

*Projet MAIA**MAchine Intelligente Autonome**Lorraine*

THÈME 3A



*R*apport
*d'**A*ctivité

2002

Table des matières

1. Composition de l'équipe	1
2. Présentation et objectifs généraux	1
2.1. Présentation	1
2.2. Défis scientifiques	2
2.2.1. Conception d'agents rationnels	2
2.2.2. Conception d'agents situés	2
2.2.3. Conception d'agents coopérants	2
2.2.4. Vers une théorie de la rationalité dans des univers multi-agents	3
3. Fondements scientifiques	4
3.1. Introduction	4
3.2. Approches d'inspiration mathématique	4
3.2.1. Objectifs	4
3.2.2. Notre approche	4
3.2.3. Calcul d'une politique optimale	5
3.2.4. Modèles distribués décisionnels de Markov pour la résolution coopérative de problèmes	5
3.2.5. Apprentissage par renforcement	6
3.2.6. Apprentissage par renforcement de modèles décisionnels de Markov pour la résolution coopérative de problèmes	6
3.2.7. Apprentissage supervisé	6
3.2.8. Identification de l'état interne d'un agent	6
3.2.9. Modèles continus	7
3.2.10. Situation dans la communauté nationale et internationale	7
3.3. Modèles d'inspiration biologique	8
3.3.1. Objectifs	8
3.3.2. Notre approche	8
3.3.3. Comprendre l'émergence	8
3.3.4. Transposition	8
3.3.5. Formalisation, généralisation, abstraction	9
3.3.6. Situation dans la communauté nationale/internationale	9
4. Domaines d'application	9
6. Résultats nouveaux	10
6.1. Modèles d'inspiration mathématique	10
6.1.1. Apprentissage par Renforcement de Comportements Complexes	10
6.1.2. Interaction vocale homme machine	10
6.1.3. Approximation d'un MDP par agrégation	11
6.1.4. Auto-organisation modulaire / Classification de MDP	11
6.1.5. Apprentissage par renforcement et connexionnisme	11
6.1.6. Modèle hybride (probabiliste/analytique) pour la planification de trajectoire	12
6.2. Modèles d'inspiration biologique	12
6.2.1. Simulation de la différenciation sociale chez les RATs	12
6.2.2. Détection de région bio-inspirée des araignées sociales	12
6.2.3. Localisation de robots mobiles dans un environnement partiellement connu en utilisant une architecture à base de Systèmes Multi-Agents.	13
7. Contrats industriels	13
7.1. Collaboration avec DIALOCA	13
8. Actions régionales, nationales et internationales	14
8.1. Actions régionales	14

8.1.1.	TAM	14
8.1.2.	Télémédecine	14
8.2.	Actions nationales	15
8.2.1.	ACI Temps et Cerveau	15
8.2.2.	RNTS DEPIC	16
8.2.3.	Pré-projet ROBEA intitulé COROCOP : Contrôle de robots coopératifs planétaires	16
8.3.	Relations internationales	16
8.3.1.	Le projet IST ELIN	16
8.3.2.	Le projet Eureka KVM	17
8.3.3.	Le projet IST OZONE	17
8.3.4.	Le projet ITEA Protéus	17
8.4.	Relations bilatérales internationales	18
8.4.1.	Projet avec Umass	18
8.4.2.	Projet avec la NASA : « self-Directed Cooperative Planetary Rovers (Cross Enterprise Space Technology Development Program de la NASA) ».	18
8.5.	Participation à des colloques, séminaires, invitations	19
10.	Bibliographie	19

1. Composition de l'équipe

Mots clefs :

Intelligence Artificielle distribuée, théorie de la décision, raisonnement sous incertitude, apprentissage, planification, diagnostic

Responsable Scientifique

François Charpillat [Directeur de Recherche, INRIA]

Assistante de projet

Martine Kuhlmann [CNRS, à temps partiel]

Personnel Université

Christine Bourjot [Maître de Conférences, U. Nancy 2]

Anne Boyer [Maître de Conférences, U. Nancy 2]

Vincent Chevrier [Maître de Conférences, U. H. Poincaré, Nancy 1]

Jean-Paul Haton [Professeur IUF, U. H. Poincaré, Nancy 1]

Alexis Scheuer [Maître de Conférences, U. H. Poincaré, Nancy 1]

Personnel INRIA

Alain Dutech [Chargé de Recherche, INRIA]

Chercheurs doctorants

Olivier Buffet [allocataire MENRT]

Iadine Chadès [ATER, U. Nancy 2, thèse soutenue le 6 janvier 2003]

Romarc Charton [boursier CIFRE MIC2]

Frank Gechter [ATER, IUFM de Lorraine]

Laurent Piraud [Université H. Poincaré, Nancy 1]

Bruno Scherrer [ATER, U.H. Poincaré, Nancy 1, thèse soutenue le 6 janvier 2003]

Vincent Thomas [allocataire MENRT]

Chercheurs post-doctorants

David Bellot [ATER, IMAG Grenoble, , thèse soutenue le 26 novembre 2002]

Laurent Jeanpierre [ATER, U. H. Poincaré, Nancy 1, , thèse soutenue le 3 décembre 2002]

Ingénieurs (Poste d'accueil)

Helder-Renato Da Costa [ingénieur U. H. Poincaré, Nancy 1, depuis decembre 2002]

Régis Lhoste [ingénieur expert INRIA]

Eric Lucchese [ingénieur d'étude U. H. Poincaré, Nancy 1]

Loïc Pelissier [ingénieur expert INRIA]

Rédimé Hadji-Pfeffer [ingénieur expert INRIA]

Stagiaires

Hubert Cecotti [stagiaire DEA, jusqu'en septembre 2002]

Cherif Smaili [stagiaire DRT, depuis novembre 2002]

2. Présentation et objectifs généraux

2.1. Présentation

Les travaux menés par le projet MAIA se situent à l'intersection des domaines de la théorie de la décision et des systèmes multi-agents. Nos recherches visent à définir les outils permettant la synthèse, la modélisation et la simulation de systèmes composés d'un ensemble d'agents autonomes situés dans un environnement physique ou virtuel. Le concept d'agent que nous considérons est large. Un agent est une entité capable de percevoir localement son environnement, capable d'interagir avec son environnement et d'autres agents, capable de

satisfaire des objectifs individuels et possédant des compétences propres qu'il peut mettre à la disposition des autres agents. Un agent peut donc tout aussi bien désigner un logiciel, un robot, un humain ou un animal.

2.2. Défis scientifiques

2.2.1. Conception d'agents rationnels

Un agent se définit donc comme un système physique ou logiciel doté de fonctions sensori-motrices et cognitives qui adopte un comportement qu'un observateur extérieur pourra qualifier d'intelligent. Cet observateur pourra alors attribuer à l'agent des qualités qu'on associe habituellement à un être humain. Il pourra dire " cet agent est serviable, autonome, flexible, réactif ", et le concepteur pourra revendiquer le statut d'agent intelligent pour le programme qui exhibe ces propriétés. Nous souhaitons dépasser cette démarche pour le moins subjective et difficilement quantifiable d'un point de vue scientifique.

Pour cela, parmi les théories candidates pour formaliser la conception d'agents intelligents, la théorie de l'intelligence artificielle fondée sur la rationalité¹, proposée par Stuart Russell en 1995, nous semble être la meilleure pour fonder nos travaux de recherche. La théorie de l'intelligence artificielle fondée sur la rationalité s'inscrit dans une démarche initiée par des philosophes, théoriciens du contrôle et économistes dont l'aboutissement s'inscrit dans la théorie de la décision par une formalisation fondée sur l'utilité. Un agent rationnel est un agent qui adopte un comportement optimal vis-à-vis d'un certain degré de satisfaction qu'il peut mesurer par ses propres moyens. La mesure de la satisfaction d'un agent est évidemment subjective, et elle dépend du contexte de l'application. Mais le biais qui est ainsi introduit permet d'exprimer formellement ce que signifie " développer un agent rationnel " (sous entendu intelligent). Un comportement se définit simplement dans ce cadre comme une fonction qui à un historique d'observations associe une action. Ce comportement peut être obtenue par calcul (planification) lorsqu'un modèle de l'environnement est disponible, ou par apprentissage dans le cas contraire.

Si cette approche fournit un cadre formel satisfaisant, elle pose encore bon nombre de questions auxquelles nous devons répondre. Parmi celles-ci, le problème de la représentation du monde et des autres agents, et en particulier dans les domaines où l'espace des états possibles est un espace continu, la définition de la notion d'utilité d'une action, le calcul d'un comportement optimal que ce soit par apprentissage ou par planification font l'objet des recherches fondamentales qui nous préoccupent.

2.2.2. Conception d'agents situés

Les agents que nous développons sont situés ou plutôt en situation, c'est-à-dire qu'ils sont immergés dans leur environnement d'intervention. Une des difficultés majeures qu'a à surmonter un tel agent réside dans le traitement de l'incertitude dans laquelle il est plongé : incertitude liée à l'imperfection des capteurs dont il est doté, incertitude liée à la localité des informations dont il dispose, incertitude liée à une mauvaise connaissance des lois régissant l'évolution de l'environnement, les autres agents ou l'utilisateur. Si nous prenons l'exemple du déplacement d'un robot mobile, la trajectoire suivie n'est pas toujours conforme à ce que le robot a planifié du fait par exemple du glissement des roues sur le sol, ou de la présence d'obstacles imprévus. Pour aborder cette notion d'incertitude, les Processus Décisionnels de Markov et les modèles graphiques fournissent un cadre théorique pour décider des actions à entreprendre lorsque leur effet est incertain. Si cette approche est séduisante sur le plan théorique, son implantation informatique pose encore aujourd'hui de nombreux problèmes, notamment en ce qui concerne les performances en temps des algorithmes sous-jacents et le problème de l'apprentissage des modèles. Pour pallier cette difficulté, nous explorons diverses voies : approximations, techniques "anytime", utilisation de plans sous-optimaux mais néanmoins satisfaisants, apprentissage supervisé, apprentissage par renforcement.

2.2.3. Conception d'agents coopérants

Un autre objectif important de notre équipe est de montrer qu'un ensemble d'entités simples, appelées agents, dotées individuellement de capacités sensori-motrices peut résoudre collectivement un problème complexe.

¹Stuart Russel, « Rationality and Intelligence », Invited paper (Computers and Thought Award), in Proc. *Fourteenth International Joint Conference on Artificial Intelligence*, Montreal, Canada, 1995

Il y a deux approches possibles pour mettre en œuvre des systèmes composés d'agents simples résolvant collectivement un problème : l'approche constructiviste (approche descendante) et l'approche émergentiste (approche ascendante).

1. approche constructiviste

Dans la première approche, le comportement individuel de chaque agent est obtenu par calcul (par planification) ou par apprentissage (apprentissage par renforcement). Nos efforts dans ce contexte, portent notamment sur la conception de modèles issus de la théorie de la décision. Nous sommes confortés dans cette démarche par des résultats préliminaires qui nous assurent qu'il est possible de calculer dans le cadre de la théorie de la décision un ensemble de comportements individuels simples qui permettent à des agents de réaliser une tâche qu'aucun ne pourrait résoudre seul. En particulier, on peut formaliser les interactions d'une société d'agents et déterminer le comportement optimal de chacun pour atteindre un objectif donné dans le cadre des modèles de décision markovien. Un défi important pour nous est de construire une théorie de l'intelligence collective fondée sur la rationalité.

2. Approche émergentiste

Dans la seconde approche, le comportement individuel de chaque agent est donné en instanciant des modèles *a priori*. Dans ce dernier cas, les deux questions duales et fondamentales que nous nous posons sont de savoir :

- avec quels composants simples en interaction (niveau micro) peut on résoudre une tâche donnée (niveau macro) ?
- comment prédire qu'un comportement global et cohérent émergera à partir d'éléments de base en interaction ?

Nous sommes confortés dans cette démarche par les travaux récents issus de la biologie du comportement animal qui nous apprennent que des êtres aussi rudimentaires que des fourmis ou des araignées sont capables à partir de mécanismes extrêmement simples de réaliser en groupe des tâches complexes comme la construction d'un abri ou le fourragement. Certains systèmes biologiques, comme les sociétés d'insectes par exemple, parviennent malgré des capacités cognitives individuelles réduites à résoudre des problèmes complexes en mettant en place une certaine organisation par l'intermédiaire des seules interactions qu'ils entretiennent avec leur environnement. Ces organisations possèdent des propriétés émergentes non présentes chez l'individu mais qui vont permettre à la collectivité de résoudre le problème auquel elle est confrontée. Ces phénomènes constituent une source d'inspiration non négligeable pour nos travaux.

2.2.4. Vers une théorie de la rationalité dans des univers multi-agents

La résolution multi-agents de problèmes est une discipline qui intéresse un grand nombre de chercheurs en France et dans le monde. Malgré cela, très peu de chercheurs ont exploité la théorie de la décision et de la rationalité pour aborder et clarifier cette problématique. Craig Boutilier fait partie de ceux-ci. Il propose notamment un modèle de décision markovien multi-agent (MMDP)² qu'il limite à des problèmes dans lesquels les agents ont tous accès à la même information. On ne peut donc plus vraiment parler de systèmes multi-agents mais plutôt de résolution distribuée de problèmes, ce qui en soit est intéressant mais ne correspond pas vraiment à notre objectif de recherche. Des travaux dans le domaine de la théorie du contrôle paraissent plus intéressants^{3, 4}.

²Craig Boutilier, *Sequential Optimality and Coordination in Multiagent Systems*, In 16th International Joint Conference on Artificial Intelligence (Stockholm). 1999

³Ooi, J.M., and Wornell, G.W. 1996. A separation theorem for periodic sharing information patterns in decentralized control. IEEE transactions on Automatic Control 42(11) : 1546-1550.

⁴Aicardi, M. ; Franco, D. ; and Mincardi, R. 1987, Decentralized optimal control of Markov chains with a common past information set. IEEE transactions on Automatic Control 32(11) :1028-1031.

Nous pensons qu'une généralisation des POMDP qui permet de distribuer le contrôle sur m agents et que l'on nommera DEC-POMDP est envisageable. Un DEC-POMDP est défini par un ensemble fini d'états S , une fonction de transition $p(s'|s, a_1, \dots, a_m)$ entre les états S et la récompense $r(s, a_1, \dots, a_m)$ qui dépendent dans ce cadre des actions des autres agents. Chaque agent i connaît la probabilité d'observer un symbole o_i en supposant les actions a_1, \dots, a_m des autres agents et l'état s' dans lequel se trouve le système : $O_i(o_i|s', a_1, \dots, a_m)$. Le but est alors de calculer une politique optimale, c'est-à-dire m fonctions qui, pour chaque agent, associe une action à un historique des observations de sorte que l'espérance de récompense, qui est globale, soit maximisée. Actuellement, aucune algorithmique n'est disponible à notre connaissance pour exploiter ce modèle. Nous envisageons de combler cette lacune. Plus généralement nous comptons construire une théorie de l'intelligence collective fondée sur la rationalité.

3. Fondements scientifiques

3.1. Introduction

La modélisation de comportements est une activité centrale de notre projet qui se décline en deux grands domaines d'activité. La modélisation et la génération de comportements intelligents sont abordées par notre groupe selon une approche originale qui nous conduit tantôt à développer des modèles d'inspiration mathématique, au rang desquels figurent en bonne place les modèles stochastiques et graphiques, l'algorithmique du mouvement, la théorie de la décision et du contrôle, tantôt des modèles d'inspiration biologique. Cette orientation vers l'algorithmique numérique nous semble essentielle pour aborder la complexité des problèmes réels auxquels nous sommes confrontés. Cependant nous n'excluons pas à cours ou moyen terme le développement de modèles mixtes symboliques numériques qui seuls nous permettront de modéliser des tâches abstraites.

3.2. Approches d'inspiration mathématique

3.2.1. Objectifs

Définir un système à agents est une tâche difficile même si nous nous limitons pour l'instant à des systèmes simples, essentiellement réactifs et sans communication explicite. Définir un tel système consiste à définir une architecture, des modèles d'agents (dynamique de l'état interne, fonction de décision, ...), des modèles de l'environnement et de sa dynamique, un signal de renforcement. Les modèles de décision séquentielle permettent d'aborder cet objectif avec des outils mathématiques efficaces. Nous nous intéressons dans ce cadre aux algorithmes d'apprentissage par renforcement et de planification stochastique.

3.2.2. Notre approche

Les modèles de décision séquentielle relèvent de diverses approches selon que la fonction d'observation de l'agent permet de déterminer à coup sûr l'état dans lequel il se trouve ou qu'au contraire il reste partiellement indéterminé, selon que l'environnement est continu ou discret, selon que l'on possède un modèle explicite ou non de la dynamique de l'environnement et de la fonction de récompense ou renforcement. Les processus de décision Markoviens (MDP) et les processus de décision Markoviens partiellement observables (POMDP) sont destinés à la modélisation du comportement d'agents évoluant dans un environnement dont on connaît la dynamique et pour lesquels on peut expliciter la fonction de récompense. Les MDP sont dédiés aux agents pouvant connaître à tout instant l'état de l'environnement dans lequel ils sont situés. Les POMDP s'attachent à modéliser des agents ayant une fonction d'observation partielle de l'environnement. Lorsque ni la dynamique de l'environnement, ni la fonction de récompense ne sont connues, on a recours aux techniques d'apprentissage par renforcement. Elles permettent à un agent d'apprendre à se comporter de manière optimale grâce à une exploration de l'environnement. Celle-ci permet soit de découvrir les lois qui régissent la dynamique de l'environnement et les récompenses correspondantes soit d'apprendre directement le comportement optimal sans construire de modèle. Les modèles de décision séquentielle discrets sont ceux que nous étudions plus particulièrement actuellement. Dans ce cadre, l'espace des états possibles est fini et le temps est discrétisé. A

chaque instant, l'agent se trouve dans un certain état $s(t)$, et il choisit une action $a(t)$. Il reçoit pour cela une récompense $r(t)$. La réalisation de l'action $a(t)$ fait évoluer le processus dans l'état $s(t+1)$. L'état résultant $s(t+1)$ ne dépend que l'état $s(t)$ et de l'action $a(t)$ car on suppose le processus markovien ce qui permet de définir une fonction de transition $T(s(t),a(t),s(t+1))=p(s(t+1)|s(t),a(t))$. Dans ce qui suit, nous exposons les problèmes que nous tentons de résoudre.

3.2.3. Calcul d'une politique optimale

Il s'agit de calculer la politique Π qui maximise l'espérance de récompense à venir. Etant donné que la récompense $r(t)$ ne dépend que de l'état $s(t)$ et de l'action $a(t)$ et comme on suppose le processus markovien on peut montrer que la fonction de valeur $V_{\Pi(x)}$ pour la meilleur politique Π vérifie l'équation : $V(x(t)) = \text{Max}_{\Pi(t)} R(x(t), \Pi(t)) + \gamma \sum_{x(t+1)} T(x(t), x(t+1), \Pi(t)) V(x(t+1))$ appelée équation de Bellman. La résolution de cette équation par programmation dynamique permet de calculer la politique optimale Π . Dans le cadre d'un POMDP les états ne sont pas identifiables à partir des observations. Un agent ne connaît donc pas à chaque instant l'état dans lequel il se trouve. La politique optimale obtenue par l'équation de Bellman n'est donc pas directement applicable pour des environnements partiellement observables. Des travaux théoriques ont montré qu'il est possible de résoudre ce problème en passant par ce qui est appelé des belief states. Cependant, les algorithmes qui en sont dérivés sont trop complexes pour pouvoir traiter des cas avec plus d'une dizaine d'états. Les algorithmes de calcul des plans ont une complexité qui les rend inutilisables en temps-réel. Cette complexité est liée notamment au nombre d'états nécessaires pour représenter un environnement, nombre qui est souvent très important (plusieurs milliers dans nos applications à la robotique). Il est alors nécessaire de trouver des méthodes permettant de calculer les plans plus rapidement. Nous nous intéressons aux méthodes permettant d'obtenir des plans sous-optimaux (mais néanmoins proches de la solution optimale), en un temps de calcul "raisonnable".

3.2.4. Modèles distribués décisionnels de Markov pour la résolution coopérative de problèmes

Il existe dans la littérature un modèle théorique qui prend en compte les contraintes de la catégorie de problèmes auxquels nous nous intéressons : il s'agit des DEC-POMDP (Decentralized Partially Observable MDP). Il a été récemment montré que la résolution est NEXP-complet lorsque le nombre d'agents considéré est supérieur ou égal à 2, en effet la résolution dépend du nombre d'états et du nombre d'actions du DEC-POMDP. Elle ne convient donc pas pour une approche réaliste de la coordination des agents dans le cadre de notre SMA évoluant dans un environnement complexe. De plus, bien que le contrôle de ce modèle soit décentralisé, sa résolution et son utilisation restent centralisées.

Nous avons donc adapté le formalisme MDP dans le cadre des SMA afin qu'il prenne en compte les propriétés essentielles qui sont pour nous : la localité des perceptions et des actions, mais aussi la poursuite d'un but local pour chaque agent. Les interactions entre nos agents permettent alors l'émergence d'une coordination qui permet d'atteindre le but global de nos agents. Notre solution, qui s'appuie sur le formalisme d'un MDP, est fondée sur deux propriétés essentielles et indépendantes de nos agents : l'**empathie** et la **subjectivité** (voir [9]).

La construction de l'architecture des agents, leur perception de l'environnement, leur comportement et l'analyse de l'environnement ont été réalisés. Le nombre d'états de chaque MDP reste constant quelle que soit la taille de notre environnement. Ceci est possible grâce à la **subjectivité** de nos agents qui place chaque agent au centre de son repère d'évolution. Nous l'avons traduite entre autres par l'utilisation d'états agrégés, et le respect d'autonomie de chaque agent. La notion d'**empathie** se traduit, elle, par la capacité que chaque agent a à se mettre à la place de son « frère » et à prévoir le comportement d'autrui afin d'en améliorer son comportement (sorte de méta-raisonnement). Ainsi, nous avons exhibé un algorithme de résolution sur des populations d'agents homogènes et hétérogènes.

A ce jour, nous avons montré que notre modèle constituait une bonne heuristique pour la résolution de problèmes de la catégorie des DEC-POMDPs à condition de garder nos hypothèses (restrictives) sur les capacités de nos agents (nos agents n'ont pas de mémoire) et sur la fonction de récompense globale qui peut être divisée en fonctions de récompense locales pour chaque agent.

3.2.5. Apprentissage par renforcement

Il s'agit d'apprendre la politique Π en effectuant une exploration de l'environnement. La question est de savoir si un agent peut apprendre, par lui-même, à partir d'interactions avec son environnement, le comportement optimal Π . Nous nous intéressons plus particulièrement aux problèmes partiellement observables. Dans ce cadre, après chaque interaction, l'agent reçoit une observation $o(t)$ qui lui donne une indication partielle sur l'état dans lequel il se trouve, et l'agent choisit une action a qu'il exécute. Cette action modifie l'état dans lequel se trouve l'agent et il reçoit une récompense ou un signal de renforcement $r(s)$. On souhaite que l'agent choisisse les actions qui augmentent la somme des récompenses qu'il recevra au cours du temps.

3.2.6. Apprentissage par renforcement de modèles décisionnels de Markov pour la résolution coopérative de problèmes

L'approche multi-agents pour la résolution de problèmes présente un certain nombre de caractéristiques intéressantes telles que robustesse, modularité, etc. Mais elle introduit aussi une difficulté au niveau de la conception du système multi-agents, c'est à dire des agents eux-mêmes. Il s'agit en effet que de l'interaction d'un groupe d'agents apparaisse un comportement répondant à une certaine attente, problème aujourd'hui loin d'être résolu. Peu de théorie a été développée dans ce domaine des phénomènes émergents, ce qui laisse sans outils dédiés à la conception d'agents coopérants.

Pour pallier cette difficulté, nous nous intéressons à une méthode de conception par apprentissage de tels systèmes multi-agents. Chaque agent apprend seul le comportement à adopter, et ce par renforcement, recevant une récompense quand son but est atteint et une punition en cas d'événement à éviter. Comme les agents apprennent simultanément, les comportements s'adaptent les uns aux autres et s'accordent pour atteindre un comportement de groupe. Le formalisme utilisé est celui des Processus de Décision Markoviens, en utilisant les algorithmes d'apprentissage associés (Q -learning, $TD(\lambda)$, gradient...).

L'apprentissage n'est toutefois pas une solution simple et définitive au problème de la conception d'agents coopérants. Diverses difficultés apparaissent, telles que : la non-stationnarité des modèles en raison des perceptions locales et de la présence d'autres agents, le problème de l'attribution du mérite⁵ rendu plus complexe par la présence de nombreux agents... Nous avons développé une première méthodologie pour pallier ces difficultés. Elle consiste en un apprentissage incrémental : difficulté croissante de la tâche, et nombre croissant d'agents dans l'environnement.

Nos travaux en cours [30] s'intéressent pour leur part à la complexité de l'environnement. En effet, dans un cadre multi-agents le nombre d'objets (agents compris) qu'un agent doit considérer est souvent grand, ce qui va compliquer l'apprentissage. Ainsi, au lieu d'apprendre directement le comportement à adopter dans le cas général, nous cherchons à reconstituer ce comportement à partir de comportements élémentaires appris indépendamment les uns des autres. Mais d'autres approches utilisant communication ou stigmergie par exemple sont envisagées pour continuer à améliorer l'apprentissage dans les systèmes multi-agents.

3.2.7. Apprentissage supervisé

Il s'agit d'apprendre à partir d'exemples d'interactions agent/environnement les paramètres d'un modèle. Plus formellement, il s'agit, étant donnés une séquence de vecteurs d'observation ($o_1, a_1, \dots, o_t, a_t$) et un modèle λ , d'ajuster ou calculer les paramètres d'un modèle de décision séquentielle pour maximiser $p(o_1, a_1, \dots, o_T, a_T | \lambda)$ [11]. Ce problème est au cœur du projet Diatelic (télé-surveillance des malades traités en dialyse péritonéale à domicile) ou de problèmes liés à la robotique. Un autre axe de recherche porte sur l'utilisation de réseaux bayésiens hybrides (on sait que les RB généralisent de nombreux formalismes comme les HMM, certains réseaux neuro-mimétiques) et hiérarchiques (pour travailler à des niveaux d'abstraction différents).

3.2.8. Identification de l'état interne d'un agent

Etant donnés une séquence de vecteurs d'observation ($o_1, a_1, \dots, o_T, a_T$) on se pose la question de savoir quelle est la probabilité que le système se trouve dans l'état s à l'instant t . Cette problématique se pose par exemple

⁵Credit Assignment Problem.

dans le domaine de la robotique mobile où il s'agit de localiser un robot à partir des seules perceptions qu'il a de son environnement. Il se pose également dans le domaine médical lorsqu'il s'agit d'identifier l'état pathologique d'un patient à partir des mesures de certains paramètres physiologiques.

3.2.9. Modèles continus

La modélisation des modèles de décision séquentielle dans les domaines où l'espace des états possibles est un espace continu [12],[31] est un sujet important pour appréhender la réalité physique des phénomènes que nous étudions. Notre approche consiste à définir des modèles subjectifs qui approximent la réalité par des modèles de décision de taille réduite fixée. Nous avons dérivé des critères qui permettent de déterminer les zones de l'espace d'état qui nécessitent d'être décrites avec plus de précision ou qui a contrario ne sont pas nécessaires pour l'approximation. Ainsi, de manière progressive, nous pouvons faire évoluer le modèle abstrait approxinant le modèle réel : nous mettons à jour la représentation, ce qui permet d'améliorer les performances sans changer la complexité de résolution.

3.2.10. Situation dans la communauté nationale et internationale

En ce qui concerne la planification stochastique, depuis le milieu des années 90, les modèles décisionnels de Markov connaissent un succès grandissant dans la communauté de recherche en IA et de plus en plus de chercheurs y font référence. Notre groupe a participé^{6,7,8,9}(notamment au cours l'ARC INRIA LIRE en particulier avec Philippe Chassaing du projet Oméga, par la coorganisation de Workshop à AAAI, IJCAI et ECAI) et continue à contribuer au développement de cette théorie au travers d'une collaboration avec Shlomo Zilberstein (2 projets NSF-INRIA) et d'une autre sous l'impulsion de la Nasa intitulée " self-Directed Cooperative Planetary Rovers " avec Shlomo Zilberstein, et Victor Lesser de l'University of Massachusetts, Eric Hansen de Mississippi State University, Richard Washington, du NASA Ames Research Center et Abdelillah Mouaddib du CRIL (Lens).

Dans les deux grandes familles d'apprentissage et de planification que nous utilisons, les fondements théoriques ont été exploités pour proposer des algorithmes exacts mais de complexités trop élevées pour être utilisés en pratique (algorithme »JLO » pour les réseaux bayésiens, « Q-Learning », « $TD(\lambda)$ », « Witness » pour le formalisme markovien). Si la plupart de ces travaux sont américains, la communauté française se développe rapidement. Notre équipe contribue directement à ce développement, notamment par son implication dans la formation d'un groupe de travail français sur les MDP. Nous sommes aujourd'hui un certain nombre d'équipes en France qui étudions les processus décisionnels de Markov. On trouve ainsi, à Toulouse des chercheurs de l'IRIT, du CERT, de l'INRA, du LAAS, à Lens des chercheurs du CRIL quelques chercheurs parisiens (Remy Munos, Olivier Sigaud).

La plupart des travaux actuels se concentrent donc sur des méthodes approchées. Notre particularité, en sus de l'intégration de ces méthodes dans un cadre multi-agents, est de diminuer la complexité de mise en œuvre des algorithmes en tirant parti de méta-connaissances sur le système à modéliser. Ainsi, nos algorithmes s'appuient sur les régularités temporelles, spatiales et structurelles du problème pour obtenir plus facilement et plus rapidement des solutions approchées au problème donné. De plus, nous avons comme volonté de coupler étroitement ces deux formes d'apprentissage et de les utiliser ensuite pour des applications de planification ou de reconnaissance pour des agents situés.

⁶Zilberstein, Shlomo and Charpillet, François and Chassaing, Philippe. Real-Time Problem-Solving with Contract Algorithms. In 16th International Joint Conference on Artificial Intelligence. (Stockholm). 1999. vol 2. pp. 1008-1013.

⁷F. Charpillet and A. Boyer. Progress : an Approach for Defining and Monitoring Non-deterministic Design to Time methods. In proc. of the ninth IEEE International Conference (ICTAI'97), Newport Beach, California, November 4-7, 1997.

⁸J.-M. Gallone and F. Charpillet. Hopfield Neural Network for Scheduling non Preemptive Tasks. in proc of 12th European Conference on Artificial Intelligence (ECAI'96). Budapest, Aout 96.

⁹A. I. Mouaddib and S. Zilberstein. Knowledge-Based Anytime Computation Proceedings of the 14th International Joint Conference on Artificial Intelligence, pp 775-781, Montreal, Canada, 1995.

3.3. Modèles d'inspiration biologique

3.3.1. Objectifs

Notre objectif concerne la proposition de modèles, outils et méthodes permettant la conception de système multi-agents selon une approche émergentiste. Il y a émergence dans un système multi-agents lorsque le comportement des agents est décrit dans un certain formalisme et les propriétés collectives observées sont décrites dans un formalisme différent qui ne se déduit pas du précédent. Cela pose alors le problème du passage entre la simplicité apparente individuelle d'un agent et la complexité globale du système. Autrement dit, il s'agit de concevoir des formes d'interaction entre les agents et leur environnement qui permettent l'observation de propriétés collectives complexes.

La problématique centrale consiste donc à établir une relation entre le niveau micro (les comportements individuels) et le niveau macro (le comportement collectif). Cette problématique peut se décomposer selon 2 perspectives complémentaires :

- étant donnée une propriété ou un comportement global désiré, comment construire les comportements individuels et la dynamique des interactions (approche descendante), pour obtenir un comportement collectif donné,
- à partir d'un ensemble de comportements individuels et de la dynamique des interactions comment prédire le comportement collectif résultant (approche ascendante).

3.3.2. Notre approche

Actuellement il n'existe pas de méthodologie conduisant à la construction de système multi-agents selon cette approche émergentiste et nous entendons contribuer à l'élaboration d'une telle méthodologie. Notre démarche s'articule autour de trois points :

- Comprendre l'émergence dans le cadre de phénomènes naturels
- Transposer les principes émergentistes pour résoudre des problèmes
- Proposer un cadre permettant de systématiser la conception émergentiste de système multi-agents.

3.3.3. Comprendre l'émergence

L'étude des modèles biologiques est une source d'inspiration qui semble relativement naturelle (sans pour autant être la seule possible). La réflexion débutera sur la base des travaux originaux déjà initiés dans l'équipe et qui concernent les araignées sociales. Chez les insectes sociaux, les interactions inter-individuelles reposent en grande partie sur l'utilisation d'un éventail de signaux chimiques caractérisés par des durées de vie, des concentrations et des critères de diffusion particulier. De plus, le groupe social est souvent divisé en castes constituées d'individus spécialisés. Le "vocabulaire chimique" des araignées sociales est beaucoup plus réduit, les interactions reposant essentiellement sur des informations émanant de la soie. Le piège correspond à un réseau soyeux tridimensionnel propice à la diffusion d'informations vibratoires. C'est une structure dynamique continuellement remaniée par les activités et les déplacements des individus, qui est également susceptible de remplir une fonction de "mémoire" collective en intégrant dans sa géométrie et ses propriétés mécaniques l'histoire du groupe et de son environnement.

L'étude de l'émergence vise à caractériser les processus sous-jacents à l'apparition de propriétés collectives émergentes à partir d'exemples concrets. L'objectif est d'obtenir un ensemble de caractérisations de mécanismes émergentistes à partir d'exemples existants. Une illustration de cette démarche se retrouve dans les travaux que nous avons déjà entrepris tels que la modélisation et la simulation d'activités collectives chez les araignées sociales : agrégation et synchronisation par exemple.

3.3.4. Transposition

L'étude et la modélisation de phénomènes naturels nous fournit le système de représentation et les comportements individuels, les mécanismes d'interaction et de réorganisation qui vont régir le fonctionnement du système afin qu'il produise une réponse au niveau collectif. Il s'agit alors d'en dégager les principes de

fonctionnement en faisant le plus possible abstraction du contexte naturel puis d'adapter ceux-ci à un problème ou une famille de problèmes.

Nous avons ainsi transposé avec succès le comportement de construction collective de toile chez les araignées pour détecter des régions homogènes dans des images à niveaux de gris.

3.3.5. Formalisation, généralisation, abstraction

Cette dernière étape vise à dépasser un cadre purement énumératif (un catalogue de " procédures " émergentistes permettant la construction de systèmes multi-agents) pour fournir des méthodes d'évaluation de l'émergence par l'élaboration de mesures de la complexité structurelle et/ou fonctionnelle, mesures s'apparentant à la notion d'entropie.

Le concept d'émergence est pluri-disciplinaire et fait référence à des travaux dans différents domaines (phénomènes non linéaires, théorie du chaos, systémique, biologie du comportement, etc.) où ses définitions ne sont pas utilisables en tant que telles pour construire et faire s'exécuter des systèmes informatiques. Nous estimons nécessaire d'étudier ces différentes définitions pour en extraire les éléments permettant la proposition d'un cadre constructif en informatique.

3.3.6. Situation dans la communauté nationale/internationale

Parmi les deux principales approches d'études des systèmes multi-agents nous avons orienté nos recherches vers les systèmes " collectifs " qui privilégient les notions d'interaction et d'organisation.

Cela s'est manifesté par de nombreuses collaborations avec les chercheurs francophones de ce domaine, notamment par une implication active dans différents groupes de travail :

- groupe de travail *MARCIA* (dès 1993 sous l'égide du GDR-PRC IA et pour lequel V. Chevrier a assuré la co-responsabilité de 95 à 97) et dont l'un des thèmes a été l'auto-organisation ;
- groupe de travail *Interaction* de l'AFCEC-AFIA (94-96),
- et groupe de travail *Colline* (sous l'égide du GDR I3 et de l'AFIA) à partir de 1997.

C'est ainsi que des modèles d'interaction et d'organisation ont été proposés dans les travaux de thèse de R. Foisel et A. Dury.

L'approche que nous suivons actuellement se fonde sur la notion d'auto-organisation pour la conception de systèmes multi-agents avec notamment l'étude de leurs propriétés émergentes. Même si la communauté informatique dans le domaine est encore réduite, elle se développe depuis quelques années notamment au travers des recherches menées à l'Iremia (Université de la Réunion), à l'IRIT, à l'Université de Neuchâtel (Suisse) et dans quelques laboratoires aux E-U, (D.Van Parunak, R Brooks par exemple).

Parmi ces recherches, certaines s'inspirent de modèles biologiques pour envisager l'émergence de propriétés. Les travaux actuels ont été principalement inspirés des modèles de société de fourmis comme au LIP6, au Lirmm ou encore à l'IRIDIA (Bruxelles). Le modèle des araignées sociales ne l'ayant jamais été, cela constitue une originalité de notre démarche. Il est à noter que pour une grande part, ce domaine a été influencé par les travaux de Jean-Louis Deneubourg du CENOLI (Bruxelles) concernant les phénomènes d'auto-organisation dans ces sociétés et les mécanismes d'interaction par phéromones dans le cas des fourmis.

4. Domaines d'application

Les domaines d'application de notre projet sont nombreux. Nous nous focalisons dans les domaines de la robotique mobile et médical (projet ADAM sur l'aide à l'anesthésie), dans les applications distribuées exploitant le web comme le télé-diagnostic (projet ITEA Proteus), la télémédecine(projet DIATELIC, projet DEPIC), l'ambiance intelligence (projets IST OZONE et Elin). Notre apport à ces domaines est relatif à notre capacité à mettre en oeuvre des outils de modélisation des processus dynamiques, partiellement observables. Notre maîtrise des techniques d'apprentissage par renforcement, d'apprentissage dans les modèles de markov cachés et les réseaux bayésiens est à l'origine du succès de nos réalisations. Parmi ces réalisations, citons, nos travaux sur la télé-surveillance de patients insuffisants rénaux traités par dialyse péritonéale. L'aboutissement

est un système qui permet de prévenir une mauvaise hydratation de ces patients. Ce système s'appuie sur un modèle obtenue par apprentissage qui est adapté à chaque patient et qui tient compte de la pratique du médecin traitant. Le succès de l'expérimentation médicale menée ces trois dernières années et la demande de la part de médecins néphrologues, libéraux, de la part d'associations spécialisées dans le traitement de l'insuffisance rénale, nous permettent d'envisager dès maintenant un passage à l'échelle avec comme premier objectif le traitement de 150 patients en Lorraine. Le LORIA n'ayant pas vocation à assurer un tel service, une start-up DIATELIC SA et l'association Diatélic Lorraine préfigurant un GIE ont été créées pour mettre en place ce service.

6. Résultats nouveaux

6.1. Modèles d'inspiration mathématique

6.1.1. Apprentissage par Renforcement de Comportements Complexes

Participants : Olivier Buffet, François Charpillet, Alain Dutech.

Le cadre de ce travail concerne l'apprentissage de comportements pour des agents devant faire face à des tâches complexes. Nous partons de l'hypothèse que cette tâche complexe peut se voir comme la résultante de motivations simples et distinctes. Dès lors, nous proposons de concevoir le comportement complexe des agents en combinant des comportements de base plus simples adaptés aux sous-tâches identifiées. En effet, alors que ces comportements de base peuvent être appris en utilisant les techniques classiques d'apprentissage par renforcement [35][33], des comportements plus complexes sont impossibles à apprendre de manière directe (problèmes d'explosion combinatoire et de minima locaux).

Dans le but d'automatiser le processus de construction du comportement résultant, les paramètres de la combinaison de comportements sont appris en maximisant la récompense globale de l'agent. De plus, en nous appuyant sur des *comportements typés* qui peuvent être instanciés plusieurs fois pour une même situation présentée à l'agent, nous obtenons des agents génériques et robustes. A la validation expérimentale de cet algorithme s'est ajoutée une validation de la communauté scientifique [19] et [18].

Le travail en cours vise l'automatisation complète du processus de construction en cherchant à identifier et abstraire les comportements de base requis. Une fois ces comportements identifiés, ils peuvent être réutilisés pour d'autres tâches. D'autre part, la complexité des problèmes abordés permet d'envisager l'utilisation de l'algorithme dans un cadre multi-agents. Néanmoins, il faudra se pencher plus avant sur les difficultés inhérentes à la coordination d'agents autonomes par des méthodes distribuées en s'inspirant, par exemple, des travaux de [20].

6.1.2. Interaction vocale homme machine

Participants : Anne Boyer, François Charpillet, Romaric Charton.

Dans le cadre de travaux sur la collaboration entre agents hétérogènes, nous avons étudié une approche centrée sur l'interaction entre un agent humain et un service de recherche d'informations. Nous avons proposé l'introduction d'un agent médiateur dont le rôle est de formaliser la requête de l'utilisateur et de lui proposer un nombre restreint et ciblé de réponses pertinentes. A cet effet, nous avons étudié une approche à base de modèles stochastiques (MDP) afin de déterminer à partir d'un modèle de service (description d'un référentiel, liste des interactions possibles et buts applicatifs) la meilleure stratégie de médiation. Ceci a conduit à l'implantation d'un prototype d'agent médiateur à base de Q -learning [36], [32], [22], [21]. L'espace d'états est construit à partir du référentiel et de la capacité de la source à apporter une réponse à la requête en cours de formalisation. Les actions consistent soit à poser une question à l'utilisateur soit à sonder la source (opération qui a un coût). Les récompenses sont données en fonction de la satisfaction de l'utilisateur, la longueur du dialogue et le coût de rapatriement des données de la source à l'utilisateur. Nous avons choisi d'utiliser un algorithme d'apprentissage par renforcement car d'une part nous ne disposons pas d'un modèle a priori, et d'autre part il permet une adaptation en ligne. L'application cible consiste à déterminer en fonction du profil de

l'utilisateur (habitudes et préférences constatées au cours de sessions passées) et de l'environnement (bruité ou non) la meilleure séquence de questions et la meilleure modalité pour la réaliser (dialogue oral ou texte par exemple). Les résultats obtenus sont encourageants et semblent bien adaptés à la gestion d'interactions ciblées. C'est pourquoi nous orientons nos travaux vers l'utilisation de modèles hiérarchiques pour des tâches plus complexes.

6.1.3. *Approximation d'un MDP par agrégation*

Participants : François Charpillet, Bruno Scherrer.

La complexité algorithmique pour exploiter un modèle de décision markovien est intrinsèquement dépendante de la taille de son espace d'états. Nous avons cette année approfondi nos travaux concernant l'utilisation d'approximations pour aborder des problèmes de décision markoviens ayant un grand espace d'états. Nous utilisons une méthode dite d'« agrégation d'états », qui permet d'approximer un modèle de grande taille à l'aide d'un modèle de taille réduite [12]. Des travaux récents de la littérature nous ont donné les moyens d'analyser théoriquement ce procédé approximatif. Ils nous ont de plus permis de proposer des algorithmes itératifs de mise à jour de l'agrégation. Partant d'une agrégation initiale quelconque, ces algorithmes permettent d'améliorer la qualité de l'approximation tout en contraignant la complexité du modèle.

Nous avons testé cette technique expérimentalement sur un grand nombre de problèmes s'inscrivant dans des espaces d'états infinis. En particulier, cette technique a donné des résultats significatifs dans un problème de contrôle d'un véhicule de type voiture dont le modèle est tiré de la robotique mobile. Ces travaux pourraient constituer un point de départ pour des applications réelles.

6.1.4. *Auto-organisation modulaire / Classification de MDP*

Participants : François Charpillet, Bruno Scherrer.

Dans le cadre où un agent artificiel aurait à résoudre un certain nombre de tâches relativement indépendantes, nous nous sommes intéressé à l'utilisation d'une approche « diviser pour régner » : nous considérons qu'un agent n'a pas un centre de décision, mais plusieurs modules spécialisés qui sont en perpétuelle compétition/coopération [27].

Nous nous sommes alors concentré sur le problème consistant à répartir, de manière complètement automatisée, n tâches sur m modules avec $m \ll n$. Nous avons montré que ce problème consiste en une classification de MDP et qu'il se formalise simplement dans le cadre de la catégorisation par noyaux. Nous avons alors proposé un algorithme itératif qui applique la technique générale des nuées dynamiques (algorithme de la littérature pour répondre au problème de la catégorisation par noyaux) à ce problème.

Nous avons testé cet algorithme sur un ensemble de tâches de navigation dans un environnement continu. Les premiers résultats expérimentaux sont encourageants : un petit nombre de modules se répartissent équitablement les tâches à résoudre.

6.1.5. *Apprentissage par renforcement et connexionnisme*

Participants : François Charpillet, Bruno Scherrer.

Une collaboration avec l'équipe Cortex nous a amené à nous intéresser à l'approche « connexionniste » pour aborder le problème de l'apprentissage par renforcement. L'approche connexionniste, en proposant de construire des systèmes composés d'un grand nombre d'unités simples et interconnectées, rend souvent possibles deux qualités intéressantes pour la mise en œuvre informatique : le parallélisme et la tolérance aux pannes.

Nous avons montré que le calcul d'un plan optimal dans un processus décisionnel de Markov admettait une réponse qui est naturellement connexionniste [12]. Pour ce faire nous avons montré que l'on pouvait voir un processus décisionnel de Markov comme une architecture connexionniste. Les unités de cette architecture sont les états. Les connexions sont le support de codage. La récompense est estimée de manière répartie sur l'ensemble des unités. Enfin la fonction de valeur optimale, qui est la représentation duale du plan, est une activité distribuée dans le réseau et l'algorithme de la littérature Value Iteration est une loi de propagation de cette activité. Nous avons alors argumenté qu'une telle vision permet d'envisager de répondre à des

problèmes d'apprentissage par renforcement d'une manière massivement parallèle (donc potentiellement rapide) et tolérante aux pannes.

6.1.6. Modèle hybride (probabiliste/analytique) pour la planification de trajectoire

Participants : Hubert Cecotti, François Charpillet, Alexis Scheuer.

En robotique mobile, la précision des mouvements prévus a un impact considérable sur la qualité des mouvements effectués. La grande difficulté consiste en particulier à prendre en compte simultanément les contraintes de déplacement du robot, l'évitement d'obstacle et les multiples imprécisions (provenant des capteurs et des actionneurs). Dans ce cadre, nous avons voulu affiner les résultats obtenus en planification de mouvement par les processus de décision markoviens (MDP), tout en évitant l'explosion combinatoire liée à une réduction du pas de discrétisation.

Pour cela, nous avons proposé un sujet de DEA consacré au couplage hiérarchique de méthodes analytiques et probabilistes, à la recherche du meilleur rapport qualité / temps de calcul [31]. Le planificateur basé sur les MDP prend en compte un modèle simplifié du mouvement, ainsi que l'évitement d'obstacle et les imprécisions. Le plan fourni par ce procédé est ensuite raffiné par des méthodes analytiques, utilisant un modèle très précis du mouvement du robot. On a ainsi obtenu en quelques secondes un mouvement d'une qualité comparable à celle d'un mouvement prévu par un MDP basé sur le modèle précis du mouvement, ce dernier mouvement nécessitant près d'une heure de calcul.

Les résultats obtenus seront mis en pratique dans le cadre d'un nouveau sujet de DEA, visant maintenant à considérer le suivi de trajectoires.

6.2. Modèles d'inspiration biologique

6.2.1. Simulation de la différenciation sociale chez les RATs

Participants : Christine Bourjot, Vincent Chevrier, Vincent Thomas.

Dans le but de mettre à jour les mécanismes fondamentaux à l'œuvre dans les phénomènes d'auto-organisation, nous avons entrepris un travail sur la simulation de l'apparition de la différenciation sociale chez le rat lors de « l'expérience de la piscine » avec le professeur D. Desor du Laboratoire de Biologie du Comportement de l'Université Henri Poincaré de Nancy. L'hypothèse biologique à valider est que la solution apportée par le groupe de rats à un problème écologique n'est pas planifiée ni contrôlée par un ensemble réduit d'individus mais émerge des nombreuses interactions entre les individus et entre les individus et leur environnement. Une modélisation basée sur le paradigme multi-agent a été développée pour simuler ce phénomène. Ce cadre de travail a permis de définir les différentes entités intervenant dans le phénomène : les rats, le milieu dans lequel ils évoluent et leurs interactions. Afin de prouver l'hypothèse biologique, les rats sont représentés par des agents réactifs et aucune référence à la cognition sociale (représentation des autres) n'est prise en compte. Les comportements individuels ont donc été définis par des réflexes de type stimulus/réponse et des règles simples de renforcement s'appuyant sur 4 variables internes. Un item comportemental est complètement défini par la donnée des stimuli déclencheurs des actions, le calcul des probabilités de réalisation des actions et par le résultat effectif de chaque action (comprenant un renforcement des probabilités). Les premières analyses statistiques des résultats de la simulation s'accordent avec succès avec les observations in vivo. Ils conduisent à la différenciation comportementale tant du point de vue des profils émergents dans le groupe que des proportions obtenues. Enfin, le modèle de simulation a été utilisé comme modèle heuristique pour aider à la formulation d'hypothèses biologiques. Plus particulièrement nous avons étudié le lien entre les contraintes environnementales et l'émergence de la structure sociale du groupe [28].

6.2.2. Détection de région bio-inspirée des araignées sociales

Participants : Christine Bourjot, Vincent Chevrier, Vincent Thomas.

La simulation de phénomènes biologiques est une source d'inspiration pour déterminer des comportements individuels simples qui produisent collectivement un phénomène complexe. Nous avons en collaboration

avec des biologistes du laboratoire de biologie et physiologie du comportement mis en évidence un modèle de comportement simulant la construction collective de toile chez les araignées sociales. Nous avons exploité les propriétés de ce mécanisme pour aborder le problème de détection de régions dans des images à niveaux de gris. Nous avons adapté l'environnement du système pour qu'il corresponde au problème à résoudre et sa solution. Enfin nous avons conservé le processus d'agrégation obtenu et sa dynamique et l'avons enrichi afin de l'adapter à un environnement nouveau. Nous avons entrepris des expérimentations [17][13] ayant pour objet de vérifier le principe de l'algorithme en l'appliquant à différentes images réelles ou provenant d'un éditeur. Nous avons obtenus plusieurs résultats qualitatifs. Le processus fournit de bons résultats lorsqu'il s'agit de déterminer une région donnée initialement, à la condition d'avoir correctement ajusté les paramètres. Cette obligation empêche encore de traiter intégralement une image et de pouvoir comparer cette approche aux approches traditionnelles. Concernant le modèle de résolution collective de problème, son originalité par rapport aux autres modèles inspirés des sociétés d'insectes est qu'il autorise, lors du choix de déplacement, la prise en compte d'informations non locales au travers des fils de soie.

6.2.3. Localisation de robots mobiles dans un environnement partiellement connu en utilisant une architecture à base de Systèmes Multi-Agents.

Participants : François Charpillat, Vincent Chevrier, Franck Gechter.

Le but de ces travaux est d'appliquer ce principe de conception de systèmes intelligents à un cadre applicatif complexe et à la réalisation d'une application concrète récurrente en robotique mobile. L'objectif fixé est de proposer une architecture permettant de localiser, suivre, identifier et contrôler une population hétérogène de robots mobiles à l'aide de percepts fournis par des capteurs dont le type, le nombre et la position ne sont pas définis a priori.

L'architecture proposée peut être décrite par les quatre points clefs suivant : son environnement, ses agents et leurs comportements, ses interactions et son organisation résultante des différentes dynamiques mises en jeu :

- L'environnement est une représentation discrète du monde réel sous la forme d'un ensemble d'états représentant chacun une position possible d'une cible dans l'environnement réel. La dynamique de l'environnement doit rendre compte des propriétés de celle du monde extérieur. Le lien entre les deux dynamiques se fait par accumulation de percepts dans les différents états.
- Le modèle comportemental des agents est inspiré du « flocking » [34]. L'ensemble des comportements a pour but à la fois une focalisation des agents sur les percepts proches et une répartition spatiale homogène dans les régions dépourvues de percepts.
- Les interactions ont lieu à la fois entre les agents, comportement de répulsion, et entre les agents et les percepts, comportement d'attraction.
- L'organisation qui émerge de cet ensemble est la focalisation des agents sur les plots formant ainsi des groupes. Cette formation de groupe matérialise la présence d'une cible sur une position du monde réel correspondant à l'état sur lequel les agents se sont focalisés.

La focalisation des agents sur les régions d'intérêt permet d'accomplir la tâche de localisation, le suivi des cibles est quant à lui le résultat de l'interaction entre les deux dynamiques qui régissent le système : la dynamique des agents et la dynamique propre de l'environnement réel.

7. Contrats industriels

7.1. Collaboration avec DIALOCA

Participants : Anne Boyer, François Charpillat, Romaric Charton, Jean Paul Haton.

La collaboration avec la société Dialoca porte sur l'utilisation d'agents intelligents dans l'élaboration de services autour de la plate-forme Unimédia développée par la société. L'objectif est que ces services s'adaptent

d'avantage aux aléas de l'environnement multimédia dans lequel ils se déroulent, tout en prenant en compte le profil des utilisateurs (leurs habitudes, leurs préférences...) et plus généralement des ressources utilisées (leur disponibilité, leur coût...). D'autre part, afin qu'ils puissent s'adapter à leurs utilisateurs, on souhaite rendre ces services capables de les accueillir et de les assister avec des conseils et des réponses personnalisées, voire une démarche adaptée. Notre travail a un objectif double : réduire la complexité de conception induite par le grand nombre de cas particuliers d'interactions d'une part, modéliser le comportement des utilisateurs pour rendre les services adaptatifs d'autre part.

Nous avons proposé une architecture à quatre niveaux (media, ressources, agents et services) permettant la collaboration d'agents hétérogènes, incluant notamment un modèle de conception de services multimédia adaptatifs où chaque agent (humain, virtuel, contrôlé ou non) est impliqué dans la réalisation d'un rôle particulier et interagit avec les autres pour réaliser le service. Nous avons défini un ensemble de classes de services permettant de modéliser les différentes applications envisagées. Le déroulement d'un service est décrit comme le suivi d'un déplacement dans un espace d'états permettant ainsi de se ramener au problème d'évitement d'obstacles et de planification de trajectoires. Pour concevoir le comportement des agents, nous avons étudié une approche par planification. En effet, les agents devront fournir des services composés, presque dynamiquement, à partir de comportements construits à partir d'une bibliothèque d'opérateurs permettant d'agir sur les média. Cette planification doit respecter des contraintes de différentes natures, depuis la sécurité et la confidentialité, aux contraintes de ressources temporelles et matérielles. D'autre part, nos études sur gestion intelligente d'un dialogue entre l'utilisateur et le service ont été appliquées afin de déterminer la meilleure stratégie d'interaction. Ce travail a été appliqué au service de réservation de billets d'avion.

8. Actions régionales, nationales et internationales

8.1. Actions régionales

8.1.1. TAM

Participants : François Charpillet, Vincent Chevrier, Franck Gechter, Laurent Jeanpierre, Bruno Scherrer.

L'action TéléAssistance Mobile vise à développer des systèmes d'entités douées d'un certain degré d'autonomie (de perception, de décision et/ou d'action) devant effectuer individuellement ou collectivement une tâche dans un environnement dynamique, bruité et incertain (monde industriel réel). Elle regroupe des équipes de recherche du CRAN, du LORIA et de Supélec (campus de Metz). Les efforts des partenaires se coordonnent autour de la thématique « Rondiers Autonomes Coopératifs » avec comme scénario démonstratif la surveillance de bâtiment. Dans ce contexte, les contributions principales de l'équipe MAIA sont :

- la proposition d'une architecture multi-agent pour la localisation et le suivi de cibles mobiles et la réalisation de primitives de perception exploitant différents types de capteurs.
- la modélisation à l'aide de Processus Décisionnel Markovien de l'environnement d'un agent ayant des objectifs multiples.

8.1.2. Télémédecine

Maintien à domicile de personnes en perte d'autonomie et prévention de la chute

Participants : Anne Boyer, François Charpillet, Franck Gechter, Loïc Pelissier.

Les personnes âgées ou handicapées sont le plus souvent désireuses de garder leur indépendance, quels que soient les efforts qu'il leur faut fournir pour se débrouiller seules à la maison. Malheureusement, faute de pouvoir garantir un niveau de confort et de sécurité suffisant, le placement dans des institutions spécialisées se révèle souvent être la seule possibilité. Le concept de télé-médecine intelligente fournit une solution à cette attente en proposant des services basés sur les technologies de l'information. Les systèmes de détection et de contrôle couplés à des systèmes de télé-médecine peuvent contribuer à rendre les maisons plus sûres et plus efficaces.

Les alternatives à l'assistance humaine à domicile peuvent se voir sous différents aspects. Il est possible d'équiper le logement de capteurs effectuant différentes mesures et on se situe alors dans le cadre de l'habitat intelligent. Ces mesures permettent de déterminer des évaluations de l'activité d'une personne. On se situe dès lors dans le cadre de l'actimétrie. Il est possible de compléter ce système par d'autres évaluations plus complexes comme la construction de modèles de comportement ou de modéliser l'état de santé d'un patient dans le but d'aider au diagnostic ou de notifier des alarmes ou des alertes. Il s'agit d'assistants intelligents.

C'est dans ce dernier cadre que se situe notre étude. Notre objectif est de réaliser une analyse comportementale de la marche fondée sur une notion de signature. A l'aide de capteurs placés judicieusement dans l'environnement, nous pouvons suivre la trajectoire d'une personne en étudiant un certain nombre de paramètres comme la vitesse, le déhanchement,... Ces différents paramètres peuvent être vus comme la signature de la marche de la personne.

L'étude a donc pour objectif de vérifier l'existence d'une signature de la marche, d'en valider ensuite la pertinence et la robustesse en milieu ouvert. Nous souhaitons déterminer si l'étude des variations de cette signature permet de déterminer des populations à risque de chutes.

Aujourd'hui, les seuls capteurs d'actimétrie non contraignants (infrarouges, sonars) disponibles montrent vite leur limite car ils ne fournissent qu'une information binaire (présence/pas de présence) dans leur rayon d'action. Cela soulève trois problèmes :

- On ne dispose pas d'informations sur la cible (détermination de son identité et activité).
- On sait s'il y a quelqu'un dans un secteur défini, mais l'on a aucune information sur sa position réelle dans cet espace (problème de précision).
- S'il y a une présence détectée, comment peut-on être sûr qu'il n'y a effectivement qu'une personne, ou alors qu'il ne s'agit pas d'un animal de compagnie (problème de différentiation) ?

Pour pallier ces limites, nous avons imaginé utiliser des caméras. La caméra est alors transformée en un capteur de présence à même de détecter une ou plusieurs personnes, de différencier ces personnes et de fournir des informations pertinentes et fiables telles que la trajectoire, la taille, la vitesse ou encore l'orientation des personnes suivies. La caméra servira à créer un détecteur de mouvements évolué capable de résoudre les problèmes de précision, de différentiation et de détermination de l'activité.

A terme, notre but est de transformer la caméra en un capteur intelligent capable de détecter des dérives de comportement de déplacement et de déclencher, le cas échéant, des alarmes et autres notifications auprès des médecins concernés.

Une première expérimentation a été menée en laboratoire. Une vingtaine de personnes a été filmée marchant dans un couloir par une seule caméra fixe. A partir de ces données ont été déterminées : la trajectoire au sol de la personne, la vitesse moyenne, vitesse instantanée et accélération, la taille et l'orientation de la personne par rapport au plan du sol. Cet ensemble de paramètres (première estimation de la signature) a donné des résultats tout à fait encourageants quant à la précision des paramètres calculés.

8.2. Actions nationales

8.2.1. ACI Temps et Cerveau

Participants : Olivier Buffet, Alain Dutech.

Dans le cadre de l'ACI « Temps et Cerveau », le projet MAIA participe un pré-projet autour des systèmes autonomes dynamiques. Cette collaboration entre l'ETIS de Cergy (équipe de P. Gaussier), l'Université de la Méditerranée (E. Daucé), l'INLN de Nice (B. Cessac), Supaéro Toulouse (M. Samuelides) et le LORIA (projet MAIA) a pour but d'étudier la notion d'émergence dans les systèmes dynamiques. Cette étude se fera sous le double aspect des réseaux de neurones artificiels dynamiques de grande taille et des systèmes multi-agents.

En particulier, nous voulons étudier l'émergence d'une modularité au sein d'un réseau de neurones artificiels de grande taille. Cette émergence sera confrontée avec l'organisation de systèmes multi-agents

développés pour des problèmes équivalents. L'application commune sur laquelle s'appuiera le projet concerne la manipulation d'un bras robotique articulé.

8.2.2. *RNTS DEPIC*

Participant : François Charpillet.

La dialyse péritonéale permet le traitement à domicile de malades insuffisants rénaux chroniques et concerne plus de 2500 patients en France et environ 120 000 dans le monde. La DPCA (Dialyse Péritonéale Continue Ambulatoire) se développe actuellement, elle s'inscrit dans les tendances modernes de médecine délocalisée et ambulatoire. Cette méthode qui présente de nombreux avantages médicaux et économiques nécessite la présence d'un cathéter intra-péritonéal, ce qui induit un risque d'infection qui, si elle n'est pas traitée à temps, peut se compliquer d'une péritonite et nécessiter une hospitalisation associée à une intervention chirurgicale et un séjour d'une dizaine de jours à l'hôpital.

L'objectif du projet DEPIC est de Concevoir et réaliser un capteur, de son instrumentation associée et d'un dispositif automatique d'alerte du risque d'infection, directement compréhensible et utilisable par le patient (signaux lumineux et sonore). Appareil portable, autonome, utilisable en milieu hospitalier, mais également à domicile, le DEPIC est prévu pour être relié au système "DIATELIC" de télésurveillance de dialyse à domicile qui analyse les données pour détecter les alertes et prévenir les aggravations. Le système prévient alors le néphrologue, le patient et le médecin généraliste traitant. Ce dispositif devrait permettre le diagnostic très précoce d'une infection, avant l'apparition des signes cliniques et des complications, et alerter le patient et le médecin qui instituera un traitement préventif.

8.2.3. *Pré-projet ROBEA intitulé COROCOP : Contrôle de robots coopératifs planétaires*

Participant : François Charpillet.

Ce projet du programme ROBEA s'inscrit dans le cadre d'une collaboration avec Abdel-Allah Mouaddib du GREYC/UMR 6072, à l'Université de Caen. Dans ce ce projet nous proposons de développer une architecture pour la planification et la supervision de l'exécution de robots mobiles d'exploration planétaire. Les résultats attendus consistent en un système autonome permettant : (1) d'augmenter la flexibilité du système global en utilisant plusieurs robots, où chacun est capable d'adapter son comportement aux changements de l'environnement et de se coordonner avec les autres robots, et (2) de développer une représentation des tâches, des plans et des stratégies de contrôle qui prennent en compte les ressources limitées et le haut niveau d'incertitude qui affecte les opérations des robots planétaires. Ces derniers, comme Sojourner, doivent opérer avec des contraintes de ressources comme l'énergie, la capacité de stockage de l'information et aussi la bande passante de communication. En plus, il existe une incertitude sur la consommation de ces ressources par les activités des robots et sur les évolutions de l'environnement.

8.3. Relations internationales

8.3.1. *Le projet IST ELIN*

Participants : Anne Boyer [coordinateur MAIA], Romaric Charton, Rédimé Hadji, Régis Lhoste.

Le projet IST-2000-30188 Elin (The electronic newspaper initiative) est un projet européen dont l'objectif principal est d'améliorer les journaux électroniques existants en introduisant les notions d'interactivité, de personnalisation avancée, de mobilité et de nomadicité. Dans le projet Elin, l'apport de MAIA consiste à spécifier et mettre en œuvre un ensemble de services web dédiés à la recherche de l'information multimédia la plus pertinente pour un consommateur, en prenant en compte ses préférences, ses habitudes, sa localisation et le contexte ambiant d'une manière générale. Notre contribution se situe à deux niveaux :

- Scientifique : nous nous sommes intéressés au filtrage collaboratif et à l'apport que les techniques de modélisation stochastique (réseaux bayésiens, modèles décisionnels de Markov) développées dans l'équipe peuvent apporter à la mise en œuvre de scénarios de push and pull personnalisés.

- Technologique : afin d'implanter un lieu de vie pour les agents intelligents que nous développons, nous utilisons une plate-forme à agents respectant les standards FIPA pour le développement de systèmes multi-agents interopérables (JADE). L'objectif est de simplifier le développement d'un ensemble d'agents et de services tout en garantissant le respect d'un standard européen. Jade est totalement neutre vis à vis de la définition d'un agent et ne traite que des capacités bas niveau permettant la mise en œuvre de communication et d'interaction entre les agents.

Comme le modèle de comportement d'un agent JADE permet l'intégration simple de code dans les tâches des agents, nous travaillons à la conception d'une bibliothèque d'algorithmes de filtres collaboratifs utilisables comme moteurs de raisonnement pour les agents. Nous concevons également une bibliothèque d'outils permettant le développement d'interactions intelligentes entre les services et l'utilisateur. Cela comprend par exemple la sélection de la modalité de dialogue la plus adaptée au contexte et à l'utilisateur.

8.3.2. *Le projet Eureka KVM*

Participants : Anne Boyer, François Charpillet, Vincent Chevrier [coordinateur MAIA].

Le projet KVM (Knowledge Valorisation Matrix) est un projet européen EURECA qui regroupe les sociétés Kappa Ice (Paris) et HTTP Software (Londres) ainsi que l'INRIA au travers de 3 projets du LORIA (MAIA, ECOO et Orpailleur). Le thème du projet concerne la valorisation de la connaissance au travers d'un outil reposant sur une cartographie particulière des connaissances d'une entreprise. Pour cela, KVM intègre un guide d'analyse, un outil de cartographie des connaissances associé à une méthodologie de représentation selon trois référentiels et un support au diagnostic stratégique et à l'aide à décision.

L'apport de MAIA à ce projet se situe principalement autour de l'exploitation stratégique des connaissances. Il concerne les techniques d'agents intelligents permettant l'aide à la décision pour le diagnostic stratégique.

8.3.3. *Le projet IST OZONE*

Participants : Anne Boyer, François Charpillet, Alain Dutech [coordinateur MAIA], Rédimé Hadji, Régis Lhoste.

Le projet OZONE (IST - 2000 - 30026) se donne pour objectif de démontrer la faisabilité de ce que l'on peut qualifier d'*Intelligence Ambiante*. Il s'agit de permettre aux utilisateurs d'être en permanence « connectés » à un réseau informatique, en utilisant pour cela la modalité (voix, image, ...) la plus adaptée. Grâce à cette connexion permanente, les logiciels s'appuyant sur OZONE doivent aussi pouvoir s'adapter et aider intelligemment les utilisateurs. Il serait ainsi possible de faire suivre le mail des gens, de gérer l'éclairage d'une maison, d'organiser des réservations dans des restaurants, de faciliter les mini télé-conférences entre amis, etc...

Ainsi, OZONE veut pouvoir offrir une bibliothèque de services à des applications qui profiteraient ainsi de son « intelligence ». Notre contribution se situe au niveau de l'apprentissage passif d'un modèle d'utilisateur, c'est-à-dire sans action active de sa part. En utilisant des méthodes d'apprentissage stochastiques, il s'agit d'apprendre les préférences d'un utilisateur en terme de modalité d'interactions et de s'appuyer sur ce modèle pour ensuite choisir la modalité la plus adaptée à une tâche donnée. Ces méthodes seront validées sur des démonstrateurs intégrant les diverses fonctionnalités d'OZONE.

8.3.4. *Le projet ITEA Protéus*

Participant : François Charpillet.

L'objectif du projet ITEA-PROTEUS est de créer un environnement intégré adapté au métier de maintenance de systèmes industriels. Cet environnement comportera une base de connaissances créée à partir de l'ensemble des informations issues des différents niveaux de maintenance, et permettra de tirer de ces données des enseignements utiles et de les rendre consultables depuis tous les niveaux d'intervention du processus de maintenance, et, notamment, depuis le lieu même où sont implantés les équipements jusqu'au bureau du gestionnaire de service.

Ce projet réunit 12 partenaires dans deux pays européens : la France et l'Allemagne. Ce consortium est conduit par CEGELEC, une compagnie Française, leader européen de la maintenance et des services.

L'information stockée dans les bases est issue des niveaux les plus bas du système de maintenance. Il s'agit d'informations telles que informations mesure venant des capteurs, informations constructeurs à partir des documents attachés aux équipements, des informations contractuelles recueillies dans les contrats et notamment des engagements pris en ce qui concernent les résultats à atteindre, des informations du service qualité de l'entreprise définissant les règles de l'art en la matière et enfin les informations recueillies par les GMAO installées sur les sites en exploitation relatant le comportement des équipements dans leur contexte. PROTEUS prévoit l'accès aisé à ces informations via des méthodes multimédia de type vidéo interactive, par des représentations 3D ou par le biais de la « réalité augmentée » utilisée par les intervenants sur site . L'accès classique à ces informations, organisées par un logiciel approprié, sera facilitée par des moteurs de recherches multi-critères ou en plein texte via une interface homme/machine soignée.

La richesse du contenu de la base sera ensuite augmentée des résultats obtenus par son interprétation qui exploitera les méthodes d'intelligence artificielle, le but étant d'améliorer les soutiens logistiques à tous les niveaux du processus de maintenance avec un souci permanent d'améliorer l'aspect prédictif de la maintenance ainsi qu'un support au curatif pour un maintien en condition opérationnel optimum.

8.4. Relations bilatérales internationales

8.4.1. *Projet avec Umass*

Participants : François Charpillat, Anne Boyer.

Un projet NSF-INRIA avec l'équipe « *Resource Bounded Reasoning Research Group* » de l'université de Massachusetts a démarré en mars 1997, pour une durée de 3 ans. Un renouvellement pour une durée de deux ans a été ensuite obtenu. Nous travaillons sur la construction, la composition et le contrôle d'algorithmes *anytime*[14].

8.4.2. *Projet avec la NASA : « self-Directed Cooperative Planetary Rovers (Cross Enterprise Space Technology Development Program de la NASA) ».*

Participant : François Charpillat.

Avec :

Prof. Shlomo Zilberstein, *University of Massachusetts*,

Prof. Victor Lesser, *University of Massachusetts*,

Prof. Eric Hansen, *Mississippi State University*,

Dr. Richard Washington, *NASA Ames Research Center*,

Dr. Abdel-illah Mouaddib CRIL (Lens).

Les robots d'exploration planétaires "Planetary rovers" sont des véhicules sans pilote commandés à distance. Ils sont équipés de caméras, de capteurs variés et d'effecteurs. Ces capteurs et effecteurs sont destinés à la réalisation d'expériences scientifiques planifiées depuis la terre. Ces robots ont démontrés leur efficacité et devraient jouer un rôle important dans les prochaines missions d'exploration planétaire de la NASA, en particulier sur Mars. Cependant les missions récentes ont révélées la faiblesse des systèmes existants qui sont complètement " téléguidés " depuis la terre. Ces systèmes comme " Sojourner " n'ont que peu d'autonomie. Il se contente d'enchaîner les ordres qui leur sont envoyés depuis la terre. L'incertitude inhérente à l'exploitation d'un environnement mal connu, la faible bande passante du réseau de communication et le temps nécessaire à la transmission de l'information entre la terre et le lieu de la mission (plusieurs minutes sur Mars) augmente le risque de disfonctionnement et d'immobilisation du robot. Ce projet a pour objectif de développer les techniques permettant d'augmenter l'autonomie des robots d'exploration planétaire. Il s'agit de gérer au mieux les ressources disponibles afin de mieux prendre en compte les difficultés soulevées ci-dessus. L'idée est de développer des stratégies de contrôle pré-établies afin de permettre au robot de répondre rapidement aux questions suivantes : étant donné un ensemble d'alternatives pour réaliser une tâche donnée laquelle

choisir compte tenu des circonstances actuelles, comment évaluer la qualité des tâches réalisées, comment réagir face à un dysfonctionnement, faut-il recommencer une expérience dont les résultats sont de qualité modeste, comment allouer les ressources pour réaliser au mieux la mission et ce dans une fenêtre temporelle donnée. Pour atteindre ses objectifs le projet développe des solutions fondées sur la théorie de la décision et plus particulièrement sur les modèles décisionnels de Markov. Il s'agit de développer des méthodes robustes à l'incertitude inhérente à l'activité d'un robot que ce soit en terme des ressources consommées ou des effets des actions entreprises. Les techniques d'apprentissage par renforcement permettent de développer de tels modèles et d'en tirer des stratégies d'exécution robustes. Un autre aspect du projet est d'évaluer la faisabilité de systèmes d'exploration planétaire composés non plus d'un seul robot mais d'un ensemble de robots éventuellement plus simples. Il s'agit entre autre d'évaluer d'une telle approche sur la complexité, l'autonomie, la robustesse de tels systèmes.

8.5. Participation à des colloques, séminaires, invitations

- Participation à des jurys de thèses de doctorat (J.-P. Haton et F. Charpillet) ;
- On se reportera à la bibliographie pour la liste des conférences et *workshops* auxquels les membres du projet ont participé.

10. Bibliographie

Bibliographie de référence

- [1] M. BOUZID, V. CHEVRIER, S. VIALLE, F. CHARPILLET. *Parallel simulation of a stochastic agent/environment interaction model*. in « Integrated Computer-Aided Engineering », numéro 3, volume 8, juillet, 2001, pages 189 - 203.
- [2] A. DURY, F. LE BER, V. CHEVRIER. *A reactive approach for solving constraint satisfaction problems : assigning land use to farming territories*. in « Agent Theories, Architectures, and Languages - ATAL'98 », 1998.
- [3] A. DUTECH, O. BUFFET, F. CHARPILLET. *Multi-Agent Systems by Incremental Gradient Reinforcement Learning*. in « 17th International Joint Conference on Artificial Intelligence, IJCAI-01, Seattle, WA, USA », volume 2, pages 833-838, août, 2001.
- [4] A. DUTECH. *Solving POMDPs using selected past events*. in « European Conference on Artificial Intelligence, Berlin », août, 2000.
- [5] P. LAROCHE. *GraphMDP : A New Decomposition Tool for Solving Markov Decision Processes*. in « International Journal on Artificial Intelligence Tools », numéro 3, volume 10, 2001, pages 325-344.
- [6] M. MOUHOUB, F. CHARPILLET, J.-P. HATON. *Experimental Analysis of Numeric and Symbolic Constraint Satisfaction Techniques for Temporal Reasoning*. in « Constraints : An International Journal », numéro 2-3, volume 3, juin, 1998, pages 151-164.
- [7] S. ZILBERSTEIN, F. CHARPILLET, P. CHASSAING. *Real-Time Problem-Solving with Contract Algorithms*. in « 16th International Joint Conference on Artificial Intelligence, Stockholm », volume 2, pages 1008-1013, août, 1999.

Thèses et habilitations à diriger des recherche

- [8] D. BELLOT. *Fusion de données avec des réseaux bayésiens pour la modélisation des systèmes dynamiques et son application en télémédecine*. Thèse d'université, UHP Nancy, novembre, 2002.
- [9] I. CHADÈS. *Planification distribuée dans les systèmes multi-agents à l'aide de processus décisionnels de Markov*. Thèse d'université, UHP Nancy - 1, janvier, 2003.
- [10] V. CHEVRIER. *Contributions au domaine des systèmes multi-agents*. thèse de doctorat, Université Henri Poincaré Nancy 1, janvier, 2002, Habilitation à Diriger des Recherches.
- [11] L. JEANPIERRE. *Apprentissage et adaptation pour la modélisation stochastique de systèmes dynamiques réels*. Thèse d'université, UHP Nancy - 1, décembre, 2002.
- [12] B. SCHERRER. *Apprentissage de représentation et auto-organisation modulaire pour un agent autonome*. Thèse d'université, UHP Nancy - 1, janvier, 2003.

Articles et chapitres de livre

- [13] C. BOURJOT, V. CHEVRIER, V. THOMAS. *A new swarm mechanism based on social spiders colonies : from web weaving to region detection*. in « Web Intelligence and Agent Systems : An International Journal - WIAS », mars, 2003, à paraître.
- [14] S. ZILBERSTEIN, F. CHARPILLET, P. CHASSAING. *Optimal Sequencing of Contract Algorithms*. in « in Annals of Mathematics and Artificial Intelligence », 2003, à paraître.

Communications à des congrès, colloques, etc.

- [15] D. BELLOT, A. BOYER, F. CHARPILLET. *A new definition of qualified gain in a data fusion*. in « The Fifth International Conference on Information Fusion - FUSION'2002, Annapolis, Maryland, USA », juillet, 2002.
- [16] D. BELLOT, A. BOYER, F. CHARPILLET. *Designing smart agent based telemedicine systems using dynamic bayesian networks : an application to kidney disease people*. in « 4th International Workshop on Enterprise Networking and Computing in Health Care Industry - HEALTHCOM 2002, Nancy, France », juin, 2002.
- [17] C. BOURJOT, V. CHEVRIER, V. THOMAS. *How Social Spider Inspired An Approach To Region Detection*. in « International Conference on Autonomous Agents and MultiAgent Systems - AAMAS'02, Bologna, Italy », juillet, 2002.
- [18] O. BUFFET, A. DUTECH, F. CHARPILLET. *Adaptive Combination of Behaviors in an Agent*. in « European Conference on Artificial Intelligence - ECAI'02, Lyon, France », pages 48-52, juillet, 2002.
- [19] O. BUFFET, A. DUTECH, F. CHARPILLET. *Learning to weigh basic behaviors in Scalable Agents*. in « International Conference on Autonomous Agents and MultiAgent Systems - AAMAS'02, Bologna, Italy », volume 3, pages 1264-1265, juillet, 2002, [poster].
- [20] I. CHADÈS, B. SCHERRER, F. CHARPILLET. *A Heuristic Approach for Solving Decentralized-POMDP* :

Assessment on the Pursuit Problem. in « ACM Symposium on Applied Computing - SAC'2002, Madrid, Spain », mars, 2002.

- [21] R. CHARTON, A. BOYER, F. CHARPILLET. *Providing users with adapted services : Dynamic building of dialogues to make heterogeneous agents cooperate.* in « 2nd IEEE International Symposium on Signal Processing and Information Technology - ISSPIT 2002, Marrakech, Maroc », décembre, 2002.
- [22] R. CHARTON, A. BOYER, F. CHARPILLET. *Towards bringing heterogeneous agents to cooperation : an architecture for multimedia services.* in « First Joint Conference on Autonomous Agents & Multiagent Systems - AAMAS'2002, Palazzo Re Enzo, Bologna, Italy », volume 3, ACM, ACM Press, éditeurs C. C. ET W. LEWIS JOHNSON., pages 1415-1416, juillet, 2002.
- [23] L. JEANPIERRE, F. CHARPILLET. *Apprentissage de modèles en télémédecine.* in « Conférence Francophone d'Apprentissage - CAP 2002, Orléans, France », Matthieu Exbrayat, juin, 2002.
- [24] L. JEANPIERRE, F. CHARPILLET. *Hidden Markov Models for Medical Diagnosis.* in « 4th International Workshop on Enterprise Networking and Computing in Health Care Industry - HEALTHCOM 2002, Nancy, France », juin, 2002.
- [25] B. SCHERRER, F. CHARPILLET. *Coevolutive Planning In Markov Decision Processes.* in « First International Joint Conference on Autonomous Agents & Multiagent Systems - AAMAS 2002, Palazzo Re Enzo, Bologna, Italy », juillet, 2002.
- [26] B. SCHERRER, F. CHARPILLET. *Cooperative Co-learning : A Model-based Approach for Solving Multi Agent Reinforcement Problems.* in « 14th IEEE International Conference on Tools with Artificial Intelligence - ICTAI 2002, Washington, USA », IEEE, novembre, 2002.
- [27] B. SCHERRER. *A connectionist architecture that adapts its representation to complex tasks.* in « International Joint Conference on Neural Networks - IJCNN 2002, Hilton hawaïian Village, Honolulu, HI », mai, 2002.
- [28] V. THOMAS, C. BOURJOT, V. CHEVRIER, D. DESOR. *MAS and RATS : Multi-agent simulation of social differentiation in rats* groups.* . in « International Workshop on Self-Organization and Evolution of Social Behaviour, Monte Verita, Ascona, Switzerland », septembre, 2002.

Rapports de recherche et publications internes

- [29] D. BELLOT, A. BOYER, F. CHARPILLET. *Vers une approche formelle de la fusion de données en intelligence artificielle : application en télémédecine.* Rapport de recherche, janvier, 2002.
- [30] O. BUFFET. *Apprentissage par Renforcement pour la Conception de Systèmes Multi-Agents.* Rapport d'avancement de thèse, juin, 2002, Un comité de thèse a été réuni pour faire le bilan des deux premières années de travail. Ce rapport est un document préparatoire à cette réunion..
- [31] H. CECCOTTI. *Planification de chemin utilisant des méthodes analytiques et stochastiques.* Stage de DEA, juillet, 2002.

- [32] R. CHARTON. *Agent approach for Dialoca Services, Architecture and user-mediator interaction*. Rapport technique, avril, 2002.

Bibliographie générale

- [33] J. BAXTER, P. BARTLETT. *Infinite-Horizon Policy-Gradient Estimation*. in « Journal of Artificial Intelligence Research », volume 15, 2001, pages 319-350.
- [34] C. W. REYNOLDS. *Flocks, Herds, and Schools : A Distributed Behavioral Model*. in « Computer Graphics », numéro 4, volume 21, 1987, pages 25-34.
- [35] R. SUTTON, G. BARTO. *Reinforcement Learning : an introduction*. Bradford Book, MIT Press, Cambridge, MA, 1998.
- [36] C. WATKINS. *Learning from delayed rewards..* thèse de doctorat, King's College of Cambridge, UK., 1989.