Grenoble, December 1, 2010.
    – (as a committee member) Julien Guitton, *Architecture hybride pour la planification d'actions et de déplacements*, Université Toulouse, March 31, 2010.
    – (as a reviewer) Christophe Thiery, *Itération sur les Politiques Optimiste et Apprentissage du Jeu de Tetris* , Université Henri Poincaré, Nancy 1, November 25, 2010.

- François Charpillet was a member of the following HDR Committees:
    – (as a reviewer) Alain Dutech, *Apprentissage par Renforcement : Au delà des Processus Décisionnels de Markov. (Vers la cognition incarnée).* HDR de l'Université Nancy 2, December 2, 2010.
    – (as a reviewer) Jean-Charles Lamirel, *Vers un approche systémique et multivues pour l'analyse de données et la recherche d'information : un nouveau paradigme.* HDR de l'Université Nancy 2, December 6, 2010.
    – (as a reviewer) Olivier Simonin, *Contribution à la résolution collective de problème - Modèles d'auto-organisation par interactions directes et indirectes dans les SMA réactifs et robotiques -* HDR de l'Université Henri Poincaré, Nancy 1, December 10, 2010.

- Vincent Chevrier was a member of the following PhD Committee:
    – (as a committee member) Thomas Girod, *Un modèle d'apprentissage multimodal pour un substrat distribué d'inspiration corticale*, Univ. H Poincaré, November 10, 2010.
    – (as a referee) Olivier Lefèvre, *Contrôle endogène des Systèmes Multi-Agents pour la résolution de problèmes complexes*, Univ. Lyon 1 Claude Bernard, October 5, 2010.

- Alain Dutech was a member of the PhD committee (as a referee) of Olga Kozlova, *Hierarchical & Factored Reinforcement Learning*, University Pierre et Marie Curie, 7th of June 2010.

- Olivier Simonin was a member of the following PhD Committee:
    – (as a committee member) Jean-Luc Paillat, *Conception et contrôle de robots à géométrie variable : applications au franchissement d'obstacles autonome*, Université d'Angers/LISA, November 15, 2010.
    – (as a committee member) Van Tuan Le, *Coopération dans les systèmes multi-robots : contribution au maintien de la connectivité et à l'allocation dynamique de rôles*, Université de Lille Nord de France, Ecole des Mines de Douai, October 6, 2010.

8.1.4. Specialist Committees (commissions de spécialistes)

- Vincent Chevrier was a member of the "Specialist Committees" in Université Joseph Fourier, Grenoble.

- Vincent Thomas was a member of the "Specialist Committees" in Université Toulouse 1 Capitole.

- François Charpillet was a member of the "Specialist Committees" in Université Joseph Fourier, Grenoble.

- Bruno Scherrer was a member of the Committee for the Young researcher competition of INRIA Lille.

8.1.5. Other responsibilities

- Olivier Buffet is a member of the "CDT", the INRIA Nancy Grand-Est committee for technologic developments.

- François Charpillet is member of the scientific committee of the GDR Robotics for the next period (2011-2014)

- Alain Dutech is a member of the "Conseil de Laboratoire" of the LORIA, Nancy. He is also part of the "Work group on Communication" of the LORIA, Nancy. He also participates to the local AGOS section of the INRIA Nancy Grand-Est center.

- Olivier Simonin is a member of INRIA Nancy Grand Est comipers-chercheur (evaluation committee for INRIA Ph.D. and post-doc applications).

- Olivier Simonin is in charge of the *Informatique Située* Project of the CPER MISN (Contrat de Projets Etat-Région Modélisation, Informations et Systèmes Numériques) at INRIA Nancy and Lorraine region.

- Olivier Simonin is a member of the *Fédération Charles Hermite* Council (including CRAN, IECN, LMAM and LORIA laboratories from Nancy universities, Paul Verlaine Metz university and from CNRS http://www.fr-hermite.univ-lorraine.fr/)

8.1.6. Scientific Diffusion

- Alain Dutech was invited as an "expert on AI" for the viewing of the documentary "Plug an Pray" (directed by Jens Schanze) at the *Pariscience festival*. He also gave a lecture on Embodied Cognitive Robotic in a high-school in Remiremont, Vosges, France.

- Vincent Thomas participated in "Les pieds sur Terre" the fourth congress of young researchers organized in April 2010 in Nancy.

- Olivier Buffet, François Charpillet, Alain Dutech, Bruno Scherrer, Daniel Szer and Simon Le Gloannec have participated in the writing of an introductory book on MDPs in Artificial Intelligence. After a french version in June 2008, an English version has been published in 2010 [55]. Maia team members were involved as authors in five chapters [49], [50], [48], [52], [53].

- Olivier Simonin was invited to animate the *Café des Sciences* debate "Demain, l'homme robot ?", in Metz, Café Jeanne d'Arc (07/04/10).

- Romain Mauffray, Nicolas Beaufort and Olivier Simonin were invited to perform demonstrations of intelligent tiles with mobile robots at Cité des Sciences "Tech'Galerie" event (Paris La Villette, June 2010) and at Forum des Sciences Cognitives Nancy 2 (November, 25, 2010).

- Nazim Fatès was a speaker in a public debate, the *Café des techniques*, organised by the CNAM in Nancy, on the theme "Comment l'homme et la machine peuvent-ils cohabiter ?". He also participated as an animator, to the training of high-school teachers to prepare them for to teach a 'Sciences et techniques du numérique' optional module at 'Classes de seconde' level.

8.2. Teaching

- Among other courses, Vincent Thomas is responsible for three courses in masters
  - one course entitled "Modélisation de comportement" 40h in Master Science Cognitive - Nancy 2 University - in collaboration with biologist presenting the agent paradigm (20h) and biological models (20h).

– a new course built this year entitled "Optimisation et modèles dynamiques stochastiques" 15h in Master RAR in computer Science - Nancy 1 UHP University - , presenting optimization techniques and Markov decision processes.

– a new course built this year entitled "Jeu video et monde virtuel" 20h in Master Science Cognitive - Nancy 2 University - presenting Game Design and Artificial intelligence techniques for serious game development.

- Alexis Scheuer and Olivier Simonin have recently proposed a new course entitled « Intelligence Artificielle pour la Robotique Mobile » (Artificial Intelligence for Mobile Robotics) in second year of Master RAR (« Reconnaissance, Apprentissage, Raisonnement », *i.e.*, Recognition, Learning, Thinking) at Nancy 1 UHP University. Khepera III robots are used in some practical courses, allowing students to program real robots and to discover experimental conditions.

  After a first year with only research students, the course is more detailed this year (30 h instead of 20) due to a professional training.

- Nazim Fatès gave a 3-hour lecture in the Ecole des Mines in the 'Introduction au Genie Industriel' module on the theme of Multi-agent systems. He also participated in the 'Systèmes Communicants' and 'Systèmes dynamiques' modules with lectures on cellular automata.

9. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] A. DUTECH. *Apprentissage par Renforcement : Au delà des Processus Décisionnels de Markov (Vers la cognition incarnée)*, Université Nancy II, December 2010, Habilitation à Diriger des Recherches, http://hal.inria.fr/tel-00549108/en.

[2] O. SIMONIN. *Contribution à la résolution collective de problème (Modèles d'auto-organisation par interactions directes et indirectes dans les SMA réactifs et robotiques)*, Université Henri Poincaré - Nancy I, December 2010, Habilitation à Diriger des Recherches, http://hal.inria.fr/tel-00546300/en.

[3] C. SMAILI. *Fusion de données multi-capteurs à l'aide d'un réseau bayésien pour l'estimation d'´etat d'un véhicule*, Université Nancy II, May 2010, http://hal.inria.fr/tel-00551833/en.

[4] C. THIERY. *Itération sur les Politiques Optimiste et Apprentissage du Jeu de Tetris*, Université Henri Poincaré - Nancy I, November 2010, http://hal.inria.fr/tel-00550081/en.

Articles in International Peer-Reviewed Journal

[5] R. ARAS, A. DUTECH. *An investigation into Mathematical Programming for Finite Horizon Decentralized POMDPs*, in "The Journal of Artificial Intelligence Research", 2010, vol. 37, p. 329-396 [*DOI :* 10.1613/JAIR.2915], http://www.jair.org/media/2915/live-2915-4898-jair.pdf, http://hal.inria.fr/inria-00439627/en.

[6] C. CAPPELLE, M. EL BADAOUI EL NAJJAR, D. POMORSKI, F. CHARPILLET. *Intelligent geo-localisation in urban areas using global positioning systems, 3dimensional geographic information systems, and vision*, in "Journal of Intelligent Transportation Systems Technology Planning and Operations", 2010, vol. 14, n° 1, p. 3 - 12 [*DOI :* 10.1080/15472450903385999], http://www.informaworld.com/smpp/ftinterface~content=a919247906~fulltext=713240930~frm=content, http://hal.inria.fr/inria-00441814/en.

[7] N. FATÈS. *Critical phenomena in cellular automata: perturbing the update, the transitions, the topology*, in "Acta Physica Polonica B", 2010, http://hal.inria.fr/inria-00412278/en.

[8] N. FATÈS. *Solving the Decentralised Gathering Problem with a Reaction-Diffusion-Chemotaxis scheme - Social amoebae as a source of inspiration*, in "Swarm Intelligence", 2010, vol. 4, n° 2, p. 91-115 [*DOI :* 10.1007/s11721-010-0038-4], http://www.springerlink.com/content/u438082n21hw5731/fulltext.pdf, http://hal.inria.fr/inria-00440007/en.

[9] O. SIMONIN, F. CHARPILLET, E. THIERRY. *Collective Construction of Numerical Potential Fields for the Foraging Problem*, in "ACM Transactions on Autonomous and Adaptive Systems", December 2010, http://hal.inria.fr/inria-00546325/en.

[10] A. SPICHER, N. FATÈS, O. SIMONIN. *Translating Discrete Multi-Agents Models into Cellular Automata, Application to Diffusion-Limited Aggregation*, in "CCIS 67 Communications in Computer and Information Sciences series", January 2010, vol. 67, p. 270-282 [*DOI :* 10.1007/978-3-642-11819-7], http://hal.inria.fr/inria-00546424/en.

[11] I. WEBER, J. HOFFMANN, J. MENDLING. *Beyond Soundness: On the Verification of Semantic Business Process Models*, in "Distributed and Parallel Databases", 2010, vol. 27, n° 3, p. 271-343 [*DOI :* 10.1007/s10619-010-7060-9], http://hal.inria.fr/inria-00491102/en.

Articles in National Peer-Reviewed Journal

[12] G. CORONA, F. CHARPILLET. *Distribution sur les croyances pour la planification de Dec-POMDP*, in "Revue d'Intelligence Artificielle", 2010, vol. 24, n° 4, p. 525-544, http://ria.e-revues.com/article.jsp?articleId=15224, http://hal.inria.fr/inria-00439530/en.

International Peer-Reviewed Conference/Proceedings

[13] M. ARAYA-LÓPEZ, O. BUFFET, V. THOMAS, F. CHARPILLET. *A POMDP Extension with Belief-dependent Rewards*, in "Neural Information Processing Systems - NIPS 2010", Vancouver, Canada, MIT Press, December 2010, http://hal.inria.fr/inria-00535560/en.

[14] M. ARAYA-LÓPEZ, V. THOMAS, O. BUFFET, F. CHARPILLET. *A Closer Look at MOMDPs*, in "22nd International Conference on Tools with Artificial Intelligence - ICTAI 2010", Arras, France, IEEE, October 2010, http://hal.inria.fr/inria-00535559/en.

[15] A. BAUTIN, L. MARTINEZ-GOMEZ, T. FRAICHARD. *Inevitable Collision States: a Probabilistic Perspective*, in "IEEE International Conference on Robotics and Automation - ICRA 2010", Anchorage, AK, États-Unis, IEEE, 2010, p. 4022 - 4027 [*DOI :* 10.1109/ROBOT.2010.5509233], http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5509233, http://hal.inria.fr/inria-00546366/en.

[16] A. BRUN, A. HAMAD, O. BUFFET, A. BOYER. *From "I like" to "I prefer" in Collaborative Filtering*, in "International Conference on Tools with Artificial Intelligence - ICTAI 2010", Arras, France, IEEE, October 2010, 3, http://hal.inria.fr/inria-00535566/en.

[17] A. BRUN, A. HAMAD, O. BUFFET, A. BOYER. *Towards Preference Relations in Recommender Systems*, in "Workshop on Preference Learning - PL 2010", Barcelona, Espagne, Eyke Hüllermeier and Johannes Fürnkranz, September 2010, p. I-XV, Ce workshop fait partie de la conférence European Conference on

Machine Learning and Principles and Practice of Knowledge Discovery in Databases - ECML PKDD 2010, http://www.ke.tu-darmstadt.de/events/PL-10/papers/1-Brun.pdf, http://hal.inria.fr/inria-00523496/en.

[18] O. BUFFET, L. CUCU, L. IDOUMGHAR, R. SCHOTT. *Tabu search type algorithms for the multiprocessor scheduling problem*, in "10th IASTED International Conference on Artificial Intelligence and Applications - AIA 2010", Innsbruck, Autriche, 2010, http://hal.inria.fr/hal-00435241/en.

[19] O. BUFFET, L. CUCU-GROSJEAN. *Impact of job dropping on the schedulability of uniprocessor probabilistic real-time systems with variable execution times*, in "1st International Real-Time Scheduling Open Problems Seminar : RTSOPS 2010", Bruxelles, Belgique, July 2010, held in conjunction with the 22th Euromicro Intl Conference on Real-Time Systems, July 7-9, 2010, Brussels, Belgium, http://hal.inria.fr/inria-00535561/en.

[20] O. BUFFET, J. HOFFMANN. *All that Glitters is not Gold: Using Landmarks for Reward Shaping in FPG*, in "ICAPS-10 Workshop on Planning and Scheduling Under Uncertainty", Toronto, Canada, May 2010, http://hal.inria.fr/inria-00534375/en.

[21] R. CHARRIER, C. BOURJOT, F. CHARPILLET. *A Deterministic Metaheuristic Approach using "Logistic Ants" for Combinatorial Optimization.*, in "Ants 2010 Seventh International Conference on Swarm Intelligence", Brussels, Belgique, Lecture Notes in Computer Science, Springer, October 2010, vol. 6234, p. 344-351 [*DOI :* 10.1007/978-3-642-15461-4_30], http://www.springerlink.com/content/476620098120k852/fulltext.pdf, http://hal.inria.fr/inria-00549250/en.

[22] V. CHEVRIER, N. FATÈS. *An Example of a Multi-agent System described as a Discrete Dynamical System*, in "8th European Workshop on Multi-agent Systems - EuMAS 2010", Paris, France, Paris Descartes University, December 2010, http://hal.inria.fr/inria-00549216/en.

[23] G. CORONA, F. CHARPILLET. *Distribution over Beliefs for Memory Bounded Dec-POMDP Planning*, in "The 26th Conference on Uncertainty in Artificial Intelligence (UAI 2010)", Catalina Island, California, États-Unis, AUAI Press, July 2010, ISBN 978-0-9749039-6-5, http://event.cwi.nl/uai2010/papers/UAI2010_0231.pdf, http://hal.inria.fr/inria-00546912/en.

[24] E. DAUCÉ, A. DUTECH. *Cartes topographiques neuronales pour l'apprentissage par renforcement sur des problèmes de contrôle non-linéaire*, in "10e Colloque Africain sur la Recherche en Informatique et en Mathématiques Appliquées", Yamoussoukro, Côte D'Ivoire, 2010, 9 P., http://hal.inria.fr/inria-00494164/en.

[25] E. DAUCÉ, A. DUTECH. *Online Learning with Noise: A Kernel-Based Policy-Gradient Approach*, in "Conférence Française de Neurosciences Computationnelles - NeuroComp 2010", Lyon, France, 2010, http://hal.inria.fr/inria-00517006/en.

[26] 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.

[27] A. DUTECH, E. COUTUREAU, A. MARCHAND. *Reinforcement Learning Approaches to Instrumental Contingency Degradation in Rats*, in "Conférence Française de Neurosciences Computationnelles - NeuroComp 2010", Lyon, France, 2010, http://hal.inria.fr/inria-00517011/en.

[28] N. Fatès, H. Berry. *Robustness of the critical behaviour in a discrete stochastic reaction-diffusion medium*, in "IWNC'09", Himeji, Japon, Proceedings in Information and Communications Technology, Springer, 2010, http://hal.inria.fr/inria-00396473/en.

[29] N. Fatès, V. Chevrier. *How important are updating schemes in multi-agent systems? An illustration on a multi-turmite model.*, in "Ninth International Conference on Autonomous Agents and Multiagent Systems - AAMAS'10", Toronto, Canada, ACM, 2010, vol. 1, p. 533-540, http://www.ifaamas.org/Proceedings/aamas2010/pdf/01%20Full%20Papers/11_04_FP_0210.pdf, http://hal.inria.fr/inria-00546845/en.

[30] A. Glad, O. Simonin, O. Buffet, F. Charpillet. *Influence of Different Execution Models on Patrolling Ants Behavior : from Agents to Robots*, in "AAMAS'10 9th International Conference on Autonomous Agents and Multiagent Systems", Toronto, Canada, ACM, May 2010, p. 1173-1180, http://portal.acm.org/ft_gateway.cfm?id=1838190&type=pdf&CFID=2008812&CFTOKEN=25338702, http://hal.inria.fr/inria-00546344/en.

[31] J. Hoffmann, N. Fatès, H. Palacios. *Brothers in Arms? On AI Planning and Cellular Automata*, in "19th European Conference on Artificial Intelligence (ECAI'10)", Lisbon, Portugal, 2010, http://hal.inria.fr/inria-00491114/en.

[32] J. Hoffmann, A. Koller. *Waking Up a Sleeping Rabbit: On Natural-Language Sentence Generation with FF*, in "20th International Conference on Automated Planning and Scheduling (ICAPS'10)", Toronto, Canada, 2010, http://hal.inria.fr/inria-00491109/en.

[33] J. Hoffmann, I. Weber, F. Kraft. *SAP Speaks PDDL*, in "24th National Conference of the American Association for Artificial Intelligence (AAAI'10)", Atlanta, États-Unis, 2010, http://hal.inria.fr/inria-00491123/en.

[34] T. Leclerc, J. Siebert, V. Chevrier, L. Ciarletta, O. Festor. *Multi-modeling and co-simulation-based mobile ubiquitous protocols and services development and assessment*, in "7th International ICST Conference on Mobile and Ubiquitous Systems - Mobiquitous 2010", Sydney, Australie, December 2010, http://hal.inria.fr/inria-00534575/en.

[35] H. Nakhost, J. Hoffmann, M. Müller. *Improving Local Search for Resource-Constrained Planning*, in "3rd Annual Symposium on Combinatorial Search (SOCS'10)", Atlanta, États-Unis, 2010, http://hal.inria.fr/inria-00491129/en.

[36] 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.

[37] B. Scherrer. *Should one compute the Temporal Difference fix point or minimize the Bellman Residual? The unified oblique projection view*, in "27th International Conference on Machine Learning - ICML 2010", Haïfa, Israël, 2010, http://hal.inria.fr/inria-00537403/en.

[38] J. Siebert, L. Ciarletta, V. Chevrier. *Agents and artefacts for multiple models co-evolution. Building complex system simulation as a set of interacting models*, in "Autonomous Agents and Multiagent Systems - AAMAS 2010", Toronto, Canada, ACM, May 2010, vol. 1, p. 509-516, http://hal.inria.fr/hal-00452865/en.

[39] J. SIEBERT, L. CIARLETTA, V. CHEVRIER. *Agents and Artefacts for Multiple Models coordination. Objective and decentralized coordination of simulators.*, in "25th Symposium on Applied Computing - SAC 2010", Sierre, Suisse, ACM, March 2010, p. 2024-2028, http://hal.inria.fr/hal-00426601/en.

[40] C. THIERY, B. SCHERRER. *Least-Squares λ Policy Iteration: Bias-Variance Trade-off in Control Problems*, in "International Conference on Machine Learning", Haifa, Israël, 2010, http://hal.inria.fr/inria-00520841/en.

[41] N. VLASSOPOULOS, N. FATÈS, H. BERRY, B. GIRAU. *An FPGA Design for the Stochastic Greenberg-Hastings Cellular Automata*, in "International Conference on High Performance Computing & Simulation - HPCS 2010", Caen, France, 2010, p. 565-574, http://hal.inria.fr/inria-00522193/en.

National Peer-Reviewed Conference/Proceedings

[42] A. BRUN, A. HAMAD, O. BUFFET, A. BOYER. *Vers l'utilisation de relations de préférence pour le filtrage collaboratif*, in "17eme congrès francophone Reconnaissance des Formes et Intelligence Artificielle - RFIA 2010", Caen, France, Université de Caen Basse-Normandie / GREYC UMR 6072, January 2010, http://hal.inria.fr/inria-00432702/en.

[43] 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 "18èmes Journées Francophones des Systèmes Multi-Agents - JFSMA'10", Mahdia, Tunisie, October 2010, http://hal.inria.fr/inria-00534600/en.

[44] 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.

Workshops without Proceedings

[45] R. ARAS, O. PIETQUIN. *Optimal Average Reward Controllers For POMDPs*, in "5èmes Journées Francophones de Planification, Décision et Apprentissage pour la conduite de systèmes (JFPDA'10)", Besançon, France, June 2010, http://hal.inria.fr/hal-00554373/en.

[46] C. THIERY, B. SCHERRER. *Least-Squares λ Policy Iteration : optimisme et compromis biais-variance pour le contrôle optimal*, in "Journées Francophones de Planification, Décision et Apprentissage pour la conduite de systèmes", Besançon, France, 2010, http://hal.inria.fr/inria-00520843/en.

Scientific Books (or Scientific Book chapters)

[47] *Proceedings of the 20th International Conference on Automated Planning and Scheduling (ICAPS'10)*, AAAI Press, 2010, 274, http://hal.inria.fr/inria-00491134/en.

[48] A. BEYNIER, F. CHARPILLET, D. SZER, A.-I. MOUADDIB. *Dec-MDP/POMDP*, in "Markov Decision Processes in Artificial Intelligence", O. SIGAUD, O. BUFFET (editors), ISTE Ltd and John Wiley & Sons Inc, 2010, http://hal.inria.fr/inria-00439121/en.

[49] O. BUFFET. *Policy-Gradient Algorithms*, in "Markov Decision Processes in Artificial Intelligence", O. SIGAUD, O. BUFFET (editors), ISTE Ltd and John Wiley & Sons Inc, 2010, http://hal.inria.fr/inria-00439128/en.

[50] A. DUTECH, B. SCHERRER. *Partially Observable Markov Decision Processes*, in "Markov Decision Processes in Artificial Intelligence", O. SIGAUD, O. BUFFET (editors), ISTE Ltd and John Wiley & Sons Inc, 2010, http://hal.inria.fr/inria-00439123/en.

[51] N. FATÈS. *Does Life resist asynchrony ?*, in "Game of Life Cellular Automata", A. ADAMATZKY (editor), Springer London, 2010, p. p. 257-274 [*DOI :* 10.1007/978-1-84996-217-9_14], http://hal.inria.fr/inria-00433313/en.

[52] S. LE GLOANNEC, A.-I. MOUADDIB. *Resource Consumption Control for an Autonomous Robot*, in "Markov Decision Processes in Artificial Intelligence", O. SIGAUD, O. BUFFET (editors), ISTE Ltd and John Wiley & Sons Inc, 2010, http://hal.inria.fr/inria-00439130/en.

[53] S. THIÉBAUX, O. BUFFET. *Operations Planning*, in "Markov Decision Processes in Artificial Intelligence", O. SIGAUD, O. BUFFET (editors), ISTE Ltd and John Wiley & Sons Inc, 2010, http://hal.inria.fr/inria-00439126/en.

Books or Proceedings Editing

[54] N. FATÈS, J. KARI, T. WORSCH (editors). *Proceedings of AUTOMATA 2010: 16th International workshop on cellular automata and discrete complex systems*, INRIA Nancy Grand Est, 2010, 356, http://hal.inria.fr/inria-00549645/en.

[55] O. SIGAUD, O. BUFFET. , O. SIGAUD, O. BUFFET (editors)*Markov Decision Processes in Artificial Intelligence*, ISTE Ltd and John Wiley & Sons Inc, 2010, http://hal.inria.fr/inria-00432735/en.

Research Reports

[56] S. ALBOUZE. *Stratégie réactive pour le croisement sans arrêt de deux platoons de véhicules décentralisés*, INRIA, February 2010, http://hal.inria.fr/inria-00549368/en.

[57] M. ARAYA-LÓPEZ, O. BUFFET, V. THOMAS, F. CHARPILLET. *A POMDP Extension with Belief-dependent Rewards (Extended Version)*, INRIA, October 2010, n$^{\text{o}}$ RR-7433, http://hal.inria.fr/inria-00529498/en.

[58] O. BUFFET, L. CUCU-GROSJEAN. *Systematic Searches for Global Multiprocessor Real-Time Scheduling*, INRIA, September 2010, n$^{\text{o}}$ RR-7386, http://hal.inria.fr/inria-00519324/en.

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

[60] N. EL ZOGHBY. *Reinforcement Learning for Safe Longitudinal Platoons*, INRIA, September 2010, http://hal.inria.fr/inria-00551695/en.

[61] S. JACQUIN. *Planification de chemins à courbure C^n*, INRIA, September 2010, http://hal.inria.fr/inria-00550053/en.

[62] L. SARZYNIEC. *Apprentissage par renforcement développemental pour la robotique autonome*, INRIA, 2010, http://hal.inria.fr/inria-00546983/en.

[63] B. Scherrer, C. Thiery. *Performance bound for Approximate Optimistic Policy Iteration*, INRIA, 2010, http://hal.inria.fr/inria-00480952/en.

[64] A. Scheuer. *Documentation of ISeeML (Introducing a Smooth, Efficient and Easy-to-use Motion Library)*, INRIA, September 2010, n^o RT-0396, http://hal.inria.fr/inria-00527913/en.

[65] J. Yazbeck. *Decentralized local approach for lateral control of platoons*, INRIA, September 2010, http://hal.inria.fr/inria-00547567/en.

Other Publications

[66] S. Belgacem, N. Fatès. *Three Emergent Phenomena in the Multi-Turmite System and their Robustness to Asynchrony*, http://hal.inria.fr/inria-00462438/en.

[67] E. Daucé, A. Dutech. *Inverse Kinematics On-line Learning: a Kernel-Based Policy-Gradient approach*, 2010, http://hal.inria.fr/inria-00520960/en.

[68] N. Fatès. *Stochastic Cellular Automata Solve the Density Classification Problem with an Arbitrary Precision*, A paraître dans les actes de STACS 2011., http://hal.inria.fr/inria-00521415/en.

[69] N. Vlassopoulos, N. Fatès. *How Fast can Virtual Amoebae Aggregate? Analysis for the Optimal Firing Rate in an Instance of the Reaction-Diffusion-Chemotaxis Aggregation Scheme*, http://hal.inria.fr/inria-00462156/en.

References in notes

[70] R. Aras, A. Dutech, F. Charpillet. *Mixed integer linear programming for exact finite-horizon planning in decentralized POMDPs*, in "Proceedings of the Thirteenth International Conference on Automated Planning and Scheduling", 2007.

[71] R. Beckers, O. Holland, J.-L. Deneubourg. *From local actions to global tasks: stigmergy and collective robotics*, in "Artificial Life IV: Proc. of the 4th Int. Workshop on the synthesis and the simulation of living systems, third edition, MIT Press", 1994.

[72] D. Bertsekas, S. Ioffe. *Temporal differences-based policy iteration and applications in neuro-dynamic programming*, MIT, 1996.

[73] P. Daviet, M. Parent. *Longitudinal and lateral servoing of vehicles in a platoon*, in "Proc. of the IEEE Int. Symp. on Intelligent Vehicles", Tokyo (JP), September 1996, p. 41–46.

[74] L. E. Dubins. *On curves of minimal length with a constraint on average curvature, and with prescribed initial and terminal positions and tangents*, in "American Journal of Mathematics", 1957, vol. 79, p. 497–516.

[75] M. Ghallab, D. Nau, P. Traverso. *Automated Planning: Theory and Practice*, Morgan Kaufmann, 2004.

[76] 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.

[77] M. G. LAGOUDAKIS, R. PARR. *Least-Squares Policy Iteration*, in "Journal of Machine Learning Research", 2003, vol. 4, p. 1107-1149.

[78] M. LUNGARELLA, G. METTA, R. PFEIFER, G. SANDINI. *Developmental robotics: a survey*, in "Connection Science", 2003, vol. 15, n° 4, p. 151–190, http://dx.doi.org/10.1080/09540090310001655110.

[79] M. MAMEI, F. ZAMBONELLI. *Field-Based Coordination for Pervasive Multiagent Systems*, Springer, 2006.

[80] S. ONG, S. PNG, D. HSU, W. LEE. *POMDPs for robotic tasks with mixed observability*, in "Proceedings of Robotics: Science and Systems V (RSS'09)", 2009.

[81] N. PEPIN, O. SIMONIN, F. CHARPILLET. *Intelligent Tiles: Putting Situated Multi-Agents Models in Real World*, in "International Conference on Agents and Artificial Intelligence - ICAART'09", Porto Portugal, ACM AAAI, 2009, http://hal.inria.fr/inria-00339231/en/.

[82] J. PETERS, S. SCHAAL. *Natural Actor-Critic*, in "Neurocomputing", March 2008, vol. 71, n° 7-9, p. 1180–1190, http://dx.doi.org/10.1016/j.neucom.2007.11.026.

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

[84] C. REYNOLDS. *Flocks, Herd and Schools: A Distributed Behavioral Model*, in "Computer Graphics", 1987, vol. Vol. 21, No 4, p. 25-24.

[85] A. SCHEUER, T. FRAICHARD. *Continuous-Curvature Path Planning for Car-Like Vehicles*, in "Proc. of the IEEE-RSJ Int. Conf. on Intelligent Robots and Systems", Grenoble (FR), September 1997, vol. 2, p. 997–1003.

[86] A. SCHEUER, C. LAUGIER. *Planning Sub-Optimal and Continuous-Curvature Paths for Car-Like Robots*, in "Proc. of the IEEE-RSJ Int. Conf. on Intelligent Robots and Systems", Victoria, BC (CA), October 1998, vol. 1, p. 25–31.

[87] 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.

[88] R. SCHOKNECHT. *Optimality of Reinforcement Learning Algorithms with Linear Function Approximation*, in "NIPS", 2002, p. 1555-1562.

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

[90] D. SZER, F. CHARPILLET. *Point-based dynamic programming for DEC-POMDPs*, in "Proceedings of the Twenty-First AAAI National Conference on Artificial Intelligence", 2006.

[91] D. SZER, F. CHARPILLET, S. ZILBERSTEIN. *MAA*: A Heuristic Search Algorithm for Solving Decentralized POMDPs*, in "Proceedings of the Twenty-First Conference on Uncertainty in Artificial Intelligence", AUAI Press, 2005, p. 576–583.

[92] V. THOMAS, C. BOURJOT, V. CHEVRIER, D. DESOR. *Hamelin: A model for collective adaptation based on internal stimuli*, in "From animal to animats 8 - Eighth International Conference on the Simulation of Adaptive Behaviour 2004 - SAB'04, Los Angeles, USA", S. SCHAAL, A. IJSPEERT, A. BILLARD, S. VIJAYAKUMAR, J. HALLAM, J.-A. MEYER (editors), Jul 2004, p. 425-434.

[93] V. THOMAS, M. AKPLOGAN. *Using Social actions and RL-algorithms to build policies in DEC-POMDP*, in "IADIS International Journal on Computer Science and Information Systems", 2009, vol. 4, p. 82-98, I.2.11.3: Multiagent systems.

[94] H. YU, D. BERTSEKAS. *New Error Bounds for Approximations from Projected Linear Equations*, Dept. Computer Science, Univ. of Helsinki, July 2008, n^o C-2008-43.