

Rapport INRIA 1994 — Programme 3
Modèles fondamentaux et applications des
processus perceptifs et cognitifs

PROJET SYCO

3 mai 1995

PROJET SYCO

Modèles fondamentaux et applications des processus perceptifs et cognitifs

Localisation : *Nancy*

Mots-clés : acquisition de connaissances (1), compréhension de texte (1), connexionnisme (1), intelligence artificielle (1), interprétation d'image (1), interprétation de signal (1), raisonnement (1), reconnaissance de la parole (1), reconnaissance des formes (1), représentation de connaissances (1), réseau neuro-mimétique (1), système expert (1), système multi-agents (1), temps réel (1).

1 Composition de l'équipe

Responsable Scientifique

Jean-Paul Haton, Professeur, U. Henri Poincaré, Nancy I

Secrétariat

Martine Kuhlmann, CNRS
Isabelle Herlich, INRIA (mi-temps)

Personnel INRIA

Frédéric Alexandre, Chargé de Recherche
François Charpillet, Chargé de Recherche

Personnel CNRS

Anne Bonneau, Chargée de Recherche
Dominique Fohr, Chargé de Recherche

Yifan Gong, Chargé de Recherche
Amedeo Napoli, Chargé de Recherche

Personnel Université

Pascal Blanchet, Maître de Conférences, U. Nancy II
Christine Bourjot, Maître de Conférences, U. Nancy II
Anne Boyer, Maître de Conférences, U. Nancy II
Vincent Chevrier, Maître de Conférences, U. H. Poincaré,
Nancy I
Sylvie Coste-Marquis, ATER, U. Nancy II
Marie-Christine Haton, Professeur, U. H. Poincaré, Nancy I
Corinne Loesel, ATER, U. H. Poincaré, Nancy I
Jean-François Mari, Maître de Conférences, U. Nancy II
Pierre Marquis, Maître de Conférences, U. Nancy II
Odile Mella, Assistante, U. H. Poincaré, Nancy I
Abdelillah Mouaddib, ATER, U. de Lille
Kamel Smaïli, Maître de Conférences, U. Nancy II

Chercheurs doctorants

Iskander Ayari, MESR
Olivier Aycard, MESR
Lionel Beaugé, MESR
Maroua Bouzid, CIES
Nadjet Bouzid, CIES
Brigitte Colnet, DGA
Jean-Claude Di Martino, DGA
Linda Djeddar, CIES
Stéphane Durand, MESR
Rémy Foisel, MESR
Jean-Michel Gallone, MESR
Norbert Glaser, HCM
Irina Illina, CIES
Nicolas Lachiche, BDI
Yannick Lallement, MESR
Jean Lieber, BDI, Servier
Malek Mouhoub, CIES
Fériel Mouria, CIES
Roselyne Nguyen, MESR
Nicolas Pican, INRIA
Jean-François Remm, BDI, LCTAR

Olivier Siohan, BDI

Ingénieurs sur contrat

Marie-Pierre Chouvet, SOLLAC (CNRS)

Florent Fauchier, RATP (UHP)

Chercheurs invités

Jean-Marie Dirand, Université de Sherbrooke, Canada (jusqu'au 1/10/94)

Hai-Shou Li, Hong Kong University (à partir du 1/11/94)

Richard Washington, Stanford University (à partir du 1/4/94)

Collaborateurs extérieurs

Jean-Claude Junqua, STL, Santa-Barbara, USA

Pascal Lallemand, Praticien-Conseil

Florence Le Ber, INRA-Nancy

Remarques : Projet commun à l'INRIA-Lorraine et au Centre de Recherche en Informatique de Nancy (CRIN), laboratoire des universités de Nancy I, Nancy II et de l'INPL, unité associée au CNRS (UA 262).

2 Présentation du projet

Le projet SYCO est organisé autour d'un thème scientifique unique : l'étude de l'intelligence et de ses processus perceptifs et cognitifs, sur le plan des modèles fondamentaux et des applications en vraie grandeur. Ce thème à long terme se conjugue sous des formes très variées, compte tenu de l'ampleur du problème et de sa diversité, avec une concertation et une complémentarité fortes entre les différentes approches.

Un cadre fédérateur de l'ensemble des travaux menés au sein de ce projet est le concept de systèmes à bases de connaissances, avec la dualité acquisition et représentation des connaissances d'une part, et raisonnement d'autre part. On peut également noter une composante cognitive, indissociable des thèmes abordés, sur les plans de la perception (auditive et visuelle, en lien avec la reconnaissance des formes), de l'ergonomie et des neurosciences (avec les modèles connexionnistes). Plus généralement, et comme le plus souvent en Intelligence Artificielle (IA), les recherches sont par essence pluridisciplinaires. L'informatique en constitue le noyau central mais l'apport de diverses disciplines est tout aussi

important, notamment : logique, linguistique, phonétique, traitement du signal, neurosciences, psychologie cognitive.

Outre ses **aspects cognitifs** déjà mentionnés, la problématique du projet tourne autour d'un ensemble de notions clés. La première est celle d'**interprétation de formes complexes**, physiques (parole, signaux industriels, documents, images) ou abstraites (situations, états). Cette interprétation est menée en vue d'une certaine action à effectuer et nécessite une compréhension des formes fournies, fondée sur une exploitation raisonnée de connaissances, parmi lesquelles les différents types de modèles de description (et leur apprentissage automatique) jouent un rôle primordial. Le raisonnement correspondant nécessite de faire coopérer différents niveaux conceptuels de traitement et donc de concevoir des architectures multi-bases de connaissances adaptées.

L'indéterminisme inhérent aux problèmes abordés, la non-monotonie et le rôle joué par le temps nécessitent de concevoir des systèmes élaborés de gestion d'**hypothèses**. Il s'agit là d'un autre concept central de nos travaux, impliquant des aspects variés : représentation et manipulation d'hypothèses, maintien de cohérence, abduction, etc.

Une autre caractéristique d'une bonne partie des travaux du projet concerne la nécessité de **fonctionner en temps réel**. Ceci pose les problèmes relatifs à l'organisation des programmes et aux architectures matérielles, classiques en informatique, mais aussi des questions spécifiques liées à la conjonction IA et temps réel : raisonnement interruptible, adaptation d'un système à bases de connaissances à un temps contraint, etc.

La réflexion fondamentale et la formalisation des concepts vont de pair avec l'expérimentation pratique conduisant à la réalisation de maquettes opérationnelles avec différents partenaires extérieurs, dans des domaines d'application avancés. Cette volonté de maintenir un équilibre entre la théorie et la pratique a toujours été présente dans le projet. Elle nous paraît être un facteur indispensable à la progression des recherches en IA.

Les points marquants concernent cette année :

- la reconnaissance automatique de la parole :
 - extension du modèle stochastique de trajectoire (STM), modèle original constituant une alternative aux modèles de Markov HMM,

- mise au point de modèles statistiques performants issus de STM pour la reconnaissance de la parole bruitée,
- complétion des tests sur notre modèle de Markov du second ordre,
- extension du modèle de reconnaissance fondé sur une approche déductive-abductive. Ce modèle initialement appliqué au décodage phonétique a été étendu au niveau lexical,
- développement de modèles hybrides combinant les aspects stochastiques et neuromimétiques pour la reconnaissance de vocabulaires difficiles.
- les modèles connexionnistes :
 - phase de mise au point de l'environnement distribué de définition de modèles hybrides dans le cadre du projet ESPRIT MIX,
 - développement et premiers tests du modèle neuromimétique de cartes spatio-temporelles permettant l'apprentissage de séquences,
 - utilisation de notre algorithme OWE pour des problèmes de neuro-contrôle dans la sidérurgie,
 - développement d'outils neuromimétiques pour le traitement de signaux divers (sonar, radar, parole),
 - mise au point d'un environnement de programmation pour la définition de neurones formels complexes (unités sigma-pi, triades synaptiques)
- les systèmes à base de connaissances :
 - définition de modèles d'agents, étude des mécanismes d'interaction entre agents (négociation, etc.),
 - méthodologie d'analyse et d'acquisition de connaissances pour les systèmes multi-agents,
 - mise au point de la boîte à outils REAKT pour le développement de systèmes temps-réel,
 - développement de DBB, environnement temps-réel distribué,
 - modèles de fusion de données, de planification et de raisonnement temporel, avec application au robot mobile acquis cette année par le projet,

- définition d'un modèle de raisonnement à partir de cas, avec application à la synthèse en chimie organique.

3 Actions de recherche

3.1 Représentation des connaissances et raisonnement

Participants : Iskander Ayari, Olivier Aycard, Maroua Bouzid, Nadjat Bouzid, Anne Boyer, François Charpillet, Vincent Chevrier, Marie-Pierre Chouvet, Sylvie Coste-Marquis, Rémy Foisel, Jean-Michel Gallone, Jean-Paul Haton, Marie-Christine Haton, Nicolas Lachiche, Jean Lieber, Corinne Loesel, Pierre Marquis, Abdelillah Mouaddib, Malek Mouhoub, Amedeo Napoli, Roselyne Nguyen, Kamel Smaïli

3.1.1 Systèmes multi-agents

Participants : Iskander Ayari, Olivier Aycard, François Charpillet, Vincent Chevrier, Marie-Pierre Chouvet, Rémy Foisel, Jean-Michel Gallone, Jean-Paul Haton, Marie-Christine Haton, Abdelillah Mouaddib, Roselyne Nguyen, Kamel Smaïli

Nous nous intéressons à la résolution de problèmes complexes nécessitant la coopération ou la mise en concurrence de plusieurs “agents” (experts, bases de connaissances, procédures). Différents modèles d'agents ont été développés ces dernières années dans le projet et ont été validés par la réalisation d'applications industrielles de grande taille.

Depuis 1986, date de création du groupe ATOME, nous avons développé et étudié plusieurs modèles multi-agents. ATOME a cependant constitué la plateforme d'expérimentation la plus riche dans le sens où un outil pré-industriel est disponible en langage C++ ou Common-Lisp et installé sur près de trente sites à travers le monde.

Les travaux effectués cette année ont plus spécifiquement porté sur l'enrichissement de la boîte à outils GTMAS [26, 25] par l'adjonction de primitives permettant la modélisation et l'implantation d'agents réactifs et de leur environnement, ainsi que sur l'étude pluri-disciplinaire des notions de communication et d'interaction dans les systèmes multi-agents à l'éclairage de celles issues de la psycho-sociologie; ce travail a amené une redéfinition de la notion d'interaction.

Nous avons abordé l'acquisition et la modélisation de connaissances pour une société d'agents. Notre but est de développer des méthodes et des formalismes pour la modélisation, la représentation et l'acquisition de connaissances. Nous parlons d'une approche guidée par des modèles d'interprétation. La représentation est basée sur un formalisme qui permet une réutilisation et nous proposons des bibliothèques de modèles. L'acquisition est organisée en deux cycles qui construisent les modèles à partir de ces bibliothèques. Dans un environnement dynamique, l'acquisition de connaissances peut souvent ne pas être complète : les connaissances doivent être adaptées à de nouvelles situations. Nous nous intéressons particulièrement à la restructuration de connaissances nécessaires pour adapter le système et ses agents.

Les aspects temps réel des systèmes multi-agents ont également été abordés sous deux aspects :

- étude des fonctionnalités nécessaires à un agent pour un raisonnement en temps réel dans le cadre du projet ESPRIT REAKT (*cf.* § 3.1.4) et son application à la conduite de procédés industriels, d'une part, et à la navigation de notre robot mobile Gaston d'autre part (*cf.* § 3.1.5);
- conception d'un outil de développement de systèmes multi-agents temps réel distribués, à la suite des travaux de l'équipe sur ATOME-TR (projet DBB avec Ilog et Matra)

Enfin, une extension actuelle de nos recherches concerne l'étude de systèmes d'agents hétérogènes, avec coopération entre agents de conception classique et agents de conception neuro-mimétiques dans le cadre du projet ESPRIT BRA MIX (*cf.* § 3.2.2) ;

Sur le plan des applications, trois actions sont à noter :

- fusion de données et d'informations issues de capteurs différents. Nous concevons un système de fusion à deux niveaux: fusion numérique, détermination de situation pour un robot mobile autonome. Le système de fusion est formé de plusieurs ensembles d'agents répartis aux différents niveaux d'abstraction et communiquant entre eux. Ces agents ont pour objectif l'interprétation d'un modèle global des obstacles présents dans l'environnement du robot, et la représentation de la situation actuelle. Ce modèle constitue l'information utile que peut utiliser un système de planification de trajectoire ou de localisation ;

- en collaboration avec l'INRA (centre de Nancy), modélisation et représentation à l'aide des techniques multi-agents de l'expertise de chercheurs en matière de diagnostic et de simulation de systèmes agricoles dans le territoire d'un village. Cette expertise s'exprime à partir d'images satellitaires et de cartes. Deux maquettes ont été réalisées en utilisant l'architecture d'Atome [18]. Actuellement le travail porte sur la modélisation des connaissances profondes par une ontologie et des modèles du domaine. Un modèle d'agent est en cours d'implantation pour représenter ces connaissances ;
- dans le cadre du projet SACHEM (SOLLAC), réalisation de plusieurs maquettes en vue de valider puis de spécifier des interfaces homme-machine. SACHEM est un projet très important dont le but est de construire un système d'aide à la conduite de hauts-fourneaux. Nous avons, dans un premier temps, montré en quoi le générateur de systèmes multi-agents ATOME pouvait être une aide à la validation d'un modèle de la coopération homme-machine. ATOME propose en effet une structure comparable à celle d'un modèle conceptuel développé d'après la méthode KADS. Dans un deuxième temps, nous avons étendu l'architecture initiale afin de réaliser des spécifications de points précis des interfaces. Ceci a donné lieu à l'implantation de nouveaux types de communications dans ATOME.

Ces travaux seront poursuivis selon les axes complémentaires suivants :

- proposition d'une architecture multi-agent pour un système de gestion de dialogue multi-modal ;
- étude des fonctionnalités nécessaires pour la définition d'un système automatique de production construit sur une architecture multi-agent ;
- proposition et implantation de systèmes multi-agents parallèles dans le cadre du Centre Charles Hermite ;
- définition de critères discriminants pour décrire des modèles d'agents et de société d'agents, spécifications et validation par une étude comparative réalisée avec GTMAS.

3.1.2 Maintien de vérité

Participants : Maroua Bouzid, Nadjat Bouzid, François Charpillet, Sylvie Coste-Marquis, Jean-Paul Haton, Nicolas Lachiche, Corinne Loesel, Pierre Marquis

Nous travaillons à étendre les systèmes de maintien de vérité fondé sur les hypothèses (ATMS) à des langages de représentation plus expressifs que ceux habituellement utilisés. Pour cela, nous avons établi les bases formelles d'un ATMS standard et démontré comment elles pouvaient être étendues au delà du cadre propositionnel classique. Ainsi la notion de minimalité en terme d'inclusion ensembliste a été reformulée en terme de conséquence logique minimale d'un ensemble d'hypothèses. Cette base formelle a permis de réviser et d'étendre les extensions temporelles et qualitatives développées dans le projet ces dernières années. Il s'agit tout d'abord du système ATMS-WLT (Assumption-based Truth Maintenance System in a Window-based Logic) qui repose sur les notions d'intervalles et de fenêtres temporelles. La seconde extension permet de modéliser et d'implanter le processus de simulation qualitative. Ce modèle s'appuie sur un ATMS qui a été étendu de façon à prendre en compte des données ayant une composante qualitative et temporelle. La composante qualitative évolue dans un espace des quantités fondé sur les ordres de grandeur. La composante temporelle repose sur une représentation du temps sous forme d'intervalles numériques.

Nous avons poursuivi cette année le travail mené sur l'abduction et le maintien de vérité, en particulier dans un cadre temporel [22, 21], et pour la reconnaissance automatique de la parole [2, 3, 27] (cf. paragraphe 3.3.1). Ces derniers travaux s'intéressent aux techniques du raisonnement hypothétique pour émettre des hypothèses (par exemple, au niveau du décodage acoustico-phonétique, hypothèses de sons explicatives des indices observés), gérer les différentes hypothèses aux différents niveaux de la reconnaissance (interprétation des sons puis des mots), et réviser les données acoustiques quand une incohérence entre indices est détectée. Nos activités ont également porté cette année sur la mise à jour de bases de connaissances. Nous avons proposé un critère d'indépendance pour la mise à jour dans le cadre de l'approche des modèles possibles (PMA) de Winslett [51]. Dans un cadre propositionnel, ce critère permet de décider si une formule est indépendante d'une action sans

avoir à considérer les modèles (ou une forme normale disjonctive) de la formule ou de l'action.

3.1.3 Raisonnement temporel

Participants: Maroua Bouzid, François Charpillet, Jean-Paul Haton, Pierre Marquis, Malek Mouhoub

Nos travaux sur le raisonnement temporel portent sur la définition de modèles de représentation des informations temporelles. Pour exploiter ces modèles, nous étudions deux types d'approches : l'une fondée sur les techniques de propagation de contraintes, l'autre fondée sur les méthodes de démonstration automatique :

- L'approche par propagation de contraintes a été au centre de nombreux travaux qui se partagent en deux écoles selon l'ontologie de référence : instant ou intervalle. Notre démarche s'insère dans la seconde catégorie. Nous étudions une extension de l'algèbre des intervalles d'Allen de telle sorte que des informations numériques sous forme de fenêtres temporelles et des informations symboliques sous forme de relations puissent être traitées simultanément. Ce modèle est fondé sur une représentation discrète du temps dans laquelle est présente une notion de granularité qui permet la représentation de phénomènes dans lesquels différents processus évoluent avec des constantes de temps disparates. Nous avons implanté une nouvelle méthode de propagation de contraintes qui consiste à séparer la propagation des deux types de contraintes, numériques et symboliques, en deux niveaux de traitement qui interagissent jusqu'à stabilité. Un outil, TemPro, a été construit sur cette base. Il peut être utilisé pour des tâches telles que l'ordonancement, l'allocation de ressources, etc. Nous avons également étudié la problématique du couplage de TemPro avec un moteur d'inférence tel que G2.
- Nous proposons dans [66] une approche fondée sur une logique temporelle réifiée restreinte à des propositions temporelles non réflexives temporellement qualifiée sur des unions d'intervalles convexes. Nous avons également démontré qu'une logique temporelle réifiée est plus expressive qu'une logique temporelle équivalente non réifiée contrairement à ce qui avait été argumenté par

Bacchus et al. Une implantation de cette logique est en cours de réalisation.

Dans [67], nous proposons une autre logique réifiée dans laquelle le temps est représenté explicitement à l'aide des *fonctions caractéristiques*, c'est-à-dire des fonctions logiques caractérisant les propriétés d'un système dans le temps.

3.1.4 Raisonnement temps réel

Participants: Anne Boyer, François Charpillet, Jean-Michel Gallone, Jean-Paul Haton, Abdelillah Mouaddib

La complexité des systèmes temps réel a tendance à s'accroître en ce qui concerne le nombre de fonctions qu'ils assurent, le débit d'information à traiter et le nombre de facteurs à prendre en compte dans la décision. Cet accroissement de complexité rend très intéressant, parfois indispensable, le recours aux techniques d'intelligence artificielle. Le fonctionnement en temps réel d'un système à bases de connaissances est donc un enjeu important.

Cependant, s'il est clair que l'IA a été utilisée avec succès dans de nombreuses disciplines, elle n'est pas encore capable d'apporter d'emblée une solution aux applications temps réel. L'origine de cette déficience est à la fois historique (les premiers travaux concernant la problématique temps réel sont relativement récents en IA) et inhérente aux problèmes abordés par l'intelligence artificielle. Ceux-ci sont généralement de nature difficile (le plus souvent NP, c'est-à-dire non déterministe) et les recherches en IA se sont focalisées sur le développement d'algorithmes complexes dont les performances sont le plus souvent imprévisibles car on ne connaît pas a priori quelle portion de l'espace de recherche il faudra parcourir pour trouver une solution. A contrario, un système temps réel requiert une gestion fine des ressources nécessitant une bonne connaissance du comportement futur du système. Pour aborder cet antagonisme entre l'IA et le temps réel, nous focalisons nos travaux de recherche selon deux axes complémentaires : la conception d'environnements de développement et le raisonnement en temps contraint.

Conception d'environnements de développement. Depuis quelques années, nos travaux dans le domaine du temps réel nous ont

conduits à étudier et intégrer dans des plates-formes d'évaluation différents modèles issus à la fois de l'intelligence artificielle et de l'informatique temps réel traditionnelle. Les modèles multi-agents et plus particulièrement ceux fondés sur le paradigme du tableau noir, ont été à la base de plusieurs implantations pratiques :

- RT-SOS, extension temps réel du modèle de la Société de Spécialistes,
- ATOME-TR développé avec MATRA dans le cadre d'un contrat DRET à partir de notre outil ATOME ,
- REAKT, un environnement de développement de systèmes à bases de connaissances temps réel (Projet Esprit III REAKT II 7805),
- DBB (Distributed Blackboard), évolution d'ATOME-TR vers une architecture distribuée, en collaboration avec MATRA et ILOG dans le cadre d'un projet financé par le Ministère de l'Industrie et du Commerce Extérieur.

Raisonnement en temps contraint. La garantie de temps de réponse est une des caractéristiques essentielles d'un système temps réel. Ce problème a été abordé d'une façon nouvelle en intelligence artificielle. Les chercheurs de ce domaine ont dégagé un certain nombre de paradigmes comme les méthodes d'affinements successifs ou les méthodes d'approximations multiples. Toutes reposent sur un compromis entre le temps de réponse d'un traitement et sa qualité. Dans ce cadre, nous avons développé deux modèles de raisonnement progressifs : GREAT (Guaranteed REASONING Time) et REAKT (REAL time Knowledge Tool). Tous deux reposent sur une résolution progressive d'un problème de façon qu'une première solution soit disponible rapidement. Ensuite celle-ci est affinée par une succession d'étapes qui produisent chacune un résultat de qualité croissante. Ainsi, si le traitement est interrompu, le résultat de l'étape précédente peut être utilisé.

GREAT est un modèle [53, 54] qui s'appuie sur un langage de représentation à base de règles de production manipulant des concepts du type (attribut, objet, valeur, granularité, précision). La granularité représente une mesure du niveau de détail d'un attribut donné qui évolue en fonction du contexte de travail. Elle permet ainsi de classer les connaissances en fonction du niveau de détail que l'on souhaite obtenir grâce à une organisation des données en régions et de la base de connaissances en paquets de règles. A chaque région correspond un paquet de règles dont

l'exécution est contrôlée par un filtre et un évaluateur. Le filtre sélectionne progressivement les données du paquet en fonction de leur degré de précision de façon que les règles s'appuyant sur les données les plus précises soient prioritaires. Après chaque déclenchement de règle, l'évaluateur mesure la qualité de la solution. Il décide de la poursuite du raisonnement, soit au niveau courant si la solution actuelle est incomplète ou incertaine, soit à un niveau d'affinement directement supérieur si la qualité n'est pas suffisante. L'organisation hiérarchique de la résolution d'un problème dans GREAT permet un traitement progressif qui peut être interrompu à tout moment dès lors que le premier niveau est réalisé. La solution du niveau complètement achevé et de granularité la plus grande est alors exploitable. Le modèle GREAT a été étendu à la résolution d'un problème multi-agents et intégré dans l'architecture temps réel du modèle de société de spécialistes RT-SOS.

REAKT : le modèle de conduite de raisonnement de REAKT s'appuie sur une architecture organisée en deux niveaux. Un premier niveau est composé d'agents périodiques ou sporadiques, caractérisés par un temps d'exécution connu avec précision. L'enchaînement de ces agents est géré par un ordonnanceur qui, à la compilation, permet de déterminer a priori si l'ensemble des agents de ce niveau peut s'exécuter tout en respectant les dates limites d'exécution. Le second niveau de l'architecture [24, 11] est composé d'agents pas forcément déterministes. Ce second niveau est chargé soit d'améliorer les résultats du premier niveau, soit de réaliser des tâches non temps réel. Ce niveau est actif dans les plages de temps pendant lesquelles le premier niveau peut être interrompu tout en préservant toutes les échéances critiques. Afin de respecter ces dernières, nous avons développé un modèle de tâche permettant une exécution progressive.

3.1.5 Planification et robotique

La conduite d'un robot autonome évoluant dans un environnement dynamique est une tâche complexe nécessitant des capacités de raisonnement élaborées. Les évolutions de l'environnement peuvent intervenir de manière impromptue, rendant obsolète la représentation du monde que possède le robot alors que dans le même temps il doit réagir rapidement à la nouvelle situation. Les informations que délivrent ses capteurs peuvent être sous certaines conditions entachées d'erreurs. Le robot doit donc être capable de fusionner les informations dont il dispose alors qu'elles

parviennent de sources hétérogènes dont certaines sont en conflit. Par ailleurs, le monde évoluant, le robot doit être capable de mettre à jour le ou les plans d'action qu'il exécute pour atteindre le ou les buts qu'il poursuit.

La conduite d'un robot mobile est donc une tâche ambitieuse nécessitant la mise en oeuvre d'outils du type de ceux développés dans le projet SYCO. Un tel robot constitue de ce fait une plate-forme d'évaluation dont l'avantage est de pouvoir offrir des conditions proches du monde réel. Dans cette perspective, notre projet s'est équipé d'un robot mobile Nomad 200 que nous avons baptisé Gaston. Cette année, nous avons réalisé une maquette opérationnelle qui valide en partie le modèle d'agent développé dans le projet REAKT [39].

3.1.6 Formalisme à objets, classification et raisonnement à partir de cas

Participants : Jean-Paul Haton, Jean Lieber, Amedeo Napoli

Cette année nous nous sommes intéressés aux formalismes de représentations de connaissances à base d'objets que sont les représentations à objets [5] et les logiques terminologiques [59]. En particulier, ont retenu toute notre attention l'étude de l'organisation interne d'une base d'objets (multi-catégorisation, points de vue, hiérarchies croisées), l'étude des relations sous-jacentes à cette organisation hiérarchique (spécialisation, composition, subsumption), et l'étude du partage de propriétés induit (mécanisme d'héritage). Parallèlement, nous avons également travaillé sur les modes de raisonnement naturellement associés aux représentations hiérarchiques : raisonnement par classification et contraintes. Le processus de classification permet de manipuler et de contrôler une hiérarchie, mais aussi de créer et de gérer de façon dynamique plusieurs hiérarchies croisées, qui correspondent à autant de structures ordonnées décrivant un point de vue sur les connaissances.

L'étude des représentations à objets et du raisonnement par classification est complétée par un ensemble de travaux sur le raisonnement à partir de cas (une mécanisation du raisonnement par analogie) : mise en valeur des rapports existant entre classification et raisonnement à partir de cas, organisation hiérarchique des cas, recherche d'information par classification [58], étude formelle de l'organisation d'une base de cas [46]. Toujours dans ce cadre, ont débuté des travaux qui concernent

plus particulièrement la gestion de plans. Un plan est considéré comme un ensemble d'actions partiellement ordonnées qui présentent des rapports temporels, et la définition d'une relation de subsomption capable d'organiser les plans en vue de leur gestion (description, manipulation, reconnaissance, adaptation de plans analogues) est en cours de mise au point.

Ces recherches ont un caractère théorique et formel, mais aussi pratique, car elles sont destinées à être appliquées dans le cadre de la formalisation de la synthèse organique qui est en cours d'étude au sein du GDR 1093 du CNRS (Traitement informatique de la connaissance en chimie organique) [57, 12]. Nous menons par ailleurs un travail sur les tuteurs intelligents pour l'aide à l'enseignement de la chimie organique dans les premières années d'université [69].

3.1.7 IA et santé

Participants : Christine Bourjot, Marie-Pierre Chouvet, Jean-Paul Haton, Marie-Christine Haton, Pascal Lallemand

Deux membres du projet, C. Bourjot et M.-C. Haton, ont été chargées cette année d'une mission d'étude sur les recherches possibles en Informatique et Santé, en particulier en liaison avec le Pôle Européen de Santé créé à Nancy.

Le domaine de la santé a par ailleurs toujours été un domaine d'application privilégié pour les modèles développés par le projet. Cette année, les travaux marquants concernent deux points :

Aspects cognitifs du diagnostic à différents niveaux d'abstraction. Une approche neuro-analytique générique est en cours de mise au point. Le domaine d'application retenu est celui du diagnostic différentiel des algies cranio-faciales, en relation avec des spécialistes nancéiens. Notre travail s'inscrit dans le contexte général des systèmes hybrides intégrant à la fois des processeurs symboliques et des processeurs connexionnistes. Le but de cette étude est d'abord de proposer un modèle général adapté aux situations d'expertise présentant les caractéristiques suivantes : connaissances multiples (à grande variabilité, à valeur statistique nécessitant une adaptation au cas particulier traité), raisonnement guidé suivant l'instant par une présomption, par un élément nouveau ou par une démarche organisée imposée par un protocole.

Il s'agira ensuite de développer une architecture informatique supportant cette démarche et de la valider dans le domaine choisi. Le raisonnement s'appuie sur les résultats fournis par différents agents : un agent chargé de prendre en compte les symptômes identifiés à l'instant courant pour converger sur des hypothèses de pathologies ; un agent dédié à la reconnaissance de cas "remarquables" ; un module analytique d'analyse de situation et de choix de méthodologie diagnostique [42].

Etude des retombées des techniques avancées de traitement automatique de la parole dans le domaine du handicap auditif.

Après avoir traité dans le passé le problème de la rééducation vocale de jeunes enfants sourds par aide visuelle, nous abordons maintenant le problème du maintien de l'intelligibilité et de la qualité de la voix chez les sujets adultes devenus sourds ou présentant des troubles du langage. Cette étude s'effectue dans la perspective du maintien du sujet dans son emploi, en collaboration avec Thomson/DASM et avec un support financier de l'AGEFIPH.

3.2 Modèles connexionnistes neuromimétiques

Participants : Frédéric Alexandre, Lionel Beaugé, Pascal Blanchet, Brigitte Colnet, Jean-Claude Di Martino, Stéphane Durand, Dominique Fohr, Yifan Gong, Jean-Paul Haton, Yannick Lallement, Nicolas Pican, Jean-François Remm

Cette action s'articule selon deux thèmes. Nous nous employons tout d'abord à créer des modèles et des mécanismes connexionnistes d'inspiration biologique. Il en résulte des unités plus complexes permettant des comportements plus réalistes. Nous travaillons ensuite sur des modèles classiques d'inspiration mathématique, issus ou non de notre laboratoire, afin de les adapter à de nouveaux domaines et à mieux cerner leurs champs d'application. Ici, les domaines d'application sont le traitement de signal, l'analyse de données et les modèles hybrides. Ces deux thèmes sont coordonnés de manière à pouvoir s'enrichir mutuellement.

3.2.1 Inspiration biologique

Participants : Frédéric Alexandre, Lionel Beaugé, Pascal Blanchet, Stéphane Durand, Jean-Paul Haton, Nicolas Pican

Cet axe de recherche concerne en premier lieu les problèmes réputés difficiles de l'IA que sont la prise en compte du temps, l'apprentissage et l'intégration des traitements de niveau perceptif et cognitif. Les travaux que nous menons actuellement portent sur les problèmes de représentation des connaissances et de traitement temporel de l'information [17]. Ils sont principalement issus du modèle de la colonne corticale et de ses extensions. La colonne corticale est une architecture modulaire du cortex, qui constitue une unité logique plus puissante que le neurone et représente une fonctionnalité de base.

Un des buts de notre travail est d'aboutir à un modèle qui soit capable, non seulement d'intégrer de l'information spatiale (ou statique, classification de formes), mais surtout de mémoriser et de dégager un certain nombre de phénomènes à dynamique temporelle [38].

Après avoir établi une classification des différents types de systèmes connexionnistes temporels [36], nous avons développé un modèle de réseau de neurones temporel original qui s'inscrit dans le cadre d'une représentation interne et explicite du temps. En effet, celui-ci se voit interprété au niveau des unités neuronales elles-mêmes. Ce type de représentation s'inspire de la neurobiologie où l'on peut montrer qu'un neurone (ou groupe de neurones) réagit non seulement sous l'effet d'un stimulus sensoriel mais également en fonction d'un contexte imposé par les unités qui l'entourent ou par d'autres aires neuronales. Notre "carte spatio-temporelle" répond à ces critères [37].

Cette carte neuronale est composée d'unités dont la fonctionnalité se décline sous deux aspects: 1) spatial, où chaque unité réagit à une forme particulière de l'entrée, 2) temporel, où chaque unité réagit selon une position temporelle ou un contexte particulier. De façon imagée, il s'agit donc d'extraire des séquences caractéristiques de l'espace d'entrée. Ce modèle se veut générique et pourra donc être utilisé pour différents problèmes tels que la vision active, la planification de mouvements, etc. Le domaine choisi pour valider la carte est la reconnaissance de la parole.

La recherche de l'existence de séquences pertinentes pour des sonorités particulières, nous a conduit à développer un système très simple

de modélisation de mots (les dix chiffres du français) par des séquences d'événements. Ce système est une simplification de la carte "spatio-temporelle" où l'on associe une séquence autonome pour chaque chiffre à apprendre. Les résultats obtenus avec un corpus d'apprentissage (monolocuteur) relativement réduit donnent un taux de reconnaissance s'élevant à 84%. La suite de notre travail consiste à étudier et à expérimenter les différents modes de communication (fonctionnalité et apprentissage) entre d'une part les unités à l'intérieur d'une même carte "spatio-temporelle" et d'autre part les cartes (ou aires) à l'intérieur d'un réseau. Nous poursuivons de plus le développement d'un système neurobiologiquement plausible à base de telles cartes pour la reconnaissance de la parole.

Parmi les phénomènes à dynamique temporelle, citons également la boucle sensori-motrice. Nous nous sommes en particulier intéressés à l'application d'une architecture inspirée de la colonne corticale au problème de l'apprentissage de séquences d'actions par essai-erreur [19]. Le travail réalisé a permis de dégager un certain nombre de similarités avec d'autres approches telles que l'Apprentissage par Renforcement, les algorithmes aux Différences Temporelles ou encore les Systèmes de Classifieurs. Ce travail a également montré l'intérêt de l'architecture connexionniste associative par rapport à ces autres techniques, aussi bien en rapidité de généralisation qu'en précision des connaissances apprises. Le système est en effet capable d'utiliser les premières connaissances acquises pour en généraliser d'autres. De plus, ses connaissances lui permettent non seulement d'agir mais aussi de prévoir. Ceci nous permet d'envisager d'autres développements, vers un système capable de réaliser des plans d'action, par exemple.

Nous nous intéressons également aux aspects de mémorisation (explicite/implicite, court terme/long terme). En collaboration avec des pharmacologues des hopitaux de Strasbourg, des psychologues de Paris et des physiologues de Nancy, nous essayons de comprendre les mécanismes de mémoire à différentes échelles (microscopique/macroscopique). Nous modélisons des principes de modulation [16], de manière à introduire des interactions entre des structures différentes (extrinsèque/intrinsèque). De tels réseaux artificiels peuvent permettre de mieux cerner les phénomènes cognitifs chez l'être vivant et d'améliorer les capacités de mémorisation dans les modèles connexionnistes.

Ces derniers sont créés à partir de NeuroBase, une base neuronale objets que nous réalisons en C++. Elle définit un ensemble de classes et de méthodes permettant l'utilisation et la création d'éléments neuronaux simples ou complexes, de réseaux connexionnistes classiques, de structures à forte inspiration neurobiologique ainsi qu'un ensemble d'outils souvent usités. La puissance des objets (modularité, abstraction, héritage, généricité) nous permet de mettre en œuvre des éléments fortement paramétrés, offrant un éventail de possibilités très vaste. Toutefois, des caractéristiques prédéfinies permettent une utilisation aisée pour un concepteur non averti. Enfin, cette base est accompagnée de divers exemples de simulation.

3.2.2 Mise en œuvre de réseaux neuromimétiques

Participants : Frédéric Alexandre, Pascal Blanchet, Brigitte Colnet, Jean-Claude Di Martino, Dominique Fohr, Yifan Gong, Jean-Paul Haton, Yannick Lallement, Nicolas Pican, Jean-François Remm

Cet axe de recherche consiste à adapter des modèles connexionnistes de manière à pouvoir les utiliser conjointement à d'autres approches symboliques, mathématiques ou relatives au traitement du signal. Une autre adaptation concerne les études d'implantation de tels algorithmes sur des machines parallèles ou sur des réseaux de machines, dans le cadre du Centre Lorrain de Compétence Charles Hermite.

Citons tout d'abord l'utilisation du connexionnisme pour l'interprétation de signaux sonar et radar ainsi que pour la reconnaissance de la parole (*cf.* § 3.3.4 pour le problème de lettres épelées). Deux applications spécifiques sont entreprises sur la chaîne de traitement de l'information sonar (*cf.* § 4.6). La première consiste à réaliser la localisation de sources de bruit d'après des signaux bruts captés par un ensemble d'hydrophones. La deuxième se situe plus en aval et tente de détecter les raies contenues dans un lofargramme (représentation spectrale des signaux acoustiques) et de réaliser leur interprétation [32, 31, 33]. Ici, on s'intéresse principalement à l'intégration des connaissances dans des systèmes distribués. De même, des techniques neuronales classiques sont adaptées de manière à pouvoir caractériser des signatures radar de cibles diverses afin de permettre une interprétation symbolique de la scène analysée (*cf.* § 4.4).

Par ailleurs, d'un point de vue mathématique et statistique, la modélisation de fonctions dans des espaces de grande dimension pose des

problèmes difficiles. Les réseaux de neurones permettent dans certains cas d'obtenir de très bons résultats. Notre approche, afin de résoudre les cas plus difficiles, est une découpe de l'espace en sous-régions où la complexité de la fonction à modéliser est moindre. Nous avons tout d'abord modélisé chaque fonction dans chaque sous-espace par un réseau de neurones (structurellement identique dans chaque sous-espace) [63]. D'une manière plus synthétique, nous avons développé un nouveau type de réseau dans lequel les poids synaptiques sont déterminés par un estimateur neuronal dont les entrées sont une caractérisation du contexte (du sous-espace) dans lequel la fonction doit être modélisée. Ce nouveau concept de réseau est nommé "Orthogonal Weight Estimator" (OWE) [61, 13, 60]. Ces algorithmes ont été appliqués avec succès sur des cas théoriques et sur des applications industrielles réelles [62] (*cf.* § 4.2).

Enfin, les systèmes hybrides symbolico-connexionnistes sont une nouvelle voie de recherche en IA [45]. Le but est de tirer parti des caractéristiques offertes par les systèmes symboliques (par exemple les systèmes à bases de connaissances) et par les systèmes connexionnistes. Ces caractéristiques présentent en effet une certaine complémentarité, les systèmes symboliques étant performants en raisonnement, et les systèmes numériques en reconnaissance de formes [44]. Nous étudions la faisabilité de tels systèmes hybrides, ainsi que leur justification en termes de sciences cognitives. Le projet ESPRIT BRA MIX porte exclusivement sur ces modèles hybrides ; dans ce cadre, nous définissons un environnement distribué pour la définition de ces modèles avant de le valider sur des applications industrielles [43]. Cet environnement, fondé sur des techniques d'Intelligence Artificielle Distribuée nous permet actuellement de définir des couplages entre les plus répandus des modèles connexionnistes d'une part et symboliques d'autre part. Ces couplages sont définis de manière générique de manière à pouvoir les réutiliser simplement d'une application à une autre et de pouvoir en tirer des enseignements sur les implications théoriques de tels couplages.

La machine cellulaire virtuelle (MCV) développée dans notre projet et à Supélec-Metz présente également des capacités intéressantes d'environnement de travail. Nous collaborons donc toujours à son développement et à son utilisation pour ce type de modèle.

3.3 Reconnaissance de la parole

Participants: Frédéric Alexandre, Pascal Blanchet, Anne Bonneau, Christine Bourjot, Anne Boyer, François Charpillet, Sylvie Coste-Marquis, Linda Djezzar, Stéphane Durand, Dominique Fohr, Yifan Gong, Jean-Paul Haton, Marie-Christine Haton, Jean-François Mari, Pierre Marquis, Odile Mella, Roselyne Nguyen, Olivier Siohan, Kamel Smaïli

Reconnaître et comprendre la parole constitue à la fois un défi scientifique et un enjeu économique. Notre projet possède une expérience de plus de vingt ans en la matière. Nos travaux s'appuient sur une connaissance approfondie des processus de la communication parlée et sur la maîtrise des outils complémentaires de la reconnaissance des formes (modèles stochastiques, neuromimétiques, etc.) et de l'IA. Ils ont été poursuivis cette année selon plusieurs voies :

- l'extension de notre modèle abductif/déductif,
- la définition de modèles hybrides stochastiques/connexionnistes,
- la reconnaissance de parole bruitée,
- la mise en œuvre d'une machine à dicter avec une architecture multi-agents,
- la vérification du locuteur.

3.3.1 Aspects phonétiques

Participants: Frédéric Alexandre, Anne Bonneau, Christine Bourjot, Anne Boyer, François Charpillet, Sylvie Coste-Marquis, Linda Djezzar, Stéphane Durand, Dominique Fohr, Yifan Gong, Jean-Paul Haton, Marie-Christine Haton, Pierre Marquis, Odile Mella, Olivier Siohan

Reconnaissance analytique de la parole à l'aide de connaissances : Pour améliorer les algorithmes existants de décodage acoustico-phonétique, nous avons entrepris depuis plusieurs années l'analyse et la modélisation du savoir-faire d'un expert en lecture de spectrogramme, la compétence mise en jeu à ce niveau étant plus accessible à l'analyse car plus consciente [6]. Cette année, nous avons travaillé plus particulièrement sur l'amélioration de la reconnaissance des occlusives et des fricatives (logiciel APHODEX). Nous co-encadrons sur ce thème la thèse de doctorat de Dominique François (projet DIALOGUE).

Nous avons lancé un projet portant sur la définition et l'implantation d'indices acoustico-phonétiques justifiés par des connaissances sur la production, l'acoustique et, quand cela est possible, la perception de la parole. Ce projet comporte une phase d'étude et une phase de reconnaissance automatique. Ce travail est mené en collaboration avec Y. Laprie (projet DIALOGUE) pour ce qui concerne la détection d'événements acoustiques et le suivi de formants.

Indices acoustiques des consonnes occlusives.

Nous poursuivons notre travail sur le lieu d'articulation des occlusives sourdes. Nous avons approfondi nos connaissances sur les occlusives grâce à une série d'expériences originales portant sur la perception de la partie transitoire du bruit d'explosion des occlusives sourdes. Le but de cette expérience était d'apprécier les performances des auditeurs concernant l'identification du lieu d'articulation de l'occlusive indépendamment de la voyelle, d'étudier l'effet de la connaissance de la voyelle subséquente (non présente) sur les scores, et de connaître les principales confusions consonantiques [34].

Nous terminons actuellement, une base de données et de connaissances pour les occlusives afin d'évaluer tous les indices contenus dans le burst. L'objectif est de les valider sur le plan acoustique en s'appuyant sur les résultats des expériences de perception. Les résultats ont montré que les auditeurs sont capables d'identifier très correctement (87%) le lieu d'articulation des occlusives, indépendamment de la connaissance explicite du contexte vocalique. La connaissance de ce dernier permet seulement une légère amélioration des scores d'identification.

Les expériences de perception ainsi que les analyses acoustiques effectuées sur notre base de données nous ont permis de déterminer quels étaient les indices acoustiques les plus robustes. Ceux-ci ont servi de point d'entrée à un système de reconnaissance des occlusives utilisant des réseaux neuronaux. Afin d'apprécier l'utilité des indices acoustiques en reconnaissance, nous comparerons les performances de ce système à celles d'un système similaire utilisant comme point d'entrée des vecteurs neuronaux.

Raisonnement hypothétique

Nous avons poursuivi notre projet original d'utilisation des techniques d'IA consistant en l'application des modèles de raisonnement hypothétique au décodage des occlusives en contexte vocalique [3] et à

l'interaction entre le décodage et le niveau lexical [2]. Ce travail a conduit au système Daphné.

Nous avons poursuivi la définition d'indices pertinents pour l'identification des occlusives. Nous avons défini et testé une première version des indices localisés sur la barre d'explosion et des indices situés sur les transitions formantiques pour les occlusives sourdes en contexte de voyelles centrales.

Nous avons également testé l'ensemble du module de décodage acoustico-phonétique (y compris le rôle de l'étape de révision permettant de lever les incohérences détectées entre indices) sur environ 400 occlusives sourdes en contexte vocalique extraites de deux corpus différents.

En parallèle, la définition d'un module lexical qui tire profit des indices forts pour éliminer directement (indices d'exclusion) ou indirectement (indices de préférence) des mots candidats [29, 28] a été menée à bien. Nous avons montré, dans ce contexte, que l'utilisation des techniques du raisonnement hypothétique permet de maintenir la cohérence des solutions trouvées et d'élaguer l'espace de recherche en éliminant les solutions inutiles [27].

3.3.2 Modèles stochastiques

Participants: Dominique Fohr, Yifan Gong, Jean-Paul Haton, Jean-François Mari, Olivier Siohan

Ces modèles permettent de coder implicitement les connaissances qui interviennent en reconnaissance. Un avantage de ces modèles est que les paramètres peuvent être estimés de façon automatique dans un cadre probabiliste. Nous avons développé de tels modèles, depuis plusieurs années, en parallèle, et parfois de façon complémentaire, avec les modèles à bases de connaissance.

En reconnaissance de la parole, la plupart des modèles stochastiques actuels modélisent les vecteurs d'observation par un processus Markovien caché de premier ordre (HMM1) qui suppose l'indépendance statistique entre les vecteurs consécutifs. Par conséquent, un HMM1 est incapable d'exploiter proprement la corrélation entre les vecteurs acoustiques. De plus, la distribution de la durée des segments acoustiques est mal modélisée, ce qui limite la performance de HMM1 pour la reconnaissance de la parole continue.

Nous avons mené des travaux sur une nouvelle génération de modèles acoustiques, portant sur deux axes : les modèles stochastiques de trajectoire et les HMM du second ordre.

Modèles stochastiques de trajectoire

Une séquence d'observations associées à une unité phonétique constitue une trajectoire. Nous considérons que chaque trajectoire observée est une réalisation d'une variable aléatoire représentée par un modèle stochastique de trajectoire (STM) [41, 7]. Ce modèle est décrit par une loi de mixture définie sur la séquences d'états. A chaque état est associé une fonction de densité de probabilité gaussienne multidimensionnelle, modélisant les vecteurs d'observations dans un espace de paramètres de parole. Dans ce formalisme, les modèles de phonèmes sont donc définis et optimisés au niveau d'une séquence d'états et non au niveau des états individuels. La variation de la durée phonétique est également prise en compte, et est modélisée par des lois de distribution probabiliste.

Le système VINICS, fondé sur ce principe, a fourni en mode dépendant du locuteur, une performance de taux d'erreur de mots inférieur à 1% pour une tâche de reconnaissance de parole continue d'un vocabulaire de 2000 mots.

Le système est rendu plus robuste vis-à-vis de la variabilité de la parole et de l'environnement d'utilisation par différentes méthodes : introduction d'un alignement non linéaire [14] et modélisation explicite au niveau symbolique du contexte phonétique [56]. La reconnaissance de la parole bruitée dans le cadre de STM a également été étudiée (*cf.* § 3.3.4). Ce modèle a été appliqué à la reconnaissance de l'écriture manuscrite cursive [72].

Modèles de Markov cachés d'ordre 2 (HMM2)

Nous avons étudié la modélisation de mots par des processus de Markov cachés continus d'ordre 2 [50], [48] et comparé les résultats de reconnaissance avec ceux fournis par les modèles de Markov d'ordre 1. Les modèles d'ordre 2 montrent une capacité supérieure à modéliser la durée des mots [49].

Ces travaux se poursuivent dans les directions suivantes :

- dans le cadre de la reconnaissance de lettres épelées au téléphone, nous avons intégré conjointement avec des HMM2 des modèles connexionnistes pour discriminer des mots acoustiquement similaires [47].

- utilisation des HMM2 dans un système de reconnaissance de parole continue indépendant du locuteur.
- génération automatique de l'ensemble d'unités de reconnaissance en tenant compte de la précision du modèle et de la quantité de parole du corpus d'apprentissage.
- exploration de la corrélation entre une trajectoire et les contextes phonétiques, les classes de locuteurs, les intonations.
- développement des applications de la reconnaissance de la parole continue et de la localisation des mots en parole continue.
- recherche de méthodes pour combiner les réponses fournies par un HMM2 avec les choix issus d'une classification statistique où chaque mot est représenté par un vecteur de traits d'origine acoustique ou provenant encore d'autres traitements comme des réseaux neuronaux. Ces travaux ont débouché sur la création d'une maquette de système mono-locuteur de reconnaissance de lettres épelées.

3.3.3 Modèles de langage pour la reconnaissance de la parole

Participants : François Charpillet, Dominique Fohr, Yifan Gong, Jean-Paul Haton, Pierre Marquis, Roselyne Nguyen, Kamel Smaïli

Les systèmes de reconnaissance de la parole continue nécessitent des modèles de langage opérationnels permettant au système de reconnaissance de filtrer les hypothèses fournies par les niveaux inférieurs de la chaîne de communication. Pour ce faire, les modèles les plus usités sont de type probabiliste. Ces systèmes sont fondés sur une phase d'apprentissage dans laquelle on apprend à reconnaître des suites de mots de taille fixe ou variable, pour les utiliser ensuite à prédire soit les prochains mots à reconnaître, soit les prochaines classes dans lesquelles se trouvent les mots candidats à la reconnaissance. Ces classes peuvent être purement syntaxiques ou de nature syntaxico-sémantique. Cela permet de poser le problème de la classification : quelle classification choisir.

Dans nos travaux sur la machine à dicter MAUD, nous avons proposé une classification manuelle fondée sur la répartition de mots en classes selon quelques propriétés syntaxico-sémantiques (une dizaine de propriétés ont été prises en compte). L'inconvénient de cette classification est qu'elle est assez générale, autrement dit elle s'adapte mal à des ap-

plications spécifiques. Pour pallier ce problème, il serait commode de pouvoir changer de classification en fonction du domaine étudié et donc finalement de changer de modèle de langage. Pour ce faire, nous avons développé une méthode de classification automatique du lexique. Cette méthode est fondée sur l'algorithme du recuit simulé auquel on a intégré quelques connaissances. Cet algorithme a comme objectif de minimiser l'énergie du système qu'on simule à des températures de plus en plus décroissantes. Dans notre cas, l'énergie à minimiser est la perplexité du langage. La simulation dans notre approche n'est pas complètement aléatoire mais orientée. En effet, on ne teste pas n'importe quel mouvement d'un mot d'une classe à une autre mais seulement ceux qui sont les plus plausibles. Nous avons implanté cette méthode sur de petits corpus, les résultats sont fort satisfaisants, on arrive ainsi à construire automatiquement des classes sémantiques (couleur, direction, verbe de mouvements, prénom, etc.) et des classes syntaxiques (verbes à la troisième personne, noms communs, etc.). Il reste à étudier l'influence du nombre de classes sur la perplexité et d'essayer de trouver d'autres critères de minimisation que la perplexité. La perplexité reste le moyen le plus pratique d'évaluer un modèle de langage, néanmoins il n'a pas été établi que ce critère permet de faire une analyse fine du comportement du modèle de langage.

Nous abordons aussi l'utilisation de la phonologie dans un système de reconnaissance de la parole continue, en collaboration avec l'équipe de G. Pérennou à l'IRIT-Toulouse. Le lexique a été enrichi par des GPC en finales permettant de prendre en compte les variations contextuelles de fin de mots. Un millier de règles phonologiques sont en cours d'exploitation dans la version multi-agents de MAUD utilisant l'outil Atome. Le problème consiste actuellement à trouver la stratégie permettant de traiter au mieux l'ensemble des connaissances du système.

3.3.4 Reconnaissance bruitée

Participants : Dominique Fohr, Yifan Gong, Jean-Paul Haton, Olivier Siohan

La plupart des systèmes de reconnaissance de la parole entraînés dans un milieu calme et employés dans un milieu bruité donnent un taux de reconnaissance fortement dégradé. La cause principale est la divergence acoustique entre l'environnement d'entraînement et celui d'utilisation.

Pour réduire cette divergence, nous avons étudié et comparé un ensemble de techniques de reconnaissance de la parole bruitée.

- Combinaison de modèles stochastiques : STM et HMM

Notre but est de combiner un modèle stochastique de trajectoire (STM) (*cf.* § 3.3.2) entraîné sur de la parole propre, avec un HMM de bruit, afin de reconnaître de la parole bruitée. Nous nous limitons ici à des bruits décorrélés du signal de parole, mais qui peuvent être non stationnaires. Le but de cette approche est de déterminer l'expression de la fonction de densité de probabilité d'un vecteur de cepstre de parole bruitée, à partir de l'expression des fonctions de densité de probabilité d'un vecteur de cepstre de parole propre et de bruit. L'intérêt de cette méthode est que l'on garantit que les performances obtenues avec les modèles compensés, pour des rapports signal-à-bruit élevés ne sont pas dégradées. Des résultats encourageants ont été obtenus sur une tâche de reconnaissance de 206 mots isolés, mono-locuteur, dans un environnement de bruit additif blanc [64].

- Filtrage MMSE par états sur des modèles STMs

Avec cette approche, notre objectif est de définir un filtrage du signal de parole bruitée, afin d'utiliser le système de reconnaissance de parole, entraîné sur de la parole propre, pour reconnaître la parole filtrée. Le cadre de la soustraction spectrale du bruit est la base de ce travail. L'inconvénient principal de la soustraction spectrale est d'introduire une distorsion sur le signal filtré, perçu comme un bruit musical. Une façon de réduire ce bruit consiste à estimer le spectre du signal filtré dans le domaine logarithmique. Ceci est en concordance avec les théories sur la perception de l'oreille humaine, qui précisent qu'il est préférable de minimiser les distorsions dans le domaine du logarithme du spectre, plutôt que dans le domaine spectral linéaire. Nous avons donc utilisé un estimateur de la parole débruitée, défini à partir de modèles de parole propre et d'un modèle de bruit. Cet estimateur minimise l'erreur quadratique moyenne (MMSE) entre le signal propre et le signal filtré dans le domaine spectral logarithmique.

- Adaptation à l'environnement par régression linéaire

L'objectif de cette méthode est de calculer l'expression d'une transformation linéaire des modèles de référence de parole propre, afin

d'effectuer la reconnaissance de parole bruitée. Les transformations linéaires sont calculées à partir d'un corpus d'adaptation de taille réduite. Ces transformations sont telles qu'elles maximisent la probabilité d'observer le corpus d'adaptation en utilisant les modèles compensés. Plusieurs transformations sont estimées, chacune étant spécifique à une classe de sons. L'intérêt de cette méthode est que les transformations sont définies de façon optimales selon un critère objectif. De plus, cette approche n'est pas limitée à une adaptation au bruit, mais peut servir de cadre pour une adaptation au locuteur et au canal d'enregistrement et de transmission.

- Utilisation de l'analyse linéaire discriminante

Cette étude porte sur l'amélioration des performances du système VINICS, pour la reconnaissance de parole propre, en utilisant le cadre de l'analyse linéaire discriminante (LDA). L'objectif de la LDA est de transformer un espace de paramètres constitué de différentes classes, en un nouvel espace, dans lequel les variances inter-classes sont maximales, et les variances intra-classes minimales. Dans notre formulation, une classe est associée à un symbole phonétique. Nous avons évalué cette méthode sur une tâche de reconnaissance de 206 mots isolés, en mode mono-locuteur. Cela a conduit à une amélioration très sensible des performances. En particulier, nous avons obtenu une diminution de l'erreur de reconnaissance de 50%. Dans deux nouvelles formulations, nous associons une classe à chaque trajectoire acoustique spécifique à chaque symbole phonétique, ainsi qu'une classe à chaque état de chaque trajectoire acoustique spécifique à chaque symbole phonétique. Une évaluation de la robustesse de la LDA dans le bruit vient d'être également achevée. Nous avons montré que le paramétrage LDA est bien adapté au bruit, mais qu'il est par contre très peu robuste aux variations du rapport signal-à-bruit.

- Transformation de base

Cette dernière approche consiste à "filtrer" la parole bruitée afin de pouvoir la reconnaître en utilisant un système entraîné sur de la parole propre. Dans ce but, nous avons développé une méthode permettant de projeter un espace de parole bruitée sur un espace de parole propre. Un vecteur de parole peut s'exprimer dans un espace comme une combinaison linéaire des vecteurs d'une base de cet espace. Notre formalisme suppose que les distances rela-

tives entre vecteurs de parole sont indépendantes de l'espace. Sous cette hypothèse, la transformation d'un vecteur d'un espace vers un autre s'effectue en transformant uniquement la base de l'espace. La transformation consiste donc à rechercher une correspondance entre les vecteurs constituant les différentes bases. Les éléments d'une base correspondent aux phonèmes issus d'un corpus d'adaptation de taille réduite. Le corpus d'adaptation étant étiqueté phonétiquement, il est possible d'obtenir directement une correspondance entre les éléments issus d'une version propre et bruitée du corpus d'adaptation. Cette méthode a été évaluée dans [65] sur une application de reconnaissance de mots isolés, et a été comparé avec la combinaison de modèles dans [64].

Nous travaillons aussi, en collaboration avec Jean-Claude Junqua (STL Santa Barbara), sur la reconnaissance de la parole en milieu bruité à l'aide de réseaux connexionnistes. Nous nous intéressons plus particulièrement aux mots acoustiquement similaires tels que les lettres épelées. Notre méthode se distingue des méthodes classiques par le fait que seules des trames discriminantes bien choisies sont utilisées lors de la reconnaissance. Nous avons testé différents types d'apprentissage et montré leurs influences sur les taux de reconnaissance [23]. Nous co-encadrons sur ce thème la thèse de doctorat de Laurent Buniet (projet DIALOGUE).

3.3.5 Vérification et identification du locuteur :

Participants : Florent Fauchier, Yifan Gong, Jean-Paul Haton, Marie-Christine Haton, Odile Mella, Olivier Siohan

Dans le domaine de la reconnaissance automatique du locuteur, deux démarches complémentaires peuvent être envisagées, l'une fondée sur la mise en oeuvre de méthodes de type reconnaissances des formes, l'autre sur l'extraction de connaissances acoustico-phonétiques. Ces deux démarches sont développées conjointement au sein du projet SYCO.

Ainsi, nous avons développé un système de vérification automatique du locuteur capable, en mode dépendant du texte, de vérifier l'identité d'une personne à partir d'une courte phrase d'environ une seconde. Un système de vérification de locuteur se compose essentiellement de deux parties : les modèles acoustiques des locuteurs et un système de décision. Notre approche pour les modèles acoustiques consiste à exploiter la corrélation acoustique, spécifique au locuteur, entre deux ensembles

de vecteurs paramétriques relatifs au même locuteur. Pour cela, nous avons modélisé cette corrélation à l'aide de notre technique d'interpolation non-linéaire vectorielle. Cette technique implantée sous forme de perceptrons multicouches, minimise l'erreur d'interpolation au sens des moindres carrés. A partir de la sortie du modèle acoustique, le système de décision calcule, de façon spécifique au locuteur, un seuil de décision (acceptation/rejet) optimal pour chaque locuteur. Nous avons étudié deux types de critères de décision bayésienne, celui du risque minimal et celui de l'erreur minimale. Les expériences de validation du système de vérification ont porté sur 72 locuteurs. L'apprentissage des modèles acoustiques a utilisé quatre prononciations de quatre phonèmes contenus dans une phrase et la vérification une autre prononciation de la même phrase. Nous avons obtenu un taux d'erreur total de 0.45% [40].

Parallèlement, une étude sur la caractérisation acoustico-phonétique des locuteurs en vue de leur identification automatique a permis d'étudier la pertinence des voyelles orales au niveau de leurs trois premiers formants. Les résultats obtenus ont été comparés avec ceux des études similaires sur la langue anglaise et interprétés en tant que sources de différences articulatoires entre locuteurs (degré de labialisation des voyelles arrondies, degré d'ouverture des voyelles ouvertes,...). Les résultats ont également été comparés aux célèbres facteurs d'échelle k_i entre les fréquences formantiques féminines et masculines[52].

4 Actions industrielles

4.1 Coopération de réseaux neuroniques

CONVENTION AVEC LE CNET LANNION LAB/OCM/TEP

Participant : Frédéric Alexandre

Cette action, qui s'est achevée cette année, porte sur l'étude de protocoles de communication entre réseaux connexionnistes. L'approche choisie a consisté à utiliser des unités Sigma-Pi (combinant somme et produit de leurs entrées) comme modèle de ce type de communication, par référence biologique à ce processus dans les aires corticales. L'application retenue, la localisation et la reconnaissance de caractères, a mis en évidence une plus grande robustesse à la translation de notre modèle par rapport aux modèles classiques.

4.2 Etude connexioniste pour le contrôle de machines industrielles

CONVENTION AVEC SOLLAC

Participants : Frédéric Alexandre, Nicolas Pican

SOLLAC, filiale d'Usinor-Sacilor pour les produits plats, réalise dans ses usines lorraines la totalité de la filière de production de l'acier. Un problème essentiel qui se pose aux industriels consiste à réaliser le contrôle le plus optimal possible de leurs machines. Une première étude nous a permis de déterminer le pré réglage d'un *Skin-Pass* (machine effectuant l'écrouissage) en fonction de l'acier présenté et de la qualité désirée. Jusqu'à présent, ces calculs étaient réalisés par un modèle mathématique approché qui fournit des résultats avec une erreur moyenne de 24%.

Une technique neuromimétique a été testée (*cf.* § 3.2.2) et a permis rapidement d'obtenir des résultats supérieurs au modèle mathématique (erreur moyenne : 12%). Après l'implantation de ce réseau sur la chaîne de production de l'acier nous avons pu valider ces résultats et tirer les conclusions permettant de les améliorer. Actuellement, 80% de la production est traitée par notre réseau et permet à SOLLAC de diminuer substantiellement ses pertes de matériaux et de mieux satisfaire ses clients. Le système a donné lieu à la prise d'un brevet international.

Ce succès nous fait nous orienter maintenant vers une autre application: le contrôle d'un four de recuit continu mixte, très grosse installation à réguler en température de façon à préchauffer de manière homogène les tôles avant leur écrouissage.

4.3 Estimation de champ radioélectrique par réseaux neuroniques

CONVENTION AVEC LE CNET BELFORT (BEL/POH/CDI)

Participant : Frédéric Alexandre

Les communications avec les mobiles constituent un point stratégique pour France Télécom. Un des problèmes consiste à placer les antennes d'émission, de manière à couvrir le territoire de manière optimale. Cette décision dépend de nombreux paramètres et fait généralement

appel à des modèles physiques. Profitant de l'existence d'une importante campagne de mesure, nous avons testé l'apport éventuel de réseaux neuromimétiques pour cette application. Les résultats ont été positifs, puisqu'en quelques mois de travail, nous avons dépassé les meilleurs modèles empiriques.

4.4 Interprétation de signaux radar

CONVENTION AVEC LCTAR

Participants : Frédéric Alexandre, Jean-François Remm, Jean-Paul Hanton

Le Centre Thomson d'Application Radar (LCTAR) souhaite pouvoir traiter automatiquement et rapidement des signaux radar complexes dont seule une petite partie contient les informations pertinentes. Nous avons étudié des méthodes supervisées et non supervisées issues de modèles classiques pour détecter et identifier ces cibles immobiles (*cf.* § 3.2.2). Ce premier travail à court terme nous permet déjà d'atteindre les résultats des modèles mathématiques. Nous obtenons en particulier de bons résultats quand l'orientation de la cible est connue. Nous abordons maintenant le cas des cibles mobiles.

4.5 Ordonnancement de pistes par réseaux neuronaux

CONVENTION AVEC LE CPM (CENTRE DE PROGRAMMATION DE LA MARINE)

Participants : Frédéric Alexandre, François Charpillet, Jean-Michel Gallone

Une des tâches du CPM consiste à développer des systèmes d'exploitation navale des informations tactiques. Un des aspects les plus cruciaux consiste à localiser, suivre, classer les objets présents dans la zone d'intérêt d'un navire. Nous nous sommes intéressés à la planification de tâches permettant de traiter de manière optimale les cibles détectées d'après leurs caractéristiques. Nous avons testé et comparé l'intérêt de techniques d'optimisation connexionnistes pour ce genre de tâches avec des techniques classiques.

4.6 Interprétation de signaux sonar

CONVENTION AVEC LE CERDSM (TOULON)

Participants : Frédéric Alexandre, Brigitte Colnet, Jean-Claude Di Martino, Jean-Paul Haton

Le thème général de cette étude a trait à l'utilisation des modèles neuromimétiques dans les différentes phases de l'interprétation de signaux sonar (*cf.* § 3.2.2). Le travail actuel concerne la localisation de sources et la détection de raies sur des lofargrammes, représentations fréquentielles de ces signaux. Différentes techniques distribuées (cartes auto-organisatrices, système de vote, traitement d'image puis poursuite) ont été développées et ont donné de bons résultats dont certains sont transposables au domaine de reconnaissance des formes.

4.7 Traitement informatique de la connaissance en chimie organique

CONVENTION AVEC LE CNRS, CRIN, LIRMM, CCIPE, INSTITUT DE RECHERCHES SERVIER, FRAGMENTEC-COGNITECH, ROUSSEL-UCLAF, SANOFI-CHIMIE, SEMA GROUP, MELODIE

Participants : Jean Lieber, Amedeo Napoli

L'objectif à long terme de ce projet est de construire un système d'aide à la planification de synthèse de molécules en chimie organique. Notre but est de représenter les plans et stratégies d'un chimiste dans le cadre d'un système à bases de connaissances (en s'appuyant sur les représentations à objets et la classification), pour pouvoir les réutiliser de façon automatique à l'aide de méthodes d'analogie et de raisonnement à partir de cas.

4.8 Détection et interprétation de cris

CONVENTION AVEC LA RATP

Participants : Florent Fauchier, Dominique Fohr, Yifan Gong, Jean-Paul Haton, Jean-François Mari, Odile Mella, Olivier Siohan

L'objectif de cette étude est d'identifier des cris de détresse dans les couloirs et stations du Métropolitain de façon à améliorer la surveillance.

Ce problème est très difficile du fait des phénomènes de bruitage, réverbération, etc. qui déforment les signaux.

4.9 Reconnaissance de la parole

CONVENTION AVEC LE CEA-CADARACHE

Participants : Yifan Gong, Jean-Paul Haton

L'étude fait partie du projet IMAGIN mené par le CEA dans le domaine des bases de données de centrales nucléaires. Nous nous intéressons à la reconnaissance de la parole, fondée sur notre système VINICS, au sein de ce projet. Cette année, le système a été réécrit en C++ et son utilisation dans l'ambiance bruitée d'une centrale a été envisagée.

4.10 Outils de raisonnement temps-réel

PROJET ESPRIT II REAKT AVEC THOMSON, SYSECA, GEC, GMV, UNIV. POL. VALENCIA, ETNOTEAM, COMPUTAS

Participants : Anne Boyer, François Charpillet, Jean-Michel Gallone, Jean-Paul Haton, Abdelillah Mouaddib

L'objectif de ce projet est le développement de systèmes multi-agents temps-réel, notamment pour la conduite de procédés industriels, la robotique, etc. Cet objectif a été atteint avec la mise au point de la boîte à outils REAKT. Le projet s'est terminé en novembre 1994. Nous y avons abordé les aspects de raisonnement en temps contraint, de garantie de temps de réponse et de planification.

4.11 Système multi-agents distribué : DBB

CONVENTION DU MINISTÈRE DE L'INDUSTRIE AVEC ILOG ET MATRA

Participants : François Charpillet, Jean-Paul Haton, Abdelillah Mouaddib

A la suite des travaux que nous avons menés précédemment sur la version temps réel d'ATOME (appelée ATOME-TR), notamment avec l'aide de la DRET, ce projet a pour but de développer l'outil DBB qui apporte en plus le caractère distribué sur un ensemble de machines. Le projet

SYCO intervient en tant que conseil des industriels Ilog et Matra qui mènent ce projet.

4.12 Reconnaissance de la parole via le téléphone

PROJET EUROPÉEN COST 249 RÉUNISSANT UNE QUINZAINE DE PARTENAIRES

Participants: Dominique Fohr, Yifan Gong, Jean-Paul Haton, Jean-François Mari, Olivier Siohan

Ce projet s'inscrit dans le programme européen de télécommunications COST. Il a pour but d'étudier les techniques de reconnaissance de parole via le réseau téléphonique. Nous comptons tester nos modèles (notamment HMM et STM) dans les conditions de bruit engendrées par de tels réseaux.

5 Actions nationales et internationales

- Participation active à des groupes de travail et associations : GDR-PRC Communication Homme-Machine (J.-P. Haton membre du bureau de direction) et IA (F. Charpillet membre du bureau de direction), International Association for Pattern Recognition (J.-P. Haton, membre du Governing Board), Association française d'intelligence artificielle (AFIA) (J.-P. Haton, membre du bureau), GRIPIC (F. Alexandre, P. Blanchet, V. Chevrier), ILGU (M.-C. Haton, J.-P. Haton), ILDI (J.-P. Haton), AFCET-AFIA *Multi-Agents* (V. Chevrier), groupe de travail sur l'IA distribuée MARCIA du GDR-PRC IA (F. Charpillet, V. Chevrier, R. Foisel, J.-P. Haton), groupe "Objets et Classification" du GDR-PRC IA (A. Napoli), groupe "Évolution des langages à objets" du GDR-PRC Programmation (A. Napoli), GDR 1093 du CNRS "Traitement informatique de la connaissance en chimie organique" (A. Napoli), groupes GT2 (S. Coste-Marquis) et GT7 (K. Smaïli) du GDR-PRC Communication Homme-Machine, ACTH (L. Beaugé, S. Durand, Y. Lallement, N. Pican, J.-F. Remm).
- Participation à des comités de lecture de revues: Pattern Recognition, Journal of Intelligent Manufacturing, IECE Transactions,

Génie Educatif, Journal d'Acoustique, Speech Communication, Revue d'Intelligence Artificielle, Computer Speech and Language, Int. J. Engineering Intelligent Systems.

- Participation à des comités de programme: NeuroNîmes, Décembre 94; IA'94, juin 94, Paris; CHI, Avril 95, Boston; Journées Francophones d'IA distribuée, Avril 94, Voiron; INFORSID'94 Mai 94, Aix-en-Provence; LMO'94, Octobre 94, Grenoble.
- Participation à des comités scientifiques de divers laboratoires et organismes: CNET, CRIM (Montréal), DRET, INRETS, IRSID, IRCAM, IRISA, IRIT, EERIE, IRSIT (Tunis).
- Participation au conseil scientifique de l'Office Parlementaire d'évaluation des choix scientifiques et techniques (J.-P. Haton), au conseil d'administration de la Fondation Bull (J.-P. Haton).
- Participation aux projet européens :
ESPRIT (projets REAKT et MIX, réseaux d'excellence ELSNET et NEURONET), HCM (projet SPIN), Projet MORE.
- Coopérations avec Centro di Informatica de Valencia, ENIT et IRSIT, Tunis (échanges en parole et image; convention pour le développement d'un système de compréhension de documents arabes, dans le cadre du projet national pour l'arabisation), Université d'Oran (accueil de doctorants), Ecole Polytechnique Fédérale de Lausanne, Université de Constantine (accueil de stagiaires), Université de Xi'an (Chine) (accueil de professeur invité), Université de Nankin (projet commun), Université de Erlangen (accueil de stagiaires), Institut National des Télécommunications, Le Caire (accueil de doctorants et de stagiaires), Speech Technology Laboratory à Santa-Barbara, USA (projet commun), Université de Sherbrooke, Canada (accueil d'un professeur invité), Université de Stanford, USA (envoi d'un stagiaire post-doctoral et accueil d'un professeur invité), Université de Hong-Kong (accueil d'un stagiaire post-doctoral).

6 Diffusion des résultats

6.1 Diffusion des produits

Nous nous efforçons de maintenir une politique de diffusion vers les utilisateurs des logiciels que nous développons. En plus d'actions ponctuelles dans le cadre de projets cités au 4, les logiciels suivants ont été transférés :

- ATOME, outil de développement de systèmes multi-agents fondé sur le modèle de tableau noir. Cet outil a déjà été installé dans plus de trente centres dans le monde ;
- éditeurs interactifs de parole et de signaux : à partir de la souche commune SNORRI développée à Nancy, nous mettons au point des éditeurs adaptés aux besoins d'un utilisateur (notamment avec le CEA-Cadarache).

6.2 Manifestations

- Séminaire MARCIA (projet du PRC IA) sur les systèmes multi-agents, Paris, Décembre 1994 ;
- Rencontres des projets IA de l'INRIA, Nancy, Décembre 1994 ;
- Séminaire "Reconnaissance automatique de la parole" du GDR-PRC "Communication Homme-Machine", Mars 94 ;
- Séminaire "IA et Défense" dans le cadre de l'Institut des Hautes Etudes de la Défense Nationale (IHEDN-Lorraine), Janvier 94 ;
- Participation à l'organisation de manifestations : Congrès International IA'94, Paris (J.-P. Haton président du comité de programme) ; Congrès "Interfaces des mondes réels et virtuels", Montpellier, mars 94 ; 6th IEEE Int. Conf. on Tools with AI, New-Orleans, Nov. 1994 (P. Marquis, vice-chairman) ; Secondes rencontres des jeunes chercheurs en IA, Marseille, Sept. 1994 ; 2nd European Workshop on Case-Based reasoning, Royaumont, Nov. 1994 (J.-P. Haton co-chairman) ; AIRTC'94, Valencia, Octobre 1994 ;

6.3 Actions d'enseignement

- Forte participation à divers enseignements dans les établissements lorrains (Université de Nancy I et II, INPL, Supélec) : Maîtrise et DEA d'Informatique, IUT, MIAGE, DESS d'Informatique, DESS IMOI, DESS de Productique, Maîtrise de Biologie, Ecole des Mines, ENSEM, Ecole de Géologie, ESIAL, ESSTIN ;
- Responsabilité du DESS d'Informatique de Nancy et de l'option "BD, GL et IA" de ce DESS (M.-C. et J.-P. Haton) ;
- Responsabilité de Département d'Informatique de l'Université Henri Poincaré Nancy I (M.-C. Haton) et de la Commission Informatique du Pôle Lorrain de Gestion de l'Université de Nancy II (C. Bourjot) ;
- Cours et séminaires d'IA et de reconnaissance des formes : CNAM, III-Brest, CNIL, ORSTOM ;
- Participation à des jurys de thèse de doctorat (F. Alexandre, F. Charpillat, D. Fohr, M.-C. Haton, J.-P. Haton, A. Napoli).

7 Perspectives

Le projet SYCO compte poursuivre à l'avenir ses recherches sur la "machine intelligente". Ces travaux comportent deux aspects complémentaires que le projet s'est toujours efforcé de mener en parallèle :

- un aspect fondamental d'étude des processus perceptifs et cognitifs de l'être humain, ainsi que de mise au point des modèles (connexionnistes, statistiques, symboliques ou hybrides) correspondants,
- un aspect appliqué de développement de maquettes de systèmes pour tester les modèles et les méthodes mises au point.

Parmi les travaux marquants à venir on peut citer :

- le test sur une grande échelle de nos systèmes de reconnaissance de la parole, dans le cadre de l'AUPELF pour la langue française, mais aussi avec les grandes bases de données américaines,
- l'étude de modèles hybrides connexionnistes-stochastiques-symboliques, notamment dans le projet ESPRIT BRA MIX,

- le développement de fonctionnalités (fusion de données, navigation, planification, etc.) autour de notre robot mobile Gaston,
- l'étude de la distribution et de la parallélisation de certains de nos modèles en liaison avec le centre lorrain Charles Hermite,
- l'extension des méthodes et modèles relatifs aux systèmes multi-agents (modèles d'agents et de systèmes, interactions, acquisition de connaissances),
- la poursuite de la réflexion lancée en 1994 sur les domaines potentiels de recherche sur informatique et santé.

8 Publications

Thèses

- [1] N. BOUZID, *Modèle Qualitatif: Analyse, Formalisation et mise en oeuvre*, thèse de doctorat, Université Henri Poincaré, Nancy I, Novembre 1994.
- [2] S. COSTE-MARQUIS, *Utilisation des techniques du raisonnement hypothétique pour la reconnaissance automatique de la parole*, thèse de doctorat, Université Henri Poincaré, Nancy I, Octobre 1994.

Articles et chapitres de livre

- [3] A. BONNEAU, F. CHARPILLET, S. COSTE-MARQUIS, J. P. HATON, Y. LAPRIE, P. MARQUIS, «Towards a Multilevel Model for Hypothetical Reasoning in Continuous Speech Recognition», in : *Levels in Speech Communication: Relations and Interactions*, C. Sorin, J. Mariani, et H. Méloni (éd.), Elsevier, 1994.
- [4] F. CHARPILLET, P. THÉRET, «Dossier: IA et temps réel», *Bulletin de l'AFIA*, 17, Avril 1994, p. 19-42.
- [5] L. DEKKER, A. NAPOLI, «Report on the IJCAI-93 Workshop on "Object-Based Representation Systems" », *SIGART Bulletin* 5, 1, 1994, p. 57-59.
- [6] D. FOHR, J.-P. HATON, Y. LAPRIE, «Knowledge-based Techniques in Acoustic-Phonetic Decoding of Speech: Interest and Limitations», *International Journal of Pattern Recognition and Artificial Intelligence* 8, 1, 1994, p. 133-153.
- [7] Y. GONG, J. P. HATON, J. F. MARI, «Issues in Acoustic Modeling of Speech for Automatic Speech Recognition», in : *Progress and Prospects of Speech Research and Technology*, H. Niemann, R. de Mori, et G. Hanrieder (éd.), Infix, 1994, p. 34-44.

- [8] Y. GONG, J.-P. HATON, «Principles and Applications of the VINICS Continuous Speech Recognition Systems», *in: Advanced Speech Applications: European Research on Speech Technology*, K. C. Varghese, S. Pfleger, et J. P. Lefevre (éd.), Springer Verlag, 1994.
- [9] J.-P. HATON, «Problems and Solutions for Noisy Speech Recognition», *Journal de Physique IV, Colloque C5*, Mai 1994, p. 439-448.
- [10] A. I. MOUADDIB, «Progressive Goal-Directed Reasoning for Real-Time AI Systems», *International Journal of Engineering Intelligent Systems*, 1994, To appear.
- [11] A. MENSCH, D. KERSUAL, A. CRESPO, F. CHARPILLET, E. PESSI, «REAKT: A Real-Time Architecture for Time-Critical Knowledge-Based Systems», *Intelligent Systems Engineering Journal*, 1994, To appear.
- [12] A. NAPOLI, C. LAURENÇO, R. DUCOURNAU, «An object-based representation system for organic synthesis planning», *International Journal of Human-Computer Studies* 41, 1/2, November 1994, p. 5-32.
- [13] N. PICAN, J.-C. FORT, F. ALEXANDRE, «An On-Line Learning Algorithm for the Orthogonal Weight Estimation of MLP», *Neural Processing Letters* 1, 1, September 1994, p. 1-4.

Communications à des congrès, colloques, etc.

- [14] M. AFIFY, Y. GONG, J.-P. HATON, «Nonlinear Time Alignment in Stochastic Trajectory Models for Speech Recognition», *in: Proceedings International Conference on Spoken Language Processing*, p. 291-293, Yokohama (Japan), September 1994.
- [15] F. ALEXANDRE, F. GUYOT, «Evaluation d'un modèle connexionniste simple pour la reconnaissance automatique de caractères», *in: Actes Colloque National sur l'Écrit et le Document*, Rouen, Juillet 1994.
- [16] L. BEAUGÉ, F. ALEXANDRE, «Utilisation d'unités Sigma-Pi pour la modélisation de processus de mémorisation», *in: Actes 7èmes Journées Neurosciences et Sciences de l'Ingénieur*, P. Marchal (éd.), p. 51-54, Chamonix, Mai 1994.
- [17] L. BEAUGÉ, S. DURAND, Y. LALLEMENT, «Approches pluridisciplinaires pour l'étude de la mémoire, du temps et du raisonnement en neuromimétisme», *in: VALGO*, p. 9-22, Chamonix, Mai 1994.
- [18] F. L. BER, J. BACHACOU, «Un système multi-agents d'aide à l'interprétation d'images satellitaires», *in: Actes AFCET*, Paris, Janvier 1994.
- [19] P. BLANCHET, «An Architecture for Representing and Learning Behaviors by Trial and Error», *in: Proceedings Congress SAB'94: from*

- Animats to Animats III , Simulation of Adaptive Behavior*, Brighton (England), August 1994.
- [20] M. BOASSON, A. I. MOUADDIB, «Comparison of two Solutions to a Reactive Planning Problem», *in: Proceedings Conference on Expert Systems and their Applications*, Paris, May 1994.
- [21] M. BOUZID, F. CHARPILLET, P. MARQUIS, «Assumption-Based Truth Maintenance in Presence of Temporal Assertions», *in: Proceedings 6th IEEE International Conference on Tools with Artificial Intelligence*, C. Koutsou, Geras (éd.), IEEE Computer Society Press, New Orleans (Louisiane, USA), 1994.
- [22] M. BOUZID, «Extension d'ATMS à une logique temporelle numérique et symbolique», *in: Actes Secondes Rencontres Nationales des Jeunes Chercheurs en Intelligence Artificielle*, p. 333, Marseille, 1994.
- [23] L. BUNNET, D. FOHR, J. C. JUNQUA, «Une méthode connexionniste pour la reconnaissance de mots enchaînés en milieu bruité», *in: Actes 7èmes Journées Neurosciences et Sciences de l'Ingénieur*, P. Marchal (éd.), p. 27–30, Chamonix, Mai 1994.
- [24] F. CHARPILLET, A. BOYER, «Incorporating AI Techniques into Predictable Real-Time Systems: Reakt Outcome», *in: Actes 14èmes Journées Internationales d'Avignon*, Paris, Juin 1994.
- [25] V. CHEVRIER, «GTMAS: A Tool for Prototyping and Assessing Design Choices in Multi-Agent Systems», *in: Actes Quatorzièmes Journées Internationales d'Avignon IA 94*, EC2, p. 161–170, Paris, Mai 1994.
- [26] V. CHEVRIER, «Un mécanisme pour la mise au point de systèmes multi-agents», *in: Actes Deuxièmes Journées Francophones sur l'Intelligence Artificielle Distribuée et Systèmes Multi-Agents*, Grenoble, Mai 1994.
- [27] S. COSTE-MARQUIS, «Hypothetical Reasoning for Automatic Recognition of Continuous Speech», *in: Proceedings 11th European Conference on Artificial Intelligence*, A. Cohn (éd.), John Wiley & Sons, Ltd., p. 565–569, Amsterdam (The Netherlands), 1994.
- [28] S. COSTE-MARQUIS, «Interaction between Most Reliable Acoustic Cues and Lexical Analysis», *in: Proceedings 3rd International Conference on Spoken Language Processing*, Yokohama (Japan), 1994.
- [29] S. COSTE-MARQUIS, «Interaction entre le décodage acoustico-phonétique et l'analyse lexicale grâce aux techniques du raisonnement hypothétique: le système DAPHNE», *in: Actes XXèmes Journées d'Etude sur la Parole*, p. 473–478, Trégastel, Juin 1994.
- [30] J. DI MARTINO, J.-F. MARI, B. MATHIEU, K. PEROT, K. SMAÏLI, «Which Model for Future Speech Recognition Systems: Hidden Markov Models or Finite-State Automata», *in: Proceedings International Confe-*

rence on Acoustics, Speech and Signal Processing, Adelaide (Australia), avril 1994. à paraître.

- [31] J. C. DI MARTINO, F. ALEXANDRE, «Curve Line Extraction using Statistical Perceptual Grouping», *in : Proceedings EUSIPCO*, Edinburgh (Scotland), September 1994.
- [32] J. C. DI MARTINO, B. COLNET, M. DI MARTINO, «The Use of Non Supervised Neural Networks to Detect Lines in Lofargram», *in : Proceedings International Conference on Acoustics, Speech and Signal Processing, II*, IEEE, p. 293–296, Adelaide (South Australia), April 1994.
- [33] J. C. DI MARTINO, B. COLNET, «Image Segmentation by Non Supervised Neural Networks», *in : Proceedings Symposium on Electronic Imaging: Science and Technology, 2182 Image and Video Processing II*, p. 350–356, San José (USA), February 1994.
- [34] L. DJEZZAR, A. BONNEAU, Y. LAPRIE, «Effet de la connaissance de la voyelle sur la perception du bruit d'explosion des occlusives», *in : Actes XXèmes Journées d'Etude sur la Parole*, p. 234–238, Trégastel, Juin 1994.
- [35] N. DOGGAZ, J. P. HATON, «Approximate Reasoning in Multi-Agent Systems: the ATOMIC Model», *in : Proceedings Second World Congress on Expert Systems*, janvier 1994.
- [36] S. DURAND, F. ALEXANDRE, «A Neural Network based on Sequence Learning; Application to Spoken Digits Recognition», *in : Actes Septièmes Journées Internationales sur les Réseaux Neuromimétiques et leurs Applications*, EC2, Marseille, Décembre 1994.
- [37] S. DURAND, F. ALEXANDRE, «Mécanisme d'apprentissage spatio-temporel pour l'organisation sensori-motrice», *in : Actes 7èmes Journées Neurosciences et Sciences de l'Ingénieur*, P. Marchal (éd.), p. 71–74, Chamonix, Mai 1994.
- [38] S. DURAND, «Prise en compte de la dimension temporelle dans les réseaux neuromimétiques», *in : Actes Seondes Rencontres Nationales des Jeunes Chercheurs en Intelligence Artificielle*, p. 297, Marseille, 1994.
- [39] J.-M. GALLONE, F. CHARPILLET, V. CHEVRIER, «Un modèle d'agent pour un raisonnement contraint par le temps», *in : Actes Journée PRC-IA sur les Systèmes Multi-Agents*, Paris, Décembre 1994.
- [40] Y. GONG, J.-P. HATON, «Non-Linear Interpolation Methods for Speaker Recognition and Verification», *in : Proceedings ESCA Workshop on Automatic Speaker Recognition, Identification and Verification*, Martigny (Switzerland), April 1994.
- [41] Y. GONG, J.-P. HATON, «Stochastic Trajectory Modeling for Speech Recognition», *in : Proceedings IEEE International Conference on Acous-*

- tics, Speech and Signal Processing, 1*, IEEE, p. 57–60, Adelaide (South Australia), April 1994.
- [42] M.-C. HATON, P. LALLEMAND, P. LONCHAMP, « Définition d'une architecture hybride pour le diagnostic médical par analyse de la connaissance d'experts », in : *Actes Septième's Journées Internationales sur les Réseaux Neuromimétiques & leurs Applications*, EC2, Marseille, Décembre 1994.
- [43] M. HILARIO, C. PELLEGRINI, F. ALEXANDRE, « Modular Integration of Connectionist and Symbolic Processing in Knowledge-Based Systems », in : *Proceedings International Symposium on Integrating Knowledge and Neural Heuristics*, Pensacola Beach (Florida, USA), May 1994.
- [44] Y. LALLEMENT, S. DURAND, « Ancrage de symboles dans des systèmes connexionnistes dédiés à des tâches de perceptions sensorielles », in : *Actes 7èmes Journées Neurosciences et Sciences de l'Ingénieur*, P. Marchal (éd.), p. 221–224, Chamonix, Mai 1994.
- [45] Y. LALLEMENT, « Modèles hybrides et intelligence artificielle », in : *Actes Premier Colloque des Jeunes Chercheurs en Sciences Cognitives*, p. 268, La Motte d'Aveillans, Mars 1994.
- [46] J. LIEBER, « A Criterion of Comparison between two Cases Bases », in : *Proceedings Second European Workshop on Case-Based Reasoning*, Chantilly, 1994.
- [47] J.-F. MARI, D. FOHR, Y. ANGLADE, J.-C. JUNQUA, « Hidden Markov Models and Selectively Trained Neural Networks for Connected Confusable Word Recognition », in : *Proceedings International Conference on Spoken Language Processing*, p. S26.11, Yokohama (Japan), September 1994.
- [48] J.-F. MARI, D. FOHR, J.-P. HATON, « Modèles stochastiques d'ordre 1 et 2 », in : *Actes XXèmes Journées d'Etude sur la Parole*, p. 199–202, Tregastel, Juin 1994.
- [49] J.-F. MARI, D. FOHR, J.-P. HATON, « Quelques aspects des modèles de Markov en reconnaissance de la parole », in : *Actes Séminaire sur la Reconnaissance Automatique de la Parole*, Nancy, Mars 1994.
- [50] J.-F. MARI, J.-P. HATON, « Automatic Word Recognition Based on Second-Order Hidden Markov Models », in : *Proceedings International Conference on Spoken Language Processing*, p. S07.17, Yokohama (Japan), September 1994.
- [51] P. MARQUIS, « Possible Models Approach via Independency », in : *Proceedings 11th European Conference on Artificial Intelligence*, A. Cohn (éd.), John Wiley & Sons, Ltd., p. 336–340, Amsterdam (The Netherlands), 1994.

- [52] O. MELLA, «Extraction of Formants of Oral Vowels and Critical Analysis for Speaker Characterization», *in: Proceedings ESCA Workshop on Automatic Speaker Recognition, Identification and Verification*, Martigny (Switzerland), April 1994.
- [53] A. I. MOUADDIB, F. CHARPILLET, J. P. HATON, «GREAT: A Model of Progressive Reasoning for Real-Time Systems», *in: Proceedings 6th IEEE International Conference on Tools with Artificial Intelligence*, C. Koutsou, Geras (éd.), IEEE Computer Society Press, New Orleans (Louisiane, USA), 1994.
- [54] A. I. MOUADDIB, F. CHARPILLET, J. P. HATON, «La conduite d'un raisonnement progressif pour la garantie de temps de réponse», *in: Actes 9ème Congrès Reconnaissance des Formes et Intelligence Artificielle AFCET, 2*, p. 669–675, Paris, janvier 1994.
- [55] A.-I. MOUADDIB, J.-MARTIN, P. ROBAUX, «Methode déductive pour les risques professionnels : système multi-agents», *in: Proceedings International Conference on Ergonomy and Industrial Applications, VI*, Toronto (Canada), August 1994.
- [56] F. MOURIA, Y. GONG, J.-P. HATON, «Un modèle phonétique tenant compte explicitement du contexte pour la reconnaissance de la parole continue», *in: Actes 9ème Congrès Reconnaissance des Formes et Intelligence Artificielle*, p. 265–275, Paris, janvier 1994.
- [57] A. NAPOLI, J. LIEBER, «A First Study on Case-Based Planning in Organic Synthesis», *in: Proceedings First European Workshop on Topics in Case-Based Reasoning*, S. Wess, K.-D. Althoff, M. Richter (éd.), Springer Verlag, p. 458–469, Kaiserslautern (Germany), 1994. Lecture Notes in Artificial Intelligence, 837.
- [58] A. NAPOLI, «Catégorisations, raisonnement par classification et raisonnement à partir de cas», *in: Journées Acquisition – Validation – Apprentissage (JAVA '94), Strasbourg*, p. E1–E14, Strasbourg, 1994.
- [59] A. NAPOLI, «Studies about the Integration of Classification-Based Reasoning and Object-Oriented Programming», *in: Working Notes of the 1994 Description Logics Workshop*, F. Baader, M. Lenzerini, W. Nutt, P. Patel-Schneider (éd.), DFKI Saarbrucken, p. 60–62, 1994.
- [60] N. PICAN, F. ALEXANDRE, «Highly Adaptative Neural Networks for Adaptative Neuro-Control: the OWE Architecture», *in: Proceedings IEEE SMC'94*, San Antonio (Texas, USA), October 1994.
- [61] N. PICAN, F. ALEXANDRE, «Weight Interpretation with Orthogonal Weight Estimator (OWE) Architecture», *in: Proceedings IMACS International Symposium on Signal Processing, Robotics and Neural Networks*, IEEE, p. 550–555, 1994.

- [62] N. PICAN, P. BRESSON, F. ALEXANDRE, J. P. HATON, E. COURIOT, «Une solution neuromimétique pour le pré-réglage d'un SkinPass adaptée au contexte d'utilisation: NeuroSkin adaptatif», *in: Actes Colloque FIRTECH*, p. 63-72, Nancy, Avril 1994.
- [63] N. PICAN, J. C. FORT, F. ALEXANDRE, «A Lateral Contribution Learning Algorithm for Multi MLP Architecture», *in: Proceedings European Symposium on Artificial Neural Networks*, M. Verleysen (réd.), D. Facto Publications, p. 229-234, Brussels (Belgium), 1994.
- [64] O. SIOHAN, Y. GONG, J.-P. HATON, «A Comparison of three Noisy Speech Recognition Approaches», *in: Proceedings International Conference on Spoken Language Processing*, Yokohama (Japan), 1994.
- [65] W.-C. TREURNIET, Y. GONG, «Noise Independent Speech Recognition for a Variety of Noise Types», *in: Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing, 1*, IEEE, p. 437-440, Adelaide (South Australia), April 1994.

Rapports de recherche et publications internes

- [66] M. BOUZID, P.-B. LADKIN, «Simple Reasoning with Time-Dependent Propositions», *Rapport interne*, CRIN-INRIA, Université Henri Poincaré, Nancy I, 1994.
- [67] M. BOUZID, A. LIGEZA, «Temporal Logic Based on Characteristic Functions», *Rapport interne*, CRIN-INRIA, Université Henri Poincaré, Nancy I, 1994.
- [68] A. BOYER, V. CHEVRIER, K. SMAÏLI, «De l'algorithmique fondamentale à l'algorithmique appliquée à la gestion des fichiers», *Rapport enseignement*, CRIN-INRIA, 1994.
- [69] D. FISCHER, M. KIRCH, A. NAPOLI, «ATOMOL = Tuteur Intelligent + Objets + Chimie», *Rapport interne*, CRIN-INRIA, 1994.
- [70] A.-I. MOUADDIB, «Knowledge-Based Anytime Computation», *Rapport interne*, CRIN-INRIA, 1994.
- [71] F. MOURIA, «Le corpus AGIMMO et une application pour le système VINICS», *Rapport interne*, CRIN-INRIA, Juin 1994.
- [72] G. SAON, A. BELAÏD, Y. GONG, «Off-line Handwriting Recognition by Statistical Correlation», *Rapport interne*, CRIN-INRIA, 1994.
- [73] S. VIALLE, T. CORNU, Y. LALLEMENT, «ParCeL-1, User's Guide and Reference Manual», *Rapport interne*, CRIN-INRIA, 1994.

9 Abstract

The SYCO group is centered around a single scientific theme, i.e. the study of perceptive and cognitive processes as well on fundamental basis as for practical applications. Its techniques and models are therefore mainly situated in the fields of artificial intelligence and pattern recognition.

The activities of the group are described by three key-words :

- interpretation of complex patterns like speech, industrial signals, documents, but also abstract patterns like situations, states, etc.. This interpretation implies a certain level of understanding that can only be achieved by exploiting multiple knowledge sources, especially in multi-agent architecture;
- handling multiple hypotheses, made necessary by the non-determinism of interpretation processes and the non-monotonicity of phenomena. This is an important aspect of the work which underlies research in neural net models, reason maintenance, abduction, etc.;
- real time operation, with implications on system architectures and on algorithms (interruption, guaranteed response time, etc.).

This year, noticeable achievements are in :

- neural networks : implementation of new cell models, test of the OWE model in steel industry, start of the ESPRIT BRA MIX project on hybrid connectionist/symbolic architectures;
- knowledge-based models : completion of the ESPRIT REAKT project toolkit for real time knowledge-based systems, advances in classification and case-based reasoning, real time reasoning and guaranteed response time, reason maintenance and abductive reasoning, planning for robotics;
- speech recognition : stochastic models (trajectory model and second order HMM), hybrid models for difficult vocabularies, noisy speech recognition, speaker verification, abductive/deductive model for continuous speech recognition.

Table des matières

1	Composition de l'équipe	1
2	Présentation du projet	3
3	Actions de recherche	6
3.1	Représentation des connaissances et raisonnement	6
3.1.1	Systèmes multi-agents	6
3.1.2	Maintien de vérité	9
3.1.3	Raisonnement temporel	10
3.1.4	Raisonnement temps réel	11
3.1.5	Planification et robotique	13
3.1.6	Formalisme à objets, classification et raisonne- ment à partir de cas	14
3.1.7	IA et santé	15
3.2	Modèles connexionnistes neuromimétiques	16
3.2.1	Inspiration biologique	17
3.2.2	Mise en œuvre de réseaux neuromimétiques	19
3.3	Reconnaissance de la parole	21
3.3.1	Aspects phonétiques	21
3.3.2	Modèles stochastiques	23
3.3.3	Modèles de langage pour la reconnaissance de la parole	25
3.3.4	Reconnaissance bruitée	26
3.3.5	Vérification et identification du locuteur :	29
4	Actions industrielles	30
4.1	Coopération de réseaux neuroniques	30
4.2	Etude connexionniste pour le contrôle de machines indus- trielles	31
4.3	Estimation de champ radioélectrique par réseaux neuro- niques	31

4.4	Interprétation de signaux radar	32
4.5	Ordonnancement de pistes par réseaux neuronaux	32
4.6	Interprétation de signaux sonar	33
4.7	Traitement informatique de la connaissance en chimie organique	33
4.8	Détection et interprétation de cris	33
4.9	Reconnaissance de la parole	34
4.10	Outils de raisonnement temps-réel	34
4.11	Système multi-agents distribué : DBB	34
4.12	Reconnaissance de la parole via le téléphone	35
5	Actions nationales et internationales	35
6	Diffusion des résultats	37
6.1	Diffusion des produits	37
6.2	Manifestations	37
6.3	Actions d'enseignement	38
7	Perspectives	38
8	Publications	39
9	Abstract	46