

DE

L'AFIA



Juillet 1996

 $N^{\circ}26$ 

# Présentation du bulletin

Le Bulletin de l'AFIA est le bulletin de l'Association Française pour l'Intelligence Artificielle. Il vise à fournir un cadre de discussion et d'échanges au sein de la communauté universitaire et industrielle. Ainsi, toutes les contributions, pour peu qu'elles aient un intérêt général pour l'ensemble des lecteurs, sont les bienvenues. En particulier, les annonces, les compte-rendus de conférences, les notes de lecture, les articles de débat sont particulièrement recherchés. Le Bulletin de l'AFIA publie également des dossiers plus substantiels sur différents thèmes liés à l'IA. Le comité de rédaction se réserve le droit de ne pas publier des contributions qu'il jugerait contraire à l'esprit du bulletin ou à sa politique éditoriale. De plus, les articles signés n'engagent que le point de vue de leurs auteurs.

### Pour contacter l'AFIA

#### Jean-Marc DAVID (Président)

RENAULT Sce Innovation et Projets d'Inf. Avancée service 12113, API QLG Q49 1 12 860 quai Stalingrad, 92109 Boulogne- Billancourt

Trésorerie Adhésions, Liens avec les adhérents René QUINIOU INRIA / IRISA Campus de Beaulieu

Parrainage de manifestations **Claudette SAYETTAT** Ecole des Mines

Centre SIMADE 158 cours Fauriel 42023 Saint Etienne Cedex 2

35042 Rennes Cedex

e-mail: quiniou@irisa.fr

### Personnes morales adhérentes à l'AFIA

AFCET, CEA-LIA, CISI, CNET PARIS, ELF-AQUITAINE, ELF-ANTAR, ENST, FRAMATOME, GIST, ILOG, INGENIA, INRETS, INRIA, INSTITUT FRANCAIS DU PETROLE, MATRA MARCONI SPACE, NEUROAGENT, PEUGEOT, RENAULT, SLIGOS, SOLLAC, Université Paris 9 Dauphine Bibliothèque.

### Bureau de l'AFIA

Jean-Marc DAVID, président Marc AYEL, vice-président René QUINIOU, trésorier Jean-François NICAUD, secrétaire

Jean-Paul BARTHES, Bertrand BRAUNSCHWEIG, JeanCHARLET, Françoise FOGELMAN SOULIE, Catherine GARBAY, Philippe JOR-RAND, Philippe LAUBLET, Eunika MERCIER-LAURENT, Marie-Christine ROUSSET, Gérard SABAH, Claudette SAYETTAT, Christian de SAINTE-MARIE, Brigitte TROUSSE, Manuel ZACKLAD

### Comité de rédaction

Philippe LAUBLET Rédacteur en chef et rubrique livres CAMS - Université Paris 4 96 boulevard Raspail, 75006 PARIS laublet@laforia.ibp.fr

**Monique BARON** «Sommaires des revues» LAFORIA, case 169 Université Paris 6 4, Place Jussieu, 75232 Paris cedex O5 baron@laforia.ibp.fr

Jean CHARLET rubrique «Annonces et comptes rendus de conférences» INSERM U194 - DIAM 91 boulevard de l'Hôpital 75634 Paris Cedex 13 charlet@biomath.jussieu.fr

Anne COLLINOT rubrique «Débat» LAFORIA, case 169 Université Paris 6 4, Place Jussieu, 75232 Paris cedex O5 collinot@laforia.ibp.fr

Alexis DROGOUL LAFORIA, case 169 Université Paris 6 4, Place Jussieu, 75232 Paris cedex O5 drogoul@laforia.ibp.fr

Gilles KASSEL rubriques «Présentation de laboraet «Présentation de sociétés» Univ. Technologique de Compiègne Centre de recherches BP 529, 60205 Compiègne Cedex gkassel@hds.univ-compiegne.fr

**Bernard LE ROUX** rubrique «Résumés de thèses» INGENIA, 52 Bd Rodin 92130 Issy les Moulineaux leroux\_b@ingenia.fr

Philippe VOLLE rubrique «IA et Internet» NON STANDARD LOGICS S.A. 57-59, rue Lhomond, 75005 Paris volle@nsl.fr

Réa. Prorata services - 45 35 94 14 Imp. Augustin Paris - 40.36.10.15

ISSN 1273-1323

### IS/SI: l'AFIA met fin aux négociations avec Montpellier et EC2.

Lors de sa réunion du 18 juin 1996, le Bureau de l'AFIA a voté à l'unanimité la motion suivante : "Compte tenu des divergences d'objectifs constatées entre l'AFIA et EC2D pour l'organisation d'une manifestation commune dénommée IS/SI 97, l'AFIA décide de mettre fin aux négociations avec EC2D, le Corum et le District de Montpellier".

Sans entrer dans le détail de discussions qui durent depuis plus de 2 ans maintenant, et ont généré un trafic très conséquent sur le réseau, il me paraît utile de rappeler brièvement ce que nous voulions faire et comment nous envisageons maintenant la suite.

L'objectif de l'AFIA était (et demeure) d'organiser annuellement une manifestation internationale d'IA - au sens large, dont elle assumerait la responsabilité sur la plan scientifique. Cette manifestation était connue sous le nom d'IS/SI (Intelligent Systems / Systèmes Intelligents). Par rapport aux manifestations existantes, et en particulier RFIA, IS/SI se distingue par son caractère international (au moins européen), et par sa volonté d'être un terrain d'échange privilégié entre la recherche et les applications de l'IA, Concrètement, 4 conférences étaient envisagées : 3 conférences centrées sur l'Intelligence Artificielle, les Réseaux de Neurones et le Traitement Informatique des Langues, la 4ème ayant pour vocation de présenter les success stories de ces trois domaines (on trouvera dans le Bulletin n° 23 de juillet dernier une présentation plus complète d'IS/SI).

L'AFIA ne souhaitant pas organiser seule une telle manifestation, nous avons, depuis 2 ans, envisagé différents montages avec la société EC2, et plus récemment la Technopole de Montpellier. Deux conventions ont successivement été signées, en 1994 et 1995, avec EC2; aucune n'a pu être appliquée du fait, par deux fois, de la défaillance d'EC2.

C'est parce qu'aujourd'hui le Bureau estime que les objectifs de nos partenaires (EC2, mais aussi le Corum et le District de Montpellier) ne sont pas compatibles avec ceux de l'AFIA que nous avons décidé de mettre fin aux discussions. Plutôt que d'organiser une manifestation dans laquelle l'AFIA ne se retrouvait pas complètement, en particulier sur le plan scientifique, nous avons préférer nous retirer et chercher à réaliser autrement notre projet, avec toutes les difficultés que cela présente.

En effet, le Bureau de l'AFIA a confirmé, lors de cette même réunion du 18 juin, l'objectif de voir déboucher ce projet, et si possible dès l'automne 1997. Vus les délais, l'édition 97 pourrait être réduite si nécessaire à une conférence sur les Réseaux de Neurones et une conférence sur les applications de l'IA.

Différents contacts ont déjà été pris ; vos suggestions sont évidemment les bienvenues.

Je remercie très chaleureusement les personnes qui ces derniers mois ont participé aux nombreuses discussions, sans ménager ni leur temps ni leurs efforts : Marc Ayel, Bertrand Braunschweig, Françoise Fogelman, Jean-Paul Krivine et Gérard Sabah.

Merci également à ceux qui avaient accepté de s'impliquer dans l'organisation de ces manifestations, par exemple en acceptant de participer aux comités de programme ; j'espère que nous pourrons bientôt les resolliciter à nouveau.

Jean-Marc David

## **URL Serveur AFIA:** http://www.irisa.fr/EXTERNE/afia/

### VIE DE L'AFIA

### Récompenses

Il nous parait intéressant de signaler les récompenses internationales obtenues par les membres de l'AFIA. N'hésitez pas à nous envoyer celles que vous pourriez connaître.

C'est donc avec plaisir que nous félicitons Marie-Christine ROUSSET pour un de ses deux papiers acceptés à AAAI-96 (Août 96). Ce papier écrit avec Alon LEVY, dont vous trouverez le résumé ci-dessous, a été selectionné comme un des trois "best paper award" de la conférence parmi plus de 650 papiers soumis.

Verification of Knowledge Bases based on **Containment Checking** 

Alon Levy (AT&T Laboratories) and Marie-**Christine Rousset (LRI, Universite Paris-Sud)** 

Building complex knowledge based applications requires encoding large amounts of domain knowledge. After acquiring knowledge from domain experts, much of the effort in building a knowledge base goes into verifying that the knowledge is encoded correctly.

A knowledge base is verified if for any correct set of inputs (i.e., problem instance), the knowledge base entails correct outputs. The correctness of the inputs and outputs is defined w.r.t. constraints that are known to hold in the domain.

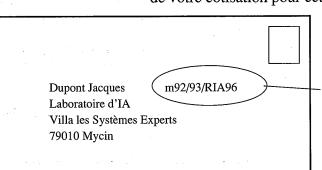
Our approach to the verification problem is based on showing a close relationship to the problem of query containment.

Our first contribution, based on this relationship, is presenting a thorough analysis of the decidability and complexity of the verification problem, for knowledge bases containing recursive rules and the interpreted predicates =, >=, =< and  $\setminus=$ .

Second, we show that important new classes of constraints on correct inputs and outputs can be expressed in a hybrid setting, in which a description logic class hierarchy is also considered, and we present the first complete algorithm for verifying such hybrid knowledge bases, thereby obtaining a method for handling this richer class of constraints.

### LE COIN DU TRÉSORIER

Les chiffres et lettres à droite de votre nom indiquent votre situation vis à vis de la trésorerie. Vous pouvez ainsi vérifier que vous êtes bien à jour de votre cotisation pour cette année.



- · Le code RIA indique une adhésion couplée avec l'abonnement à RIA.
- La lettre "m" indique une adhésion en tant que personne morale
- Les chiffres sont les deux derniers chiffres de l'année.

Ne jetez pas l'enveloppe d'expédition du Bulletin sans regarder de plus près l'étiquette portant votre nom et adresse...

### VIE DE L'AFIA

### **Groupe Terminologie et Intelligence Artificielle**

Animateurs: Didier Bourigault (EDF et CAMS), Anne Condamines (ERSS,

Didier.Bourigault@der.edfgdf.fr condamin@irit.irit.fr

Membres: B. Bachimont (AP-HP), B. Biébow (LIPN), J. Bouaud (AP-HP), J. Charlet (AP-HP), C. Enguehard (IRIN), B. Habert (ENS Fontenay-St Cloud), C. Jacquemin (IRIN), G. Otman (CTN), F. Rousselot (ERIC), J. Royauté (INIST), S. Szulman (LIPN), Y. Toussaint (INRIA, CNRS), P. Zweigenbaum (AP-HP).

Le groupe Terminologie et Intelligence Artificielle (TIA) est parrainé par l'AFIA depuis 1994 ; il est subventionné par le PRC-IA.

D'abord conçu comme un groupe de réflexion sur les possibilités de rencontre entre les méthodes et les objectifs de la terminologie et de l'IA, TIA s'est à présent structuré selon deux axes. Le premier de ces axes vise à travailler sur le concept de Base de Connaissances Terminologiques (BCT), c'est-à-dire à émettre des hypothèses semblant valides au regard des problématiques et des approches des diverses disciplines représentées et à les tester sur une expérimentation réelle. Le second de ces axes, qui s'est plus récemment défini, concerne l'animation de la communauté scientifique; il vise à créer des moments de rencontre entre intervenants des différentes disciplines concernées par la problématique « TIA »; les travaux et les réflexions du groupe TIA permettent alors de baliser ce terrain interdisciplinaire et d'articuler les diverses approches.

### I- TRAVAIL INTERNE AU **GROUPE**

Notre travail sera présenté suivant quatre points : Domaine et problématique, Axes de travail et objectifs, Travail déjà effectué et travail en cours, Plan de travail pour l'année 1996.

### 1. DOMAINE, PROBLEMATIQUE

Les relations mutuelles entre la terminologie et l'intelligence artificielle

constituent un problème fédérateur qui concerne au moins 4 disciplines : terminologie, linguistique sur corpus, traitement automatique de la langue naturelle et acquisition et représentation des connaissances en IA.

Ainsi que le montre leur appartenance, les membres de TIA sont bien représentatifs de ces disciplines.

Ce problème devient crucial aussi dans les entreprises au travers de questions qui se posent sur les mémoires d'entreprises (ou capitalisation du savoir-faire) et la gestion de documents. La plupart des membres de TIA entretiennent des relations étroites avec des entreprises : Matra Marconi Space, EDF, CNES, Aérospatiale, CNET, Assistance Publique-Hôpitaux de Paris, UAP, CEA, INIST.

Pour l'IA, la prise en compte de données terminologiques est un apport capital aussi bien pour l'acquisition que pour la représentation des connaissances dans un domaine donné. L'idée directrice qui anime le groupe vise à constituer une description de ce domaine à partir de l'analyse d'un corpus. Par hypothèse, cette description structurée, appelée base de terminologiques connaissances (BCT), constitue une source de connaissances à partir de laquelle on pourra construire différentes bases de connaissances pour des systèmes d'intelligence artificielle (systèmes à base de connaissances, SBC) de façon plus efficace que si l'on partait de zéro.

### 2. AXES DE TRAVAIL ET **OBJECTIFS**; **PROGRAMME DE** TRAVAIL:

### 1. Bases de Connaissances Terminologiques

Une réflexion théorique et appliquée s'est engagée pour tester l'hypothèse de faisabilité d'une BCT. Elle nous permet de définir les questions à résoudre et de commencer à donner des réponses.

1.1 Tester l'hypothèse de faisabilité d'une BCT

Deux éléments délimitent plus précisément cette hypothèse:

- La BCT peut être constituée à partir d'une analyse du corpus d'un domaine.
- La BCT doit permettre des utilisations multiples: elle doit donc respecter le critère très discuté en IA de réutilisabilité.

#### 1.2 Questions à résoudre :

- Quel format donner à cette BCT?
- Comment doit-elle être construite ?
- · Comment utilise-t-on une Base de Connaissances Terminologiques pour construire une Base de Connaissance?

#### 2. Programme de travail

Pour donner des réponses à ces questions, les orientations suivantes ont été définies. Réflexion:

- · Réflexion sur les définitions de « terme » de « concept » et de « relation » en terminologie et en IA.
- Etude des modèles de bases de données terminologiques.
- Etude des modèles de bases de connaissances en IA.
- Etude des méthodes d'exploration de corpus en linguistique, terminologie, informatique linguistique.
- Etude des méthodes de construction de terminologies.

#### Expérimentation:

- Choix d'un domaine et d'un corpus.
- Exploration de corpus : comparaison des méthodes en terminologie et en informatique linguistique.
- Exercice de début de construction d'un réseau notionnel : comparaison des approches et des méthodes.
- Identification et expérimentation d'outils d'exploration de corpus.

### 3. TRAVAIL DÉJÀ EFFECTUÉ, TRAVAIL EN COURS

### 1. Travail effectué

Les résultats des discussions ont permis d'affiner la notion de BCT et de donner des bases de discussions communes. Ils ont été présentés par D. Bourigault et A. Condamines lors des Journées du PRC IA à Nancy en 1995. Par ailleurs, la progression des travaux dans TIA a nourri les réflexions individuelles des membres qui ont publié des articles sur des thèmes voisins de ceux examinés dans TIA.

### VIE DE L'AFIA

#### 2. Travail en cours

Actuellement, notre effort porte plus précisément sur la question suivante : examen de méthodologies pour construire un réseau conceptuel. Nous avons défini deux types d'exercices sur le corpus sélectionné (maladies cardiovasculaires)

- · Construction d'un réseau conceptuel reliant une quinzaine de termes préalablement choisi.
- · Construction à partir d'un terme du réseau conceptuel dont il est le centre.

Ces exercices permettent à différentes équipes impliquées dans TIA d'évaluer, sur un corpus commun, l'utilisation possible d'outils tels que :

- outil d'extraction de connaissances terminologiques: Lexter et Ana;
- outils statistiques : Alceste et Lexico ;
- outils de type concordancier : Sato ; • outils de type étiqueteur : Olmes,
- outil permettant de prendre en compte la variation à l'intérieur des termes :
- · outil recherchant des marqueurs syntaxiques de relations conceptuelles : Mantex.

Les premiers résultats de ces travaux, qui sont encore en cours, ont été présentés lors d'une journée ATALA (Association pour le traitement Automatique de la LAngue) organisée conjointement par le projet européen POINTER et par le groupe TIA le 27 janvier 1996.

### 4. PLAN DE TRAVAIL POUR L'ANNÉE 1996

Le travail se poursuit sur le corpus. Les résultats de ce travail alimentent les réflexions sur différents points.

### 1. Expérimentation

Poursuite de l'exploration du corpus

- Poursuite de l'application de méthodes linguistiques, terminologiques, informatique linguistique et de l'expérimentation avec des outils existants.
- Etude de la complémentarité des méthodes manuelles et automatiques.
- Construction d'une (ébauche de) BCT sur la base du travail précédent.
- Utilisation de la BCT pour construire une base de connaissances de SBC.

#### 2. Travail de réflexion

- · Affinage du modèle de BCT retenu pour l'expérimentation.
- Méthode de construction de BCT.
- Poursuite des réflexions sur les notions de « terme », « concept » et « relation » alimentées par le travail sur

### II. ANIMATION DE LA COMMU-NAUTÉ SCIENTIFIQUE

Ce second axe est présenté suivant le même plan que le premier.

### 1. DOMAINE, PROBLEMATIQUE

IA, Terminologie, Linguistique, TALN constituent plusieurs communautés qui ne se connaissent pas suffisamment: l'objectif de TIA est d'aider à leur rapprochement. Grâce à l'expérience que nous acquérons sur l'interdisciplinarité à l'intérieur du groupe TIA, nous sommes à même d'organiser des manifestations de plus grande ampleur.

### 2. AXES DE TRAVAIL ET OBJEC-TIFS; PROGRAMME DE TRA-VAIL

Deux axes:

- Constituer un groupe de travail pluridisciplinaire, en nombre restreint afin de favoriser la réflexion.
- Organiser des journées de rencontre ouvertes.

Programme de travail:

cadres existants.

- Réunions bimestrielles du groupe de
- Organisation de journées scientifiques
- « Terminologie et IA ». • Organisation de rencontres dans des

### 3. TRAVAIL DÉJÀ EFFECTUÉ, TRAVAIL EN COURS

· Organisation des « Premières rencontres Terminologie et Intelligence Artificielle », avril 1995. Plus de 200 personnes étaient présentes, à peu près également réparties dans les trois communautés : linguistique/terminologie, informatique, industriels. Pour ces premières rencontres, la formule « conférences invitées » a été retenue, avec comme consigne à chaque conférencier d'essayer d'être accessible à chaque communauté. Cette consigne a été plus ou moins suivie selon les conférenciers. Le bilan de cette première tentative

pour fédérer plusieurs communautés autour des thèmes de TIA est largement positif, ne serait-ce que par le nombre de participants. Les discussions qui ont suivi les présentations mais aussi les réactions des participants, après les rencontres, ont confirmé le besoin de lieu de rencontres de ce type. Est apparue aussi, nettement, la nécessité de développer, lors de prochaines journées, les présentations de type didactique, permettant de donner à chaque participant les éléments de base pour comprendre les objectifs et les méthodes de chaque discipline. Ces remarques seront prises en compte pour les prochaines journées TIA, en 1997, qui s'organiseront en deux temps : des tutoriels et des communications sélectionnées par un Comité de programme.

- Coordination d'un numéro spécial de la Banque des Mots, tenant lieu d'actes des Rencontres de 1995 (responsable : Gabriel Otman), sorti en mai 1996.
- · Co-Organisation d'une journée ATALA (Association pour le Traitement Automatique de la LAngue), en collaboration avec le projet européen POINTER, le 27 janvier 1996 (responsable Pierre Zweigenbaum).

### 4. PLAN DE TRAVAIL POUR L'ANNÉE QUI VIENT

- · Co-organisation d'une journée de rencontre avec les groupes de travail PRC-IA Graphes Conceptuels et EVA.
- · Organisation des « Deuxièmes Rencontres Terminologie et Intelligence Artificielle », prévues pour les 3 et 4 avril 1997, à Toulouse (voir appel dans ce bulletin).



### **CONGRES AFIA**

### **APPEL A PARTICIPATION**

### **Troisièmes Rencontres Nationales des** Jeunes Chercheurs en Intelligence Artificielle.

Nantes, 28-31 Août 96

Les Troisièmes Rencontres Nationales des Jeunes Chercheurs en Intelligence Artificielle auront lieu à Nantes du 28 au 31 Août au matin inclus. Elles sont accueillies par l'IRIN, Institut de Recherche en Informatique de Nantes, dans les locaux de la Faculté des Sciences et des Techniques de la ville. Elles sont à ce jour parrainées par le CNRS, l'INRIA, le PRC-IA, la Ville de Nantes, l'Université de Nantes et l'Ecole doctorale "Sciences Pour l'Ingénieur" de Nantes.

Ces rencontres sont organisées tous les deux ans par l'AFIA. Leur but est de permettre à des étudiants préparant une thèse en Intelligence Artificielle, ou l'ayant soutenue depuis peu, de se rencontrer et de présenter leurs travaux. Mais elles offrent aussi aux auditeurs, universitaires ou non, un vaste panorama des recherches en cours dans le domaine.

### PROGRAMME PRELIMI-NAIRE (version courte)

Les posters seront affichés en permanence dans des salles attenantes à l'amphithéâtre. Les articles longs seront présentés en session plénière en 20mn, plus 5mn de questions. Chaque poster sera brièvement (1-3mn) présenté en session plénière. Puis son auteur se tiendra à disposition des participants dans la salle correspondante durant un temps spécialement réservé, de l'ordre de 30mn.

### MERCREDI 28 AOUT

9h - 10h15 : Accueil 10h15 - 10h45 : OUVERTURE DES RENCONTRES

10h45 - 12h : Représentation des connaissances

12h - 12h30: Posters "Représentation des connaissances"

14h - 15h45 : EIAO, Explication, Multi-agents

16h - 16h50: Déduction, Programmation logique I

16h50 - 17h 45 : Posters "EIAO, Explication, Multi-agents" et "Déduction, Programmation logique"

### JEUDI 29 AOUT

8h45 - 10h : Langage naturel 10h - 10h30 : Posters "Langage naturel et EIAO, Explication, Multi-agents"

10h45 - 12h30 : Déduction, Programmation logique II

14h - 15h20 : Conférencier invité Jean-Arcady MEYER (Ecole Nationale Supérieure, Ulm): "L'approche animat : simulation des comportements adaptatifs chez l'animal et le robot"

### VENDREDI 30 AOUT

8h45 - 10h: Planification, Diagnostic, Applications industrielles 10h - 10h30: Posters "Planification, Diagnostic, Applications industrielles" 10h45 - 12h : Apprentissage,

Acquisition des connaissances I 12h - 12h30 : Posters "Apprentissage, Acquisition des connaissances"

14h - 15h20 : Conférencier invité Daniel KAYSER (LIPN, Université Paris Nord): "L'Intelligence Artificielle est une

Science"

15h20 - 15h45 : Séance générale de 16h - 17h15: Contraintes, CSP 17h15 - 17h45 : Posters

### SAMEDI 31 AOUT

"Contraintes, CSP"

8h45 - 10h30 : Apprentissage, Acquisition des connaissances II 10h45 - 12h : Représentation des connaissances II 12h30: FIN DES RENCONTRES

Président du Comité de Programme : Torsten Schaub LERIA, Faculté des Sciences Université d'Angers 2, boulevard Lavoisier 49045 Angers Cedex 01

Conseiller scientifique senior: Laurent Siklossy Université de Savoie ESI/GEC/LIA Campus Scientifique Savoie Tech-73376 Le Bourget du Lac Cedex

Président du Comité d'Organisation Philippe LAMARRE IRIN. 2 rue de la Houssinière, 44072 Nantes CEDEX 03

### **RENSEIGNEMENTS -INSCRIPTIONS**

http://www.sciences.univnantes.fr/info/manifestations/ rjcia96/

e-mail: rjcia96.org@irin.univnantes.fr

### **DEBAT**

### L'intelligence artificielle et sa place dans les sciences de la cognition

Vincent Rialle Université Joseph FOURIER et CHU de Grenoble Laboratoire TIMC-IMAG URA CNRS 1618 Vincent.Rialle@imag.fr

#### I. Introduction

Quelle est la place de l'Intelligence Artificielle dans la mouvance des sciences de la cognition (SC) ? Une telle question inclut tout en la dépassant celle de la définition même de l'IA, mille fois rencontrée. Ce texte ne prétend pas donner une vue exhaustive de l'IA, vaste sujet qui se laisse difficilement "encapsuler" dans un article de synthèse, mais tente simplement de présenter de manière concise, donc nécessairement non exhaustive, et pédagogique le panorama des relations que l'IA entretient avec les SC.

La collusion entre les sciences de la cognition et l'intelligence artificielle est très étroite. Mais l'IA cognitive n'est pas toute l'IA: un nombre important de chercheurs ne se reconnaît pas dans un contexte de sciences cognitives. D'un autre côté, l'IA permet à un très grand nombre de thèmes de s'exprimer : de la quasi-totalité des disciplines des sciences de la cognition à la poésie, la théologie, la psychanalyse, etc. C'est dire la prégnance de l'IA non seulement dans la sphère scientifique mais aussi dans de multiples dimensions de l'expérience humaine actuelle. Commençons comme il se doit par un peu d'histoire.

### II. Une histoire enchâssée dans celle des sciences de la cognition

Pour montrer l'étroite collusion qui existe entre l'IA et les sciences de la cognition, rien de tel qu'un bond d'un demi-siècle en arrière, lorsque tout était encore en germe. Dans le creuset des années 40, et bien qu'une longue tradition scientifique remontant au moins à Aristote ait expliqué cette émergence, tout semblait naître en même temps : le concept de "machine universelle" avait été proposé par Turing en 1936 et 1937, qui n'aura de cesse d'étudier la possibilité théorique d'une machine intelli-

gente. Le célèbre "Test de Turing" scellera à sa manière l'une des plus étroites relations qui existent aujourd'hui entre IA et SC, puisqu'il n'ambitionne rien moins que de fournir le moyen de décider si une machine peut être considérée comme "intelligente" au regard de l'homme, c'est-à-dire dotée de toutes les capacités de raisonnement, de décision et, le test insiste là-dessus, de ruse habituellement reconnues à l'intelligence humaine. Von Neumann de son côté, à qui l'on doit la paternité (disputée à Eckert et Mauchly) de l'architecture générale de l'ordinateur, publiera son dernier ouvrage sous le titre "The Computer and the Brain". D'où la constatation que ces deux illustres pères fondateurs, à qui l'on attribue couramment une part importante de la paternité de l'informatique dans son ensemble, semblaient avoir pour préoccupations fondamentales la conception d'une machine possédant tous les attributs d'une intelligence humaine. Ils n'étaient déjà pas seuls et allaient être suivis d'un grand nombre : en 1943, W. McCulloch (psychiatre d'origine) et W. Pitts proposent le premier modèle mathématique de fonctionnement du neurone et montrent la capacité d'un réseau constitué de tels objets formels à reproduire certaines fonctions logiques.

Ouelques années plus tard (1949), le psychologue canadien D.O. Hebb introduit l'apprentissage par ajustement des poids synaptiques, notion qui connaîtra le succès que l'on sait sous le terme d'apprentissage hebbien ou règle de Hebb. Norbert Wiener de son côté apporte les fondements d'une théorie générale de la commande et du contrôle — la cybernétique —, dont on reconnaît aujourd'hui la portée et l'influence qu'elle eut sur l'IA. C'est la « première cybernétique », celle des années 40-50, qui semble-t-il, constitue le premier berceau de ce qui deviendra plus tard les sciences de la cognition (encadré 1).

« La cybernétique aura : introduit la conceptualisation et le formalisme logico-mathématique dans les sciences du cerveau et du système nerveux ; conçu organisation des machines à traiter l'information et jeté les fondements de intelligence artificielle ; produit la 'méta-science" des systèmes, laquelle a aissé son empreinte sur l'ensemble des sciences humaines et sociales, de la thérapie familiale à l'anthropologie culturelle en passant par l'économié, la théorie des jeux, la sociologie, les sciences du politique et bien d'autres, fourni à point nommé à plusieurs "révolutions scientifiques" du XX<sup>e</sup> siècle, très diverses puisqu'elles vont de la biologie moléculaire à la relecture de Freud par Lacan, les métaphores dont elles avaient besoin pour marquer leur rupture par rapport à un ordre ancien » ((Dupuis, 1985 ; 11 ; cité par (Pelissier et Tête, 1995 : XVIII-XIX)).

Comme chacun sait, la date de naissance de l'expression intelligence artificielle est celle du séminaire d'été tenu en 1956 au Dartmouth College, États-Unis. Mais ce qu'il faut remarquer ici est le profil de ses organisateurs, premier noyau dur de la discipline, qui signait déjà une forte relation entre informatique et psychologie expérimentale: M. Minsky (psychologue), A. Newell (psychologue), H. A. Simon (psychologue, économiste), J. McCarthy (informaticien), C. Shannon (mathématicien).

Soulignons cependant que chez ces pionniers de l'IA, l'intelligence était essentiellement celle des raisonnements logiques et des processus d'inférence heuristique. Nous concevons aujourd'hui plus clairement que l'intelligence n'est pas que cela : elle inclut les processus perceptifs, l'apprentissage, l'auto-organisation, l'adaptation... et souvent y culmine. Ces notions étaientelle concevables pour les pionniers de l'IA des années 40 ? Piaget, dès les années trente, en indiquait en tout cas la direction. La cybernétique était certainement sensible à ce point de vue, mais l'IA naissante n'en était pas là : l'intelligence à imiter était celle des "grosses têtes", savants, experts et logiciens; pas encore celle qui se développe jour après iour chez l'enfant, ou encore l'intelli-

### **DEBAT**

gence collective et distribuée de nombreuses colonies animales.

Ouelle qu'ait été l'histoire des conceptions de l'intelligence, tout s'est passé comme si dès les premières années, le projet de réaliser une machine telle que l'ordinateur avait, à côté de ses objectifs pratiques immédiats, celui de reproduire voire dépasser, l'intelligence humaine dans toutes ses dimensions. La suite sera une profusion de travaux théoriques et empiriques... et nombre de débats et de questionnements épistémo-philosophiques.

### III. L'Homme et son double, la Machine, ou le carrefour de l'IA et des SC

Pour continuer de montrer la collusion entre l'IA et les SC, il n'est pas inutile de refaire ces parcours, quelque peu classiques mais bons à rappeler, qui vont de l'homme à la machine et réciproquement.

L'ordinateur est insurpassable pour toutes les opérations purement calculatoires, mais l'opérateur humain reste difficile à remplacer par une machine pour l'accomplissement d'un certain nombre d'actions demandant beaucoup de souplesse (e.g., réponses à un guichet de renseignements, "fouille" dans des textes...), des tâches de spécialistes (application d'une expertise, traduction, interprétation, accomplissement d'actes complexes perceptifs et moteurs, ...) ou encore impliquant la vie, la justice ou l'éthique. Néanmoins certains objectifs, imposés notamment par l'essor technologique, ne peuvent être réalisés que par une machine suffisamment autonome : activités en milieux contaminés (nucléaires, biologiques, ...), en milieux inaccessibles (exploration/exploitation de planètes), activités de précision (e.g., robotique médicale), etc. Ces besoins exercent une pression permanente pour reculer les frontières de l'assumable par une machine. Mais la référence par excellence pour accomplir ces tâches reste la personne humaine, à travers notamment ses capacités de jugement en situation et de réaction face à l'imprévu (cf. le sauvetage d'Apollo 13). D'où l'importance prise ces dernières années par la notion d'ingénierie inverse, qui juxtapose deux questions:

1- Comment faisons-nous? à laquelle tentent de répondre les sciences de la cognition.

2- Comment faire faire par une machine? Dont la réponse est attendue d'une coopération entre diverses branches scientifiques et filières technologiques auxquelles concourent en priorité la robotique et l'IA.

Mais cette base pragmatique n'est pas neutre. Elle se déploie sur une toile de fond de déterminations biologiques, écologiques et politiques au sens fondamental des termes. L'explicitation de ces déterminations fait en partie l'objet des SC, notamment dans ses débats sur l'intentionnalité, la conscience et la subjectivité, et sur les rapports entre corps, cerveau, esprit, culture et champ social. Nous évoquerons, sans nous y attarder, ces débats par la voie des relations entre l'IA et la psychologie cognitive (chap. V), mais après un rappel au chapitre suivant, de l'opposition qui a prévalu il y a quelques années entre IA «classique» et connexionnisme.

### IV. IA «classique», connexionnisme, émergentisme...

Il a été de bon ton, durant la période qui a suivi le regain du connexionnisme au début des années 80, d'opposer l'IA dite classique et le connexionnisme, comme si l'une de ces deux branches d'investigation pouvait prétendre modéliser à elle seule l'ensemble des processus cognitifs humains. Le débat a été inévitable et fructueux, puisque l'on parle aujourd'hui beaucoup plus de systèmes complexes, hybridant les approches ou distribuant les tâches à des modules spécialisés, sans distinction d'obédience théorique. Il est repris ici en tant que l'un des débats animés qui a pris corps au sein des SC (e.g., (Posner, 1989)).

Le front des conflits tournait d'un côté autour de la rigidité du Système de Symbole Physique, et de l'autre côté autour de la difficulté posée par les réseaux de neurones à la compositionalité des symboles et par suite aux opérations symboliques un tant soit peu complexes.

Il résulte des meilleures analyses qu'aucune des deux approches ne saurait prétendre posséder les clés d'un modèle du

fonctionnement cognitif humain, et que les oppositions se résorbent en une différence de niveaux de résolution (Lacharité, 1994). Par ailleurs, la question de l'inclusion ou de l'exclusion des réseaux neuronaux du champ de l'IA peut être considérée comme une affaire d'école plutôt classée (encadré 2).

« (...) il est important de noter que d'un côté l'intelligence artificielle, dans sa version classique, ne comprend pas les réseaux neuronaux puisqu'elle s'est constituée en vue d'offrir une alternative à la cybernétique, mais que, d'un autre côté, dans sa version générique, l'intelligence artificielle comprend toutes les techniques susceptibles de doter les machines d'intelligence, et donc, en particulier les techniques promues par le connexionnisme. » (Ganascia, 1993 211).

La tombée en désuétude des oppositions entre IA symbolique et connexionnisme a été accélérée par l'avancée de la réflexion sur le concept d'intelligence. Au centre de cette réflexion se trouvent moins la notion d'intelligence, quelque peu indéfinissable et par trop générale, que celles, plus affinées, d'intentionnalité, d'interprétation, d'émergence, d'autonomie, de vie artificielle, de grammaire cognitive, etc. Autant de facettes qui semblent évoquer plus fidèlement les caractères de l'intelligence que l'on constate aussi bien dans la nature en général que dans les activités proprement humaines. Ces dernières tiennent dans les SC le haut du pavé, et le défi consiste à comprendre l'activité rationnelle de l'homme en tant que partie intégrante d'une dynamique biologique et pulsionnelle d'un côté, d'une dynamique sociale, culturelle et historique de l'autre côté. Ces dynamiques sont combinées, contraintes ou amplifiées par une capacité presque hypertrophiée chez l'homme : celle de produire en permanence du sens, du contre-sens ou de la croyance par le langage et l'activité représentationnelle. Vaste programme... mais qui présente l'avantage de relativiser et donc de ramener une certaine modestie au sein des débats ponctuels entre telle ou telle école.

### **DEBAT**

### V. L'IA et la psychologie cognitive

Dans l'introduction de son ouvrage collectif consacré au débat entre l'IA et la psychologie cognitive, J.-F. Le Ny, nomme intelligence artificielle cognitive, la part de l'IA qui s'intéresse à l'étude du fonctionnement cognitif humain (encadré 3):

« Appelons « cognitive » cette conception; elle peut a priori sembler plus agréable au psychologue. Mais elle n'est pas forcément plus confortable, ni pour le spécialiste de l'intelligence artificielle, ni, en définitive, pour nous. (...) Ce fonctionnement cognitif humain, encore faut-il d'abord le connaître et, cette fois-ci, au sens fort du terme. Nos collègues de l'intelligence artificielle cognitive, devenus par là nos cousins, comme je le disais en commençant, peuvent bien alors s'adresser aux spécialistes de la psychologie scientifique que nous sommes censés être, plutôt qu'à leur intuition subjective. Mais en quels termes répondrons-nous à leurs questions spécifiques ? » (Le Ny, 1993 : 13). encadré 3

Il y souligne aussi à juste titre que cette IA cognitive n'est pas forcément la plus efficace en termes d'ingénierie (e.g., le principe du réacteur est plus efficace que celui du battement d'aile pour construire un avion). De son côté, G. Tiberghien en trace succinctement le programme (encadré 4):

« (...) une partie seulement de l'IA s'intéresse aux phénomènes cognitifs et à la façon dont ceux-ci sont produits par le cerveau. L'IA « cognitiviste » est fondamentalement bionique : a) elle s'inspire de certaines propriétés symboliques ou sub-symboliques de la cognition humaine afin de produire des logiciels simulant des propriétés similaires dans des environnements de résolution de problèmes particuliers ; b) elle fournit également des formalismes qui sont des langages de description aptes à saisir certaines de prédicats, réseaux sémantiques, systèmes de production, langage objet, réseaux connexionnistes, systèmes hybrides) ; 🦠

c) enfin, l'IA offre des systèmes computationnels qui permettent de simuler, de façon précise, des modèles psychologiques ou neuropsychologiques ; cela oblige les chercheurs à définir des cahiers des charges rigoureux ce qui peut augmenter la portée heuristique de leurs tentatives de modélisation. » (Tiberghien, 1993).

encadré 4

Une phrase de S. Harnad est également significative de l'importance de l'IA pour le spécialiste en psychologie cognitive: «(...) en tant que psychologue cognitiviste, je m'intéresse à l'intelligence des machines uniquement dans la mesure où elle peut être considérée comme un modèle "réaliste" de l'intelligence humaine (ou de celle d'autres organismes)» (Harnad, 1994:

Il ne faut cependant pas renverser le mouvement : que l'intelligence des machines intéresse la psychologie cognitive ne fait pas de la modélisation psycho-cognitive une branche de l'IA, même si cette modélisation vise une simulation sur ordinateur. Le but premier de la psychologie cognitive n'est pas de fabriquer des machines cognitives, mais d'étudier la cognition humaine. Il reste néanmoins que le spécialiste de l'IA et celui de la psychologie cognitive ont en partage, jusqu'à un certain point, la cognition humaine : « Nous explorons le territoire des activités cognitives humaines (mémoire et connaissance, calcul et raisonnement, langage et communication, perception et action) pour déterminer dans quelle mesure les manifestations de certaines d'entre elles peuvent être réalisées par

une machine. » (Jorrand, 1995: 4). Le projet de l'IA considérée comme une branche des sciences de la cognition consiste donc à comprendre l'intelligence pour la reproduire sous forme de modèles implémentés dans des ordinateurs. Le but est clairement de produire de l'artificiel, comme le rappelle Ganascia: «Si l'on veut saisir pleinement ce que désigne l'IA, il faut bien voir que ce qui importe au premier chef

dans l'IA, ce n'est pas l'intelligence, mais c'est l'artificiel : l'intelligence artificielle est d'abord un artifice, une ruse au moyen de laquelle on prête une intelligence aux machines dans le but de les asservir à nos propres fins » (Ganascia, 1993 : 210). De plus, il n'est absolument pas nécessaire de reproduire les mécanismes de l'intelligence naturelle pour qu'un système interactif puisse être considéré comme "intelligent" (encadré 5).

« (...) si l'utilisateur d'un système intelligent a l'impression que celui-ci comprend une intention, qu'il manipule des significations, qu'il exprime des idées ou qu'il est conscient, il est indifférent que ces objets jouent ou non un rôle effectif dans les mécanismes implémentés dans la machine. Autrement dit, dans la mesure où un utilisateur pourra - lui - toujours attribuer des intentions, reconnaître des idées ou des significations dans le fonctionnement d'un système naturel ou artificiel, le seul problème qui se posera au concepteur d'un système intelligent sera que la machine manifeste ces propriétés au niveau phénoménal. Leguel est le seul qui importe à l'utilisateur humain, puisque c'est à ce niveau qu'il juge de l'intelligence et qu'il mobilise le schéma de prévisibilité qui s'applique à ce qui est intelligent en général ». (Jorion, 1990 : 12). encadré 5

Il demeure que, dès que l'on arrive au point c) du programme énoncé par G. Tiberghien (encadré 4), la relation entre l'IA et la psychologie cognitive débouche sur un problème central : celui de la causalité des représentations mentales, sur lequel Fodor notamment appuie sa théorie de la modularité de l'esprit et dont P. Engel en donne un résumé clair (encadré 6):

« (...) la psychologie cognitive postule que le comportement des organismes s'explique par des représentations mentales douées de contenus et causalement efficaces (des « inférences », des « plans », des « bases de données », des « représentations de connaissances »). Mais comment des représentations douées de contenus peuvent-elles être causalement effi-

### **DEBAT**

caces, s'il est vrai que, selon le postulat scientifique, toute causalité est physique? (...) La solution selon Fodor vient en deux temps. Premièrement, les attitudes propositionnelles sont des relations à des propositions, entendues au sens de symboles concrets « inscrits » dans le cerveau. (...) En d'autres termes, les états mentaux comme les croyances et les désirs, sont des états fonctionnels de l'organisme, qui se définissent par leurs relations fonctionnelles et causales à des stimuli, à des sorties comportementales, et à d'autres états mentaux (...). Deuxièmement on peut comprendre la relation causale entre des contenus sémantiques de représentation et des comportements si l'on recours à la métaphore de l'ordinateur, car les ordinateurs ont précisément la propriété de relier les propriétés causales de symboles à leurs propriétés sémantiques, par le biais de leur structure syntaxique. » (Engel, 1995: 53-54). encadré 6

Ainsi posée, cette problématique de la causalité des représentations mentales débouche sur le terrain de la philosophie de l'esprit. Mais avant d'y venir, il est nécessaire, pour comprendre certains débats animés aussi bien chez les spécialistes de l'IA qu'en philosophie et en SC, d'aborder le point de vue extrémiste de l'intelligence artificielle.

### VI. L'IA forte : de l'homo au robot sapiens

Outre les diverses tendances et positions théoriques ou épistémologiques qui constituent l'IA, on peut observer, tout au long de son histoire, deux pôles d'attraction antagonistes nommés, en particulier par J. Searle, "IA forte" et "IA modérée" (plutôt que "faible").

L'IA modérée est une position pragmatique qui consiste à ne se référer qu'à ce qui est de l'ordre du pratique et du réalisable dans un proche ou moyen avenir, et qui vise à la réalisation de tâches difficiles mais concrètes et de première importance sur les plans industriel, économique ou social (traduction automatique, "fouille" de textes, résolution de problèmes complexes dans divers domaines, systèmes de perceptionaction en temps réel, etc.).

L'IA forte de son côté offre une position extrême quant à la conception même de l'esprit, de la machine et finalement de l'homme. Actuellement, cette position se nourrit des meilleures avancées des neurosciences — celles notamment qui ouvrent des voies de recherche sur les liens entre raisonnements, décisions, émotions, et états corporels, ou celles qui explicitent le continuum entre cognition biologique et cognition sociale. Elle reprend ces thèmes pour tracer un programme d'élaboration complète d'un esprit dans une machine. L'IA forte n'hésite pas à esquisser une sorte de fantasme de réalité virtuelle absolue dont une version amusante, de type sentimentalo-humaniste, est ici citée à titre d'exemple : «The emotional states observed in higher animals and humans - curiosity, love, fear, hate, jealousy, etc. are an integral part of, and contribute to, social intelligence. Likewise, higher intelligence must also include as part of its repertoire social intelligence. From the point of view of human requirements, we need neither computers that hate, nor computers that are unduly afraid. We want computers that love computers that cherish and care for their human masters, and perhaps also for their machines colleagues». (Stonier, 1992: 169-170). Mais l'IA forte n'en reste pas là. Rem-

placer peu à peu des parties de notre cerveau et de notre corps par des éléments micro-électroniques, l'un des grands maîtres de l'IA l'envisage avec le plus grand sérieux (encadré 7).

« Mieux nous connaîtrons notre cerveau; plus nous pourrons l'améliorer (...). À mesure que ces inventions s'imposeront, nous les relierons à nos cerveaux, en insérant des milliers d'électrodes microscopiques dans le corps calleux, grand fuseau nerveux où circulent les données du cerveau. Toutes les parties du cerveau seront équipées de nouveaux accessoires. Enfin, nous remplacerons toutes les parties du corps et du cerveau et nous remédierons à tous les défauts et à toutes les blessures qui écourtent notre vie ». (Minsky, 1994 : 123). Par conséquent,

« Laisserons-nous la terre à des robots? Oui, mais ils seront nos enfants. Nous devons nos esprits à la vie et à la mort de toutes les créatures qui nous ont précédés dans la lutte nommée évolution. Tout cet effort ne doit pas se transformer en un immense gâchis : c'est notre mission: » (Minsky, 1994:126).

encadré 7

Ainsi, exempts du traumatisme de la naissance, grâce à l'accouchement sans douleur permis par le clonage robotique («un milliard de copies en moins d'un an » (Minsky, 1994: 124!), nos arrièresarrières-...- arrières petits-enfants, les robots-humains parfaits et immortels, qui conserveront dans leur prodigieuse mémoire la totalité des connaissances, réflexions philosophiques, poèmes, tragédies et partitions musicales ..., seront-ils à même de se demander ce qui a bien pu conduire ce malheureux petit homme du vingtième siècle à opérer cette curieuse mutation consistant à se transformer en robot immortel et parfait?

On retrouve, dans cette implosion radicale du questionnement ontologique, le thème de la transfiguration rencontré dans la métaphysique : l'homme, par son perfectionnement, accède à l'immortalité. Mais ici, l'ange qu'il devient est un robot. L'autre envolée de l'IA forte dans la métaphysique se fait par le thème du Golem : l'homme reste ici un homme biologique, mais il produit un double, le robot, censé au moins l'égaler en tous points.

### VII. L'IA et la philosophie de l'esprit : l'esprit est-il computationnel?

L'IA forte représente une curiosité qui peut tantôt amuser tantôt inquiéter, mais qui, en tout cas, se développe, comme tous les intégrismes, sur un arrêt de la réflexion. Il existe en revanche une problématique des plus fécondes autour de questions fondamentales telles que la conscience, l'interprétation, l'intentionnalité, etc., englobées sous le terme très générique de sciences de l'esprit. On parle aussi de philosophie de l'esprit, de philosophie cognitive (Rialle et Fisette, 1996) et de fonctionnalismes (l's est important. cf. INTELLECTICA nº 21 (2), 1995, richement consacré à ces thèmes). Dans cette veine, on assiste depuis quelques années à

dimensions des représentations sym-

boliques ou sub-symboliques (langage

### **D**EBAT

une floraison de livres et d'articles dont les titres juxtaposent le mot esprit au mot ordinateur ou à l'un de ses équivalents paradigmatiques, tel qu'intelligence artificielle, machine, "computation". La question récurrente qui s'y trouve débattue, sous diverses formes (elle est posée ici sous une forme simplifiée au maximum) est la suivante: l'esprit est-il computationnel? En clair: le cerveau, si l'on admet l'indissociabilité corps-cerveau-esprit, fonctionne-t-il dans ses diverses activités comme une machine de traitement de symboles ? Si la réponse est positive, deux conséquences s'imposent alors : d'une part, on aura réussi à percer l'énigme de l'apparente dualité corps-esprit, qui est aujourd'hui l'un des points les plus discutés en philosophie de l'esprit, avec toutes les tentatives de naturalisation de l'esprit; d'autre part, on saura construire une machine intelligente. Si a contrario la sphère cognitive n'était pas computationnelle, serait-il encore possible d'en rendre compte scientifiquement ? Mais la question recouvre plusieurs pièges, à commencer par celui du contenu du terme "scientifique".

Le jeu de la réflexion dans lequel nous entraîne la question de l'esprit computationnel nous conduit à une série d'interrogations "délicieuses", pour reprendre un terme d'Hofstadter, qui intéressent l'IA "philosophique", mais dont la place manque ici pour les aborder. Disons simplement que, dans l'hypothèse soutenue actuellement (Jackendoff, Dennett, etc) selon laquelle la conscience pourra être, tôt ou tard, expliquée — c'est-à-dire explicitée en des termes rationnels au moyen d'un langage de description approprié -, ne sera-t-elle pas du même coup programmable? Si notre conscience peut être révélée de manière entièrement rationnelle, c'est-à-dire sans recours à aucun moment de son explication à une forme de pensée magique ni à une boîte noire impénétrable, cette explication ne sera-t-elle pas du même coup réductible à un algorithme ? C'est une question à considérer comme malicieuse, comme toutes les questions faussement naïves, mais y pénétrer demanderait plus de développement.

#### VIII. Conclusion

De ce parcours "en diagonale" des relations entre l'IA et les SC, retenons que ces deux vastes domaines scientifiques sont historiquement et pratiquement indissociables. De plus, l'ordinateur constituant aujourd'hui l'aune ou la métaphore la plus employée pour comparer, opposer, jauger l'intelligence, l'IA se trouve d'emblée convoquée dans tous les débats philosophiques sur les capacités cognitives de l'homme. De fulgurants raccourcis sont pris au passage par quelques prophètes de l'IA forte, raccourcis qui ont néanmoins pour effet positif de stimuler, chez ceux qui ne souhaitent pas les emprunter, la réflexion critique et épistémologique. Mais à force de "grignoter" du terrain sur la part spécifiquement humaine de l'intelligence, on ne voit pas pourquoi nous n'arriverions pas à produire une machine qui illusionne jusqu'à un degré élevé, une machine qui apprenne comme un enfant, hésite, s'émerveille, explique ses hésitations, demande de l'aide, se met en colère lorsque les choses ne vont pas, etc. La question cependant est moins celle de savoir si cela est réalisable ou non, que celle de comprendre ce qui, dans nos fondements ou notre culture, pousse une frange non négligeable de nos contemporains à ce que ce soit une machine qui nous parle d'amour! Chemin faisant, l'IA rencontre, redécouvre et renouvelle plutôt qu'elle ne les "découvre" quelques problèmes fondamentaux : ceux notamment qui concernent la définition de l'intelligence et ses déterminations, la définition de l'esprit, de la conscience et de l'intentionnalité, la relation cerveau-esprit, la formation des concepts et des croyances, l'apprentissage de la langue et sa relation à l'intelligence, son usage en tant que médiateur dans les dialogues et dans des textes, etc. "Tout cela ne dépasse-t-il pas le cadre de l'IA ?", objectera-t-on. La réponse est bien sûr : oui. Ces questions dépassent en effet largement le champ de l'IA, mais aujourd'hui, tout le monde se retrouve "sur le pont" — neuropsychologues (Damasio, Crick, Edelman...), mathématiciens (Penrose...), philosophes (Dennett, Putnam, Searle...), etc. - pour percer les mystères des fonctions supérieures de l'homme. Tout un chacun sans exception se réfère à un moment ou à un autre à l'IA ou à l'ordinateur. Ce qui domine néanmoins, c'est que personne ne semble être vraiment mieux placé que les autres pour aborder ces questions et qu'une certaine pulsion épistémophilique pousse des gens de l'IA, comme dans

d'autres disciplines, à contribuer à cette recherche pour laquelle la multidisciplinarité constitue une condition de progrès.

Bibliographie succincte Dupuis, J.P. (1985). L'essor de la première cybernétique. Cahiers du CREA, 7. Engel, P. (1995). Introduction à la philosophie de l'esprit. La Découverte. Paris. Ganascia, J.-G. (1993). Approches du morcellage en apprentissage symbolique. In Le Ny J. F. (ed.) Intelligence Artificielle et Intelli gence Naturelle. Presses Universitaires de France. Paris. 209-230. Harnad, S. (1994). L'ancrage des symbole dans le monde analogique à l'aide de réseaux neuronaux : un modèle hybride. In Rialle, V et Payette, D. (ed.) Modèles de la cognition vers une science de l'esprit, série LEKTON Vol. IV, n°2. Département de Philosophie Université du Québec à Montréal. 65-83. Jorion, P. (1990). Principes des systèmes intelligents. Masson: Paris. Jorrand, P. (1995). On nous dit que l'IA est morte... Bulletin de l'AFIA (Association Française pour l'Intelligence Artificielle Lacharité, N. (1994). Le problème de l'archi tecture dans le débat entre l'approch connexionniste et les approches «classiques de la représentation. In Rialle, V. et Payette D. (ed.) Modèles de la cognition : vers une science de l'esprit. LEKTON, 4(2), 37-64. Le Moigne, J.L. (1986). Genèse de quelques nouvelles sciences : de l'intelligence artifi cielle aux sciences de la cognition. In Le Moigne, J. L. (ed.) Intelligence des mécaismes, Mécanismes de l'intelligence. Fayard Le Moigne, J.L. (1994). Joyèuse méditation épistémologique sur la modélisation symbolisante de l'intelligence. In Rialle, V. et Payette, D. (ed.) Modèles de la cognition : vers une science de l'esprit. LEKTON, 4(2), 115-138. Le Ny, J.F. (ed.), (1993). Intelligence Natu relle et Intelligence Artificielle, Presses Universitaires de France, Paris. Minsky, M. (1994). Laisserons-nous la terre à des robots ? Pour la Science, 206. 120-126. Pelissier, A., et Tête, A. (ed.), (1995) Sciences cognitives : textes fondateurs Presses Universitaires de France. Paris. Posner, M. (ed.), (1989). Foundation of Cognitive Science. MIT Press. Cambridge, Rialle, V., et Fisette, D. (ed.), (1996). Penser l'esprit : des sciences de la cognition à une philosophie cognitive. Presses Universitaires de Grenoble: Grenoble, France. Stonier, T. (1992). Beyond Information: the natural history of intelligence. Springer-Ver Tiberghien, G. (1993). Les sciences cogn tives : un nouveau programme scientifique ' In Sfez, L. (ed.) Dictionnaire critique de la communication. Presses Univ. de France

### PRESENTATION DES LABORATOIRES

### Activités en Intelligence Artificielle au Département d'Informatique et de Recherche Opérationnelle

Université de Montréal

### présentation adressée par Guy Lapalme

Le département d'informatique et de recherche opérationnelle (DIRO), fondé en 1966, regroupe actuellement environ 40 professeurs, 450 étudiants de premier cycle, 80 de second cycle et 80 doctorants. Il est un des plus grands départements d'informatique au Canada et il est le plus actif au Québec en termes d'activités de recherche. Ses membres sont regroupés en laboratoires rassemblant autour d'un même thème les énergies de professeurs, de chercheurs et d'étudiants. Le DIRO a emménagé en 1994 dans le pavillon André-Aisenstadt, un environnement moderne et fonctionnel qui abrite également d'autres départements et centres de recherche liés aux mathématiques et à l'informatique.

Le DIRO comprend trois laboratoires qui travaillent sur des sujets liés à l'intelligence artificielle, dont les activités sont décrites dans les pages qui suivent. Il compte également des laboratoires traitant des domaines suivants : protocoles de communication, architecture des ordinateurs, génie logiciel, biologie informatique, infographie, informatique théorique et quantique, parallélisme, vision et modélisation géométrique, transport, optimisation et simulation numérique.

### LABORATOIRE INCOGNITO



Le laboratoire Incognito (informatique cognitive) a été fondé en 1983 par les professeurs Paul Bratley, Guy Lapalme et Jean Vaucher afin d'étudier la structuration des connaissances, la programmation logique et fonctionnelle, les systèmes experts et d'autres aspects liés à l'intelligence artificielle. C'est un groupe informel sans structure administrative fixe, qui réunit actuellement les efforts de recherche de Michel Boyer, Guy Lapalme, Jacques Malenfant, Jian-Yun Nie et Jean Vaucher. Les membres d'Incognito

ont établi des collaborations suivies avec des professeurs et étudiants des départements de linguistique et philologie ainsi qu'avec des laboratoires étrangers : l'école des Mines de Nantes, le laboratoire CLIPS (Communication langagière et interaction personne-machine) de Grenoble et l'Institut de Recherche en Informatique de Toulouse.

Michel Boyer s'intéresse depuis maintenant plusieurs années à la programmation logique et ses extensions. Dans ce cadre s'inscrivent des travaux sur les grammaires d'unification, sur la méta-interprétation et maintenant sur l'étude des mécanismes de contraintes en Prolog. On y démontre l'intérêt-et la viabilité de l'ajout de contraintes quadratiques dans un Prolog qui possède déjà les contraintes linéaires.

Guy Lapalme travaille à la génération de texte, un problème complexe et important dans un contexte où les informations traitées par l'ordinateur et devant être communiquées à l'usager se complexifient. Dans ce cadre, il a créé, avec Richard Kittredge professeur au département de linguistique, le groupe, nommé SCRIPTUM (décrit plus loin) qui s'attaque à des domaines où l'information est stéréotypée ou bien structurée. Les projets comportent une étude linguistique du domaine, la mise au point d'une stratégie d'organisation de texte et l'implantation en Prolog d'un prototype opérationnel de générateur à partir des données "brutes". Cette partie implantation est une caractéristique importante des projets qui nous différencie de plusieurs autres où, trop souvent, on se contente d'une étude de faisabilité ou d'une implantation partielle de certains modules.

Jacques Malenfant s'intéresse à l'étude, la conception et à l'implantation de langages de programmation réflexifs dans le cadre du projet Réflex (décrit plus loin). La réflexivité dans les langages de programmation devient actuellement un domaine très important car elle donne au programmeur un modèle flexible lui permettant d'adapter le langage lui-même aux besoins sans cesse changeants de ses applications. La réflexivité en programmation donne un fondement global à des idées développées indépendamment dans plusieurs approches (programmation fonctionnelle, logique, par objets, etc.); elle appréhende ces phénomènes avec un bagage commun en facilitant le transfert d'expertise d'un domaine à l'autre, et solutionne des problèmes importants autant du point de vue des langages que du point de vue des applica-

Jian-Yun Nie travaille sur la recherche d'informations qui vise à extraire, parmi un ensemble volumineux de documents, ceux qui sont pertinents à une requête donnée. Sur le plan théorique, il s'intéresse à la modélisation des systèmes de recherche d'informations. L'objectif est de définir un cadre approprié qui permet de décrire tous les aspects impliqués dans une opération de recherche d'informations, tels que le contenu des documents, le besoin de l'utilisateur et le contexte de l'opération. Les logiques constituent les outils privilégiés dans cette tâche. Mais un des problèmes à résoudre est de traiter de l'incertitude dans un cadre logique, ce qui nécessite l'adjonction de la logique avec, par exemple, la théorie des probabilités. Actuellement, l'accent est mis sur l'utilisation d'un nouveau type de logique - la logique contrefactuelle, identifiée comme la plus adaptée à la recherche d'informations. Du côté pratique, il se concentre sur l'implantation d'un système de recherche d'informations basée

### PRESENTATION DES LABORATOIRES

sur la sémantique. Cette méthode de recherche d'informations repose sur la "compréhension" plus poussée des documents et des requêtes (des textes), d'une part, et d'autre part, sur l'utilisation des liens sémantiques entre les concepts (ou plus généralement, des connaissances).

Jean Vaucher travaille à l'intégration des techniques d'Intelligence Artificielle avec celles du Génie Logiciel, particulièrement au niveau de la structuration des connaissances imbriquées, de la déduction automatique et des interfaces personne-machine. D'un côté pratique, il est intéressé par les systèmes de conception (CASE) et par les systèmes experts. Du côté méthodologique, il expérimente avec la programmation par objets (depuis 25 ans) et la programmation logique (depuis 15 ans).

#### Projets spéciaux

Pour donner un cadre à une collaboration suivie entre eux et des chercheurs d'autres facultés ou universités, deux professeurs du laboratoire Incognito (G.Lapalme et J.Malenfant) ont ressenti la nécessité de former des sous-groupes de recherche à l'intérieur du laboratoire. Il s'agit des groupes SCRIPTUM et Réflex.

L'équipe Réflex du laboratoire Incognito s'intéresse à la conception et à l'implantation de langages de programmation réflexifs. La réflexivité en programmation consiste à permettre à un programme d'inspecter et de modifier l'état de sa propre exécution. Le but est de permettre à ce programme d'adapter son propre comportement aux conditions dynamiques de son exécution, ceci pour optimiser son code, expliquer son comportement (trace, "debugging"), ajouter des fonctionnalités au langage de base (gestion d'exception par exemple), etc. Un langage de programmation réflexif est un langage qui supporte la réflexivité en représentant ses programmes ainsi que leur état d'exécution à l'aide de ses propres procédures et structures de données. Ces procédures et structures de données sont rendues accessibles au programme de l'utilisateur à l'exécution. Actuellement, plusieurs langages de programmation offrent, au moins en partie, des capacités réflexives : les langages à objets Smalltalk-80 et CLOS, le langage de programmation logique Prolog, le langage fonctionnel Scheme, etc. Nos projets se développent selon trois axes de recherche principaux :

- Langage de programmation par objets réflexif.
- · Application des techniques de transformation de programmes basées sur la sémantique (évaluation partielle, interprétation abstraite) pour résoudre les problèmes d'efficacité des langages réflexifs.
- Etude des langages réflexifs par des approches théoriques (sémantique dénotationnelle, systèmes de réécriture, monades, etc.) à la fois pour en définir la sémantique et pour obtenir des méthodes d'implantation efficaces formellement démontrables.

Les projets de recherche en cours portent sur l'implantation d'un langage à prototypes (objets) réflexif réifiant les continuations comme objet de plein droit et sur l'utilisation des monades pour représenter la partie comportementale du langage.

### Scriptum<sup>1</sup>

Un aspect très important des interfaces personne-machine est la possibilité d'obtenir une génération automatique de textes en langue naturelle à partir des structures d'information conservées dans l'ordinateur. En effet, la complexité de celles-ci rend pratiquement inutilisable l'approche consistant à prévoir à l'avance toutes les possibilités d'interactions au moyen de textes préparés à l'avance. Il y a bien certaines possibilités de paramétrisation au moyen de "formats variables" mais elles restent très limitées et très en deçà des interactions qu'on serait en droit d'attendre de la part d'un système qu'on voudrait qualifier d'intelligent. La génération de texte couvre un très large éventail de problèmes intéressants : la représentation des connaissances en fonction des domaines et des applications, les problèmes conceptuels comme le choix des informations à exprimer, les problèmes linguistiques comme le découpage en paragraphes ou en phrases, la sélection des constructions et des mots, la syntaxe des phrases et la morphologie. Or ces aspects du traitement de la langue naturelle ont été peu étudiés en comparaison avec les problèmes d'analyse qui eux ont fait l'objet de multiples théories et méthodes. Nous nous concentrons sur les aspects linguistiques et informatiques en ne considérant pas les problèmes de représentation de connaissances dans leur ensemble. Nous visons surtout deux classes de domaines où l'information est stéréotypée ou bien structurée : dans la première classe, nous nous intéresserons à des situations stables ou évolutives décrites par des informations numériques souvent conservées sous forme de tables semblables à celles rencontrées dans les banques de données relationnelles. Les "procédures" forment la deuxième classe de domaines qui nous intéressent. On y trouve des instructions d'assemblage, de réparation ou d'entretien qui spécifient des suites d'actions à effectuer dans un monde de référence concret pour obtenir un résultat voulu. Ce qui caractérise ces types de textes est la présence d'informations "objectives" sous forme de nombres ou de relations entre les faits.

<sup>1</sup>Scriptum signifie "écrit" en latin ; les amateurs d'acrostiches pourront retrouver ce sigle dans "Système Capable de Reproduire Intelligemment des Phrases et des Textes à l'Université de Montréal" ou bien "Système de CRéation Informatique de Petits Textes à l'Université de Montréal".

#### Incognito par l'Internet

Cette section regroupe les différentes adresses électroniques des membres du laboratoire. Le laboratoire Incognito possède sa propre adresse de courrier électronique : cogito@IRO.UMontreal.CA à laquelle peuvent être adressées toutes les demandes qui ne relèvent pas d'un membre particulier.

Michel Boyer Guy Lapalme Jacques Malenfant Jian-Yun Nie Jean Vaucher Amina Arfi Nicolas Anquetil

bover@IRO.UMontreal.CA lanalme@IRO.UMontreal.CA malenfan@IRO.UMontreal.CA nie@IRO.UMontreal.CA vaucher@IRO.UMontreal.CA arfi@IRO.UMontreal.CA anquetil@IRO.UMontreal.CA

### PRESENTATION DES LABORATOIRES

Ming Bai Martin Brisebois Eliana Coelho François-Nicola Demers Massimo Fasciano Marco Jacques Katia Montero Gilles Roy Akila Sarirete Stéphane Somé

bai@IRO.UMontreal.CA briseboi@IRO.UMontreal.CA coelho@IRO.UMontreal.CA demers@IRO.UMontreal.CA fasciano@IRO.UMontreal.CA marcoj@IRO.UMontreal.CA montero@IRO.UMontreal.CA royg@IRO.UMontreal.CA sarirete@IRO.UMontreal.CA some@IRO.UMontreal.CA

#### Autres adresses électroniques

Le Département d'Informatique et de Recherche Opérationnelle de l'Université de Montréal dispose d'un serveur WWW (World Wide Web) (http://www.IRO.UMontreal. CA) et FTP (ftp.IRO.UMontreal.CA) sur lequel le laboratoire Incognito a deux adresses:

• une page WWW

(http://www.IRO.UMontreal.CA/labs/incognito)

• un serveur ftp autorisant les connexions anonymes (anonymous) qui permet d'accéder à certaines thèses récentes d'étudiants du laboratoire (en format PostScript) ou à des articles (ftp.IRO.UMontreal.CA:/pub/incognito).

### LABORATOIRE HÉRON



Fondé en 1986, le laboratoire vise à évaluer l'apport de l'intelligence artificielle pour les systèmes tutoriels et les systèmes de bases de données. Il fonctionne en coopération très étroite avec le laboratoire multimédia dirigé par Jan Gecsei, qui s'appuie sur de nombreux projets de développement de systèmes tutoriels.

HÉRON (Higher Educational Research ON tutoring systems) est un

groupe multi-disciplinaire qui englobe des chercheurs de plusieurs universités : l'Université de Montréal, l'Université du Québec à Montréal, Mc Gill University, Bishop's University, et l'Université du Québec à Chicoutimi.

Les systèmes tutoriels intelligents (STI) résultent d'une convergence des travaux et progrès réalisés en intelligence artificielle et éducation. Les orientations des travaux du groupe HÉRON concernent des projets relatifs à l'architecture des STI, l'étude du raisonnement, la modélisation de l'étudiant, les stratégies pédagogiques, l'apprentissage à partir de systèmes experts, la coopération entre agents intelli-

Fasciano, M., Lapalme, G. "Postgraphe: a system for the generation of statistical graphics and texte" Proceedings of the 8th International Workshop on Natural Language Generation (INLG-96) Herstmonceux, Sussex, UK, 13-15

Gagnon, M., Lapalme, G. "From conceptual time to linguistic time" Computational Linguistics, 22(1): 91-127, March 1996.

Kosseim, L., Lapalme, G. "Choosing rhetorical relations in instructional texts: the case of effects and guidances" 5th European Workshop on Natural Language Generation, pages 207-218. Faculty of Social and Behavioural Sciences, Rijks Universiteit Leiden, Leiden, The Netherlands, 20-23 May 1995.

Malenfant, J., Jacques, M., Demers, F.-N. "A tutorial on Behavioral Reflection and its Implementation" Proceedings of Reflection'96, 1995.

Nie, J.Y., Lepage, F., Brisebois, M. "Information retrieval as counterfactual" The Computer Journal, 38(8), pages 643-657, 1995.

gents, les techniques d'explication de pannes, ... Le groupe est, en fait, le premier groupe au Québec réunissant des personnes spécialisées en bases de données, intelligence artificielle, multimédia, télématique, sciences cognitives, médecine, éducation. Le groupe a également été reconnu sur le plan international puisqu'il a organisé les conférences internationales ITS-88 ITS-92 et ITS-96 sur les systèmes tutoriels intelligents.

#### Projets de recherche

Les projets de recherche entrepris au laboratoire HÉRON s'articulent autour des principaux aspects suivants :

- Modélisation contextuelle de l'étudiant : le modèle étudiant est une des parties les plus importantes d'un STI. Plus d'informations devraient être obtenues de ce modèle, en particulier les conditions d'apprentissage dans lesquelles l'étudiant a acquis ses connaissances. Ce modèle contextuel de l'étudiant devrait guider le processus tutoriel avec les styles d'ap-
- Utilisation de systèmes experts en systèmes tutoriels : il a déjà été démontré qu'il n'est pas facile d'utiliser un système expert pour construire un système tutoriel. Les connaissances dans un système expert sont trop compilées pour être facilement accessibles à un utilisateur non spécialisé. Nous utilisons ici des techniques pour décompiler les connaissances du système expert et les transformer en une forme acceptable
- Analyse et détection du raisonnement : pour adapter des explications à un étudiant en situation d'apprentissage il est important de pouvoir détecter les différents états intermédiaires de raisonnement. Nous utilisons une expression iconique permettant de symboliser les étapes pour aboutir à la solution, ceci pour représenter le processus de raisonnement et adapter la connaissance à délivrer dans chaque cas.

# BULLETIN DE L'AFIA —

### PRESENTATION DES LABORATOIRES

- · Apprentissage de la maintenance : les processus de montage ou d'entretien de pièces mécaniques ou électriques dans des installations complexes doivent être assurés par du personnel fiable, possédant un bonne formation technique, sous peine d'introduire plus de pannes que prévues dans un fonctionnement normal. Notre approche vise à détecter les degrés d'apprentissage et de connaissance des intervenants de manière à assurer des interventions efficaces sur la maintenance de matériel en particulier dans l'aéronautique.
- Modélisation des connaissances : il s'agit ici de représenter les connaissances à l'intérieur d'un STI de manière à permettre une transmission flexible selon l'usager, son contexte d'apprentissage et son environnement de fonctionnement. La coopération entre les différents types de connaissances est basée sur les principes d'agents intelligents.
- Système tutoriel réparti : il s'agit de permettre à un STI, réparti géographiquement, de s'adapter à la fois en ce qui concerne le curriculum nécessaire à l'enseignement local et le modèle étudiant concerné.

### Activités réalisées par les chercheurs

Claude Frasson s'intéresse à différents aspects des systèmes tutoriels intelligents (STI): l'architecture des STI et les différents composants qui peuvent être réunis, les stratégies d'apprentissage et leur combinaison, la modélisation des connaissances de l'apprenant, l'analyse de son raisonnement lors de la résolution d'une tâche. Il étudie en particulier comment composer les modules logiciels intervenant dans divers types de STI de manière à construire des STI efficaces et réalisables. Les formes tutorielles intervenant dans ces STI vont du système tutoriel directif au système permettant à l'apprenant d'acquérir des connaissances par coopération avec d'autres apprenants dans le milieu de formation. Les modes tutoriels associés concernent le conseiller, le critique, le démonstrateur...

Esma Aimeur s'intéresse aux questions d'acquisition et de modélisation de connaissances en vue de leur utilisation dans les systèmes tutoriels intelligents. Ses domaines de recherche concernent, au sein du projet SAFARI, la modélisation de tâches chez l'apprenant, le développement de nouvelles stratégies tutorielles et la prise en compte de contextes d'apprentissage. Elle s'intéresse aux techniques d'explicitation des connaissances (directes et indirectes) et plus précisément à l'analyse de protocoles et à l'observation directe afin de construire un modèle de l'expert et un modèle de l'apprenant.

Jan Gecsei s'intéresse aux applications multimédia pour les systèmes tutoriels intelligents (outils d'apprentissage, d'édition, de communication). Dans un système réparti supportant des services multimédia tels que collaboration à distance, télé-enseignement ou accès aux bases de données réparties, la notion de la qualité de service (QS) est d'une importance particulière. Il participe à un projet visant la gestion cohérente de la QS à travers tous les niveaux architecturaux du système. Un aspect particulièrement important est l'interface usager permettant d'influencer cette gestion. Le projet envisage également la mise en fonction d'une base de données distribuée. Dans le domaine des STI, ses recherches

portent surtout sur les façons innovatrices d'utiliser des éléments multimédia dans les interfaces usager STI. Tout ceci dans le but d'augmenter l'efficacité et l'acceptabilité industrielle de ce type de STI.

Marc Kaltenbach s'intéresse aux différentes perspectives de raisonnement et d'interaction d'un usager en situation de travail avec son environnement. Il vise également à déterminer les structures de connaissances adéquates pour faciliter l'acquisition de connaissances. Il s'intéresse aussi à la création d'interfaces qui supportent le raisonnement dans le cadre de la résolution de problèmes. Les questions abordées sont : comment représenter des informations de nature concrète ou abstraite, évoluant dans le temps (par suite de choix d'un utilisateur du système) ? Comment intégrer diverses sources de connaissances dans le cadre de systèmes tutoriels intelligents?

#### Contacts nationaux et internationaux

Au Canada, le laboratoire est en liaison avec les équipes suivantes: laboratoire ARIES (Gordon McCalla, Saskatchewan university), Alberta Research Council (Marlène Jones), Simon Fraser university (Phil Winne, Linda Harassim), Ontario Institute for Studies in Education (Barbara Wasson ). Aux Etats-Unis, le laboratoire est en liaison avec Elliot Soloway (University of Michigan), Beverly Woolf (University of Massachussett), Alan Lesgold et Kurt Van-Lehn (Learning Research and Development Center, Pittsburgh), William Clancey (Institute for Research on Learning, Palo-Alto). En Europe, Héron est en liaison avec Guy Gouardères (laboratoire API, Apprentissage, pédagogie et Intelligence artificielle, de Toulouse), Brigitte de la Passardière (MASI, université Paris 6) et le Laforia (université Paris 6).

#### Projets spéciaux

Le groupe travaille au sein d'un vaste projet (SAFARI) (Système d'Aide à la Formation par l'Analyse de Raisonnement Interactif) associant des entreprises et visant à l'élaboration de systèmes tutoriels génériques de complexité variable adaptés à la formation en entreprises. Le groupe SAFARI est formé des personnes suivantes :

### Chercheurs Claude Frasson Esma Aimeur Gilles Gauthier Ian Gecsei Gilles Imbeau Marc Kaltenbach Susanne Lajoie

Bernard Lefebvre

**Professionnels** Anatole Kengn Carmen Alexe Imed Jebel

Chercheurs post-doctoraux Jean-Yves Djamen Thierry Mengelle

#### Personnel de soutien Martine Gemme

### PRESENTATION DES LABORATOIRES

Étudiants Ph.D Tang Ho Le Daniel Leibu Robert Azevedo Robert Bouchard Daniela Lupascu Rubiao Guo Antoine Mghayar Nadjib Rabia Adil Kabbaj Abdallah Serroud Tang Ho Le André Mayers Wandreile De Sousa Steve Munsie Hichem Talbi Roger Nkambou Wei Gin Gu Marie-Chantale Beaulieu Joan Rosca Serge Tadie Charles De Lean Francis Demers

Étudiants maîtrise Laurent Duperval Boumalha Hassiba Imed Jebel

Claude Dufour Michel Lalonde Louis Martin

Jean-Yves Rossignol

Le laboratoire intervient de façon plus large à l'intérieur du GRITI (Groupe de Recherche Inter-universitaire en Tutoriels Intelligents) qui regroupe environ 70 personnes de 5 universités (Université de Montréal, UQAM, McGill, UQAC, Bishop's) et détient une expertise dans le domaine de la formation en entreprise.

#### Liens avec les entreprises

Le groupe travaille depuis plusieurs années avec des entreprises dans le cadre de recherches conjointes. Les principales entreprises concernées sont : Bell, BNR, Northern Telecom, DMR, CSST, Virtual Prototype, Novasys

### Héron par l'internet

Deux adresses de courrier électronique labmgh@iro.umontreal.ca et safari@iro.umontreal.ca sont disponibles pour joindre un membre du laboratoire. Le laboratoire possède aussi deux adresses sur le serveur WWW du département d'informatique http://www.iro.umontreal.ca et

http://www.iro.umontreal.ca/labs/its/labs.html. ailleurs, l'adresse ftp du laboratoire est : ftp.iro.umontreal.ca/pub/its.

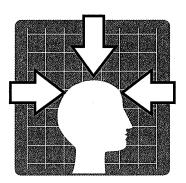
Aimeur, E., Frasson, C. "Analysing a New Learning Strategy according to different knowledge levels" Computer and Education, An International Journal, 1996 (à paraître). Frasson, C., Mengelle, T., Aimeur, E., Gouardères, G. "Ar Actor-based Architecture for Intelligent Tutoring Systems' ITS'96 Conference, Montréal, June 1996. Frasson, C., Kaltenbach, M. "Strengthening the Novice-

Expert shift using the self-explanation effect" Journal of Artificial Intelligence in Education, Special issue on student modelling, vol 3(4), pp. 477-494, 1993.

Gecsei, J., Frasson, C. "SAFARI: an Environment for Creating Tutoring Systems in Industrial Training" EdMedia, World Conference on Educational Multimedia and Hypermedia, Vancouver, 25-30 June 1994.

Gecsei, J. "Adaptation in Distributed Multimedia Systems" to appear in IEEE Multimedia 1997.

### LABORATOIRE D'INFORMATIQUE DES SYSTÈMES ADAPTATIFS (LISA)



Le LISA a été créé en 1993 pour regrouper les efforts de profesquatre seurs dont les travaux de recherche portent sur les systèmes adaptatifs. L'apprentissage automatique dans les sys-

tèmes adaptatifs est un domaine de recherche à l'intersection de l'intelligence artificielle, de l'inférence statistique et de l'optimisation. Il s'agit de construire des systèmes capables d'apprendre à effectuer une tâche à partir d'exemples et de connaissances a priori sur la tâche. Ces systèmes ont aussi souvent l'avantage de pouvoir être adaptatifs, donc de s'adapter à des changements dans leur environnement. Plusieurs des algorithmes utilisés proviennent de la recherche en réseaux de neurones artificiels, et sont inspirés de la manière dont le cerveau traite l'information. Un des avantages de cette approche est le traitement parallèle de l'information, qui se prête bien à des réalisations efficaces en matériel.

L'apprentissage automatique est important quand on n'a pas suffisamment de connaissances explicites sur un problème pour pouvoir directement écrire un programme ou un ensemble de règles permettant de résoudre ce problème. Dans certains cas, on peut cependant obtenir plus d'information sur le problème à résoudre sous la forme (implicite) d'exemples d'entrées et de sorties. Ces exemples nous donnent, pour un certain nombre de situations connues, la réponse qu'on désire que le système produise. Il s'agit ensuite de combiner les connaissances explicites sur le problème (dans le design du système et de la représentation) avec les connaissances implicites (sous la forme d'exemples).

Les professeurs du groupe concentrent leur recherche sur les algorithmes d'apprentissage, les questions reliées à leur réalisation physique, et/ou leur application à des problèmes de prise de décision et optimisation, de reconnaissance des formes, et d'analyse de séquences.

Yoshua Bengio s'intéresse aux algorithmes d'apprentissage pour la reconnaissance des formes et l'analyse de séquences. En particulier, il travaille en collaboration avec les laboratoires AT&T Bell du New-Jersey sur un projet d'analyse de documents combinant réseaux de neurones et modèles de Markov. Il applique aussi les algorithmes d'apprentissage à la prédiction et la prise de décision sur des séries chronologiques. Du point de vue

### PRESENTATION DES LABORATOIRES

plus fondamental, sa recherche se concentre sur l'apprentissage du contexte à long terme dans les systèmes adaptatifs (tels que les réseaux de neurones récurrents et les modèles de Markov).

Jocelyn Cloutier s'intéresse à deux aspects de l'apprentissage automatique. Le premier est l'apprentissage de stratégies et plus particulièrement, de stratégies dans les jeux. Un projet en cours vise l'apprentissage automatique d'une fonction heuristique évaluant les configurations du jeu othello. La deuxième direction explorée concerne les problèmes rencontrés dans la réalisation, sur une puce de silicium, de réseaux de neurones. Ces problèmes vont de la détermination de la précision nécessaire aux calculs, à l'évaluation de la tolérance aux pannes d'une puce "neurale" jusqu'à la réalisation d'un simulateur parallèle!

Jan Gecsei s'intéresse aux applications des réseaux de neurones dans des contextes inhabituels. Par exemple, dans le projet Safari (formation industrielle), il faut construire des systèmes capables de s'adapter aux préférences et autres particularités de l'étudiant (usager). Il peut s'agir d'une adaptation supervisée (l'usager demande explicitement au système de corriger certains de ses comportements) ou bien non-supervisée (en observant les interactions de l'usager avec le système de formation).

Jean-Yves Potvin est associé au Centre de Recherche sur les Transports et s'intéresse à l'intégration des techniques propres à la recherche opérationnelle, aux systèmes évolutifs et aux systèmes d'apprentissage automatique, afin de résoudre des problèmes complexes d'optimisation combinatoire, surtout dans le domaine des transports. En particulier, il s'intéresse à l'utilisation des réseaux de neurones, des algorithmes génétiques et de la programmation génétique dans le but de générer des routes optimales pour des véhicules de transport (e.g.cueillette et livraison de courrier "express", distribution de biens à des clients à partir d'un dépôt central, etc.).

Le LISA est situé au local 3332 du Pavillon André-Aisenstadt, avec plusieurs stations de travail SUN. Pour plus d'information, veuillez contacter Yoshua Bengio (bengioy@iro.umontreal.ca, ou 514-343-6804).

Bengio, Y. "Neural Networks for Speech and Sequence Recognition" International Thomson Computer Press, 1996.

Bengio, Y., Frasconi, P. "Diffusion of Context and Credit Information in Markovian Models" Journal of Artificial Intelligence Research, vol 3, p. 223-244, 1995.

Bengio, S., Bengio, Y., Cloutier, J. "On the Search for new learning rules for ANNs", Neural Processing Letters, Vol 2, no 4, p. 1-5.

### Présentations de laboratoires dans le bulletin de l'AFIA

LIFIA, Grenoble	Bulletin n°1
LRI, Orsay	Bulletin n°1
Service Systèmes Experts, Renault	Bulletin n°1
CEDIAG, Bull	Bulletin n°2
CERT, ONERA, Toulouse	Bulletin n°2
IRIT, Toulouse	Bulletin n°2
LAAS, Toulouse	Bulletin n°2
HEUDIASYC, UTC	Bulletin n°3
IFP, Rueil Malmaison	Bulletin n°3
DIAM, INSERM U194	Bulletin n°3
Lab. Math. Info., Fac Médecine de Marseille	Bulletin n°4
GMD, St. Augustin (RFA)	Bulletin n°4
ONERA, Chatillon	Bulletin n°4
KSL, Université de Stanford (USA)	Bulletin n°5
Dépt Applications de l'IA au CNET, Lannion	Bulletin n°5
LAFORIA, Univ. Pierre et Marie Curie	Bulletin n°6
L'institut FAW, ULM (RFA)	Bulletin n°6
Institut IIIA, Compiègne	Bulletin n°6
LAIR, OHIO State University (USA)	Bulletin n°7
ARAMIIHS, Labo mixte MATRA-CNRS,	Bulletin n°7
CEA, Service SERMA, Saclay	Bulletin n°8
Société ILOG	Bulletin n°8
LAIAC, Université de Caen	Bulletin n°9
Institut Français du Pétrole	Bulletin n°10
DFKI (Centre allemand de recherches en IA)	Bulletin n°11
GRTC, Marseille	Bulletin n°11
Inst. d'Analyse des Systèmes, Ac. Russe	Bulletin n°12
Georges Mason Univ., Center for AI (USA)	Bulletin n°13
IRISA, INRIA et Université de Rennes	Bulletin n°13
Société INGENIA	Bulletin n°14
LIPN, Université de Paris Nord	Bulletin n°14
Insitut EURISCO	Bulletin n°15
LRDC, Université de Pittsburgh (USA)	Bulletin n°15
Société ISOFT	Bulletin n°16
Dépt. d'Info de l'Université d'Ottawa	Bulletin n°16
Equipe CHM, Université du Colorado (USA)	Bulletin n°17
LIRMM, Montpellier	70 11 .1 010
Dikinin, montpoiner	Bulletin n°19
Institut autrichien de recherches en I.A.	Bulletin n°19 Bulletin n°20
Institut autrichien de recherches en I.A.	Bulletin n°20 Bulletin n°21 Bulletin n°22
Institut autrichien de recherches en I.A. ENST Bretagne	Bulletin n°20 Bulletin n°21
Institut autrichien de recherches en I.A. ENST Bretagne LIA - Université de Savoie	Bulletin n°20 Bulletin n°21 Bulletin n°22

### I.A. ET BIOLOGIE MOLECULAIRE

### Informatique, intelligence artificielle, et génomes

### Dossier réalisé par François Rechenmann

Les volumes de données et de connaissances produites et manipulées, la découverte constante de nouveaux objets et de nouveaux mécanismes, l'impérieuse nécessité de recouper et d'organiser des informations issues de sources dispersées et spécialisées, ont conduit la recherche en biologie, et plus particulièrement au niveau moléculaire, à faire appel très tôt aux outils et aux méthodes de l'informatique.

À l'inverse, l'informatique trouve dans la biologie moléculaire un domaine d'expérimentation extrêmement riche, en développement constant, et peuplé d'objets formels familiers que sont par exemple les séquences génomiques et protéiques, mais aussi les arbres phylogéné-

Ces dernières années ont ainsi vu émerger des domaines de recherche originaux, dans lesquels la biologie moléculaire est simultanément domaine d'application de méthodes et d'outils disponibles, d'expérimentation et d'évaluation de méthodes et d'outils originaux, et source d'enthousiasmants sujets de recherche en informatique.

Grâce à diverses actions incitatives et de formation (en particulier le GdR CNRS 1029 "Informatique et génomes" et le GIP GREG - "Groupement de Recherches et d'Études sur les Génomes"), une communauté de recherche bi-disciplinaire s'est ainsi constituée. Les chercheurs de cette communauté ont investi de leur temps et de leur énergie à comprendre la problématique biologique et à y identifier des sujets de recherche pertinents à la fois du point de vue de l'informatique et de celui de la biologie moléculaire. Ce dossier présente une partie de ces efforts de recherche, au moment où ces associations entre équipes d'informaticiens et de biologistes commencent à produire leurs premiers résultats.

#### Le génome

Le résultat le plus spectaculaire de la biologie moléculaire a été la mise en évidence d'une structure des organismes vivants dédiée au stockage d'information : le génome. Le génome n'a pas de rôle fonctionnel, il ne réalise pas par lui-même quelque fonction enzymatique complexe, pas plus qu'il ne contribue à la structure spatiale de la cellule. Par contre, il contient l'information nécessaire à la fabrication des molécules qui vont, elles, agir, constituer des membranes,

Le "cœur" du génome est une molécule linéaire constituée de l'enchaînement de quatre corps chimiques, appeles bases et symbolisés par les lettres A, C, G et T (respectivement pour adénine, cytosine, guanine et thymine). En première approximation, l'information génétique peut donc se représenter par un mot très long écrit dans un alphabet de quatre lettres.

Ce niveau d'approximation suffit à la plupart des études sur le génome, mais il est essentiel de garder en mémoire que ce qui va suivre n'est pas l'énoncé de lois de type physique. La biologie est un monde dans lequel toute règle a ses exceptions. Cette présence constante d'exceptions résulte, pour une part au moins, d'un caractère fondamental du monde vivant : il résulte d'un très long processus évolutif (environ 3 milliards d'années). Au cours du temps, face en particulier à des modifications de l'environnement, des solutions ont émergé à partir desquelles d'autres constructions se sont réalisées. Ainsi de multiples trajectoires étaient possibles, mais une seule s'est réalisée. Il n'y a aucun argument pour soutenir l'idée que la vraie trajectoire possède des propriétés particulières, d'optimalité par exemple, par rapport à l'infinité des trajectoires qui étaient possibles. D'une certaine façon, ce manque d'optimalité se traduit par l'absence de lois et son remplacement par une notion de solution consensus avec exceptions.

Le simple examen de la taille des génomes des différents organismes vivants met en évidence cette apparence irrationnelle de l'information génétique. Un raisonnement naïf peut faire croire que plus un organisme est complexe plus la quantité d'information qui lui est nécessaire sera grande. On s'attend donc à trouver de grands génomes de mammifères, des génomes plus petits chez les grenouilles, tritons et autres organismes à sang froid, encore plus petits chez les insectes et minuscules chez les bactéries et autres microbes. Si le dernier point est en général vérifié (ne pas oublier les exceptions), par contre on trouve par exemple des tritons ayant un génome beaucoup plus grand que celui de l'Homme. Les plantes fournissent également de nombreux contre-exemples à une relation monotone entre quantité d'information et complexité. Le plus célèbre est le génome minuscule de l'arabette (100 fois plus petit que celui du maïs) qui a fait de cette plante l'objet d'études favori des explorateurs des génomes végétaux. L'examen de régions pour lesquelles le décryptage des séquences a été réalisé va fournir la piste de résolution de cette contradic-

### Le gène

Le gène est un concept essentiel de la biologie, c'est l'information nécessaire à l'expression d'un caractère, le déterminant des pois ronds ou ridés de Gregor Mendel. C'est ce concept qui a assuré le lien entre les caractères visibles d'un individu et une information génétique dont on ne connaissait pas la nature. Bien avant de comprendre quel était le support physique du gène, on a su manipuler le concept et trouver des "lois" régissant, par exemple, sa transmission le long des généalogies. Curieusement la découverte de la structure fine du génome n'a pas réellement éclairci la notion de gène ; il est apparu en effet impossible de délimiter précisément une région du génome responsable de l'expression d'un caractère donné. Plus encore que dans d'autres domaines de la biologie, le foisonnement des exceptions rend difficile

l'émergence de la "solution consensus". Par contre, si on limite la notion de gène à la seule information nécessaire à spécifier l'enchaînement des acides aminés constituant les protéines telles que la cellule les synthétise, des règles formelles simples peuvent être données.

S'il ne fallait retenir que deux points de la traduction gène -> protéine, ce serait d'une part l'aspect très formel de la modélisation (l'application d'une simple table de codage : le code génétique) et d'autre part la faible densité d'information présente dans les gènes qui résulte en particulier de la structure en introns/exons. La notion de densité d'information est centrale dans la discussion sur la variabilité de la taille des génomes que nous avons esquissée au paragraphe précédent. L'absence de relation entre taille du génome et complexité résulte d'une très grande variabilité de cette densité d'information.

### La problématique bio-informatique

Cette section présente succinctement les domaines d'intervention des méthodes et outils de l'informatique en biologie moléculaire, et ce au delà de l'utilisation courante de l'informatique, commune à toute discipline scientifique.

### 1. Acquisition des séquences

L'objet de base est ainsi la séquence d'ADN, principal constituant physique du génome, "texte" écrit dans un alphabet de 4 lettres, dont certaines régions sont traduites en protéines, suites d'acides aminés, représentées par des séquences dans un alphabet de 20 lettres cette fois.

L'acquisition des séquences génomiques fait intervenir un certain nombre de manipulations physiques qui sont partiellement automatisables. Une erreur de séquençage, se traduisant par l'obtention d'une base incorrecte, peut être parfois détectée plus tard lors de la phase d'analyse de la séquence et nécessite alors de revenir à cette phase d'acquisition. L'automatisation est donc requise, à la fois pour abaisser les coûts de séquençage et pour accroître la fiabilité des résultats.

La technologie dans ce domaine est en progression constante.

L'étude des protéines pose également des problèmes d'analyse d'images et de reconnaissance de formes lors de la lecture et de l'interprétation de gels d'électrophorèse.

#### 2. Méthodes et algorithmes

Construction de cartes

L'étape suivante est de mettre bout à bout les séquences partielles, en exploitant la connaissance sur leurs recouvrements, pour obtenir des séquences d'un seul tenant de plus en plus longues, les "contigs". Plus généralement, quelles que soient les entités impliquées ("contigs", marqueurs de différentes natures, etc.), construire une carte consiste à passer de la donnée de contraintes locales sur l'ordre de ces entités à un ordre global. Dans son principe, la construction de cartes peut donc être vue comme un problème complexe de satisfaction de contraintes.

La construction de cartes génétiques impose de plus une étape préalable de calculs de probabilités complexes.

Analyse de séquences

Obtenir les séquences composant les génomes d'organismes vivants est maintenant possible à grande échelle grâce aux progrès importants des techniques de séquençage systématique. Le séquençage de génomes entiers de plusieurs organismes "modèles" est ainsi prévu à l'horizon de quelques années seulement. Reste à interpréter ces "textes". Il s'agit d'identifier des objets significatifs, tels que les gènes bien sûr, mais aussi les divers signaux de régulation de leur expression. Un problème particulièrement important est la recherche des exons et des introns à l'intérieur des gènes eucaryotes: seuls les exons, une fois mis bout à bout, sont traduits en protéines. Il faut donc mettre en évidence et rechercher dans les séquences les régularités qui sont signes de la présence de ces objets. Les méthodes utilisées sont diverses et relèvent de plusieurs domaines de recherche en informatique, automatique et mathématiques appliquées. Une classe particulièrement importante est celle des méthodes d'alignement : il s'agit de mettre en correspondance les bases

qui composent deux séquences ou plus. En pratique, ces méthodes permettent de rechercher, dans les bases de données, des séquences similaires, et donc probablement homologues, et de bénéficier des informations sur la fonction et l'évolution qui leur sont attachées.

#### Prédiction de structures

Les macro-molécules d'ARN, ainsi que les protéines, peuvent être décrites par la séquence de leurs constituants : acides nucléiques codés par les lettres A, C, G et U dans le premier cas, acides aminés codés dans un alphabet de 20 lettres dans le second. Mais ces molécules possèdent des configurations spatiales dont le rôle fonctionnel est extrêmement impor-

Sous l'hypothèse que les structures 2D et 3D sont entièrement codées dans la séquence, le but des méthodes de prédiction de structure est de déterminer ces configurations spatiales à partir de la donnée de la séquence. Là encore, les approches sont multiples et font appel à différentes techniques informatiques: propagation et satisfaction de contraintes, réseaux de neurones formels, inférence de grammaires, apprentissage symbolique, etc.

La visualisation de ces structures en deux ou trois dimensions constitue une autre classe de problèmes.

Reconstruction de phylogénies

À partir de séquences homologues de génomes de différents organismes, il s'agit de reconstruire l'arbre phylogénétique qui représente la filiation des espèces. Les méthodes développées relèvent de l'analyse de données, de la modélisation des processus temporels, mais également de l'apprentissage symbolique automatique.

### 3. Mise en œuvre des méthodes

La complexité des méthodes présentées ci-dessus (§2), ainsi que le volume des données sur lesquelles elles portent (longueur des séquences, nombre d'espèces simultanément considérées), posent un problème d'efficacité qui conduit à recourir à des implémentations parallèles, voire à des architectures matérielles spécia-

### I.A. ET BIOLOGIE MOLECULAIRE

#### 4. Bases de données

Les bases de séquences se sont développées dès le milieu des années 70. Il s'agit de stocker les séquences accompagnées d'annotations, telles que l'organisme correspondant et le laboratoire ayant effectué le séquençage, mais aussi quelques premiers résultats d'analyse. On distingue fréquemment, d'une part les banques de séquences, telles que EMBL ou GenBank, qui contiennent des séquences peu annotées, en plusieurs exemplaires, parfois erronées, et d'autre part les bases de séquences plus ou moins spécialisées, dans lesquelles un important travail de correction, d'organisation et d'annotation a été mené.

Il faut souligner que les banques et les bases de séquences constituent souvent le principal point de contact des biologistes avec l'informatique. L'opération la plus fréquente consiste effectivement à "scanner" une banque pour rechercher les séquences similaires à une séquence qui vient d'être obtenue. À titre d'exemple, si les séquences similaires retrouvées sont connues pour coder une certaine protéine, il en est vraisemblablement de même pour la nouvelle séquence.

Sous la pression du flux phénoménal de données, les biologistes ont développé de multiples banques et bases de séquences, génomiques et protéiques. La plupart d'entre elles sont des fichiers plus ou moins structurés. Il faut noter également la présence de quelques modèles spécifiques, dont le succès est dû à l'excellente adéquation aux préoccupations des biologistes. Cependant, l'utilisation du modèle relationnel, malgré son inadéquation à ce type de données, est depuis quelques années plus souvent employé. Quant au modèle à objets, il est encore extrêmement peu présent.

En effet, l'énorme volume de séquences déjà stockées, et le flux toujours croissant de nouvelles séquences (le volume des séquences contenues dans les banques double maintenant tous les ans), sont autant de freins à l'adoption de modèles de données récents, Mais le développement de nouvelles bases de séquences spécialisées, et la mise en place de structures permettant de rendre les bases existantes inter-opérables, constituent des

domaines d'intervention privilégiés de l'informatique.

Plusieurs de ces bases sont maintenant accessibles sur W3. Plus généralement, les biologistes, à travers l'accès aux bases de séquences et à la recherche de séquences homologues à distance, ont une bonne tradition d'utilisation des services Internet.

#### 5. Bases de connaissances

Au delà du stockage des séquences et de leurs annotations, existe le besoin de modéliser des connaissances plus fines sur les génomes et leur expression. Il s'agit ainsi de modéliser des connais-

- descriptives, sur les entités telles que gènes et signaux, mais aussi sur leur organisation dans des structures complexes telles que les opérons. Les modèles de connaissances à objets se révèlent bien adaptés à cette modélisa-
- comportementales, en particulier sur les mécanismes de régulation de l'expression des gènes et les liens avec le développement et le métabolisme. Il s'agit d'écrire des modèles dynamiques, mais l'insuffisance de données quantitatives peut conduire à étudier l'utilisation de modèles qualitatifs.
- méthodologiques, sur les différentes méthodes mises en œuvre (cf. §2). Il s'agit d'accompagner ces méthodes des connaissances qui permettent au système informatique d'aider un utilisateur biologiste à choisir et enchaîner les méthodes adaptées à la résolution de son problème.

Enfin, il est indispensable de conserver les liens entre ces connaissances formalisées et les connaissances textuelles (articles, thèses, monographies) qui ont servi de support à cette modélisation.

#### 6. Environnements intégrés

Il est nécessaire d'offrir des environnements complets qui permettent de représenter données et connaissances (des différentes catégories exposées au §5), de choisir, d'enchaîner et d'exécuter les méthodes (en particulier celles décrites au §2), et de visualiser et évaluer les résultats produits à l'aide d'interfaces spécialisées, en particulier cartographiques.

Les réseaux à haut débit permettent d'envisager des versions réparties de

tels environnements, qui accéderaient alors, de façon totalement transparente, aux ressources de données et de calcul disponibles auprès des centres de compétence correspondants.

#### Conclusion

La taille et la diversité des objets identifiés et manipulés par les biologistes imposent à l'évidence le recours aux outils et aux méthodes de l'informatique. Mais l'informatique, et en particulier l'intelligence artificielle, constitue également un formidable outil de modélisation et d'aide à la découverte, dont toutes les potentialités ne se sont pas encore complètement révélées en génétique moléculaire. La reconnaissance de ces apports, leur adaptation aux besoins, exprimés ou anticipés, des biologistes, et leur mise en œuvre, constituent le défi posé aux groupes bi-disciplinaires dont les travaux de recherche devront ainsi satisfaire à une double exigence de qualité. en informatique et en biologie.

### Pour en savoir plus

Pour une introduction à la biologie moléculaire :

Kourilsky P., Les artisans de l'hérédité, Éditions Odile Jacob, 1990

Pour un exposé de la problématique de l'utilisation de méthodes et d'outils informatiques en biologie moléculaire :

Schulze-Kremer S., Molecular Bioinformatics - Algorithms and Applications, Walter de Gruyter, 1996

Chaque année depuis 1993, la conférence ISMB (Intelligent Systems for Molecular Biology) rend compte des travaux de recherche les plus récents dans ce domaine. Elle s'est tenue en 1993 à Bethesda (MD), en 94 à Stanford, en 95 à Cambridge (UK), et en 96 à St. Louis (USA). Les actes sont publiés par AAAI Press.

### Cette introduction reprend des éléments de l'article suivant :

Christian Gautier, François Rechenmann, "Organisation, évolution et modélisation de l'information génomique", Congrès AFCET, Versailles (FR), (8-10 juin) 1993.

### Contributions

- Laboratoire d'Informatique Fondamentale de Lille et Service Informatique de l'Université de Mons Hainaut - INRA. Institut National de la Recherche Agronomique - Groupe "Analyse de séquences"
- LIRMM, Montpellier Systèmes et Modèles à base d'Objets et d'Agents
- LIRMM, Montpellier Systèmes d'Information Coopératifs et Bases de Données
- Laboratoire de Biologie Informatique et Théorique, Montréal, Québec, Canada
- IRISA/INRIA Rennes, projet REPCO "Représentation des connaissances"
- Laboratoire de Génétique et Physiologie du Développement, Marseille
- Laboratoire de biométrie, génétique et biologie des populations, université Claude Bernard, Lyon
- Inria Rhône-Alpes, Grenoble projet Sherpa "Modèles et bases de connaissances à objets"

Laboratoire d'Informatique Fondamentale de Lille et Service Informatique de l'Université de Mons Hainaut

Jean-Paul Delahaye, Max Dauchet, Éric Rivals Laboratoire d'Informatique Fondamentale de Lille URA CNRS 369. Université de Lille I 59655 Villeneuve d'Ascq e-mail: delahaye@lifl.fr

> Olivier Delgrange Service Informatique Université de Mons Hainaut 15. avenue Maitriau 7000 Mons, Belgique

### L'analyse des séquences génétiques par méthodes de compression

L'approche que nous proposons pour l'étude des séquences génétiques est inspirée de la théorie de la complexité de Kolmogorov qui définit le contenu en information d'une suite finie de digits (cas auquel bien sûr il est facile de se ramener en biologie) par la longueur du plus petit programme pour une machine universelle fixée qui produit cette suite.

Cette approche est impraticable tel quel à cause de la non calculabilité de la complexité de Kolmogorov, mais peut être mise en œuvre par l'utilisation d'algorithmes de compression : tout algorithme de compression est en fait un moyen de calcul approché de la complexité de Kolmogorov. C'est ce que nous avons tenté.

Ceci nous a conduit à divers types de résultats. D'abord nous avons découvert que les algorithmes classiques de compression de texte étaient inadaptés aux séquences génétiques. Nous avons donc mis au point des algorithmes spécialisés donnant des taux de compressions meilleurs sur les séquences génétiques que les algorithmes classiques. Ensuite nous avons réalisé que certains types d'algorithmes de compression (basés sur la recherche de répétitions approximatives, courtes, successives et nombreuses) pouvaient être utiles pour explorer des séquences longues : on se déplace le long de la séquence en faisant glisser une fenêtre de longueur fixe et en mesurant à chaque étape les taux de compression obtenus ; l'étude des fenêtres ayant donné les meilleurs taux de compression et les raisons structurales qui ont permis la compression des meilleures fenêtres constituent alors des informations utiles pour le biologiste qu'il peut comparer avec celles obtenues par des méthodes statistiques (par exemple).

Une idée très importante à notre sens et qui justifie l'approche "compression" est que le taux de compression est une mesure objective (ou quasi-objective) des régularités observées. Lorsque le taux de compression est important, il est certain que la régularité détectée par l'algorithme de compression est significative (ce qui n'interdit pas qu'une compression plus importante puisse être obtenue par une autre méthode : c'est la traduction dans la méthode de la non calculabilité de la complexité de Kolmogorov). Aussi les méthodes de compression fournissent la garantie que les régularités détectées sont réelles (et non pas dues à des illusions comme on en a parfois lorsqu'on recherche des "régularités" dans un objet observé avec attention), mais aussi fournissent une mesure absolue des régularités observées permettant de

les comparer quantitativement entre elles quelle qu'en soit la nature : statistique ou structurale.

E. Rivals, O. Delgrange, M. Dauchet, J.P. Delahaye: A First Step Toward Chromosome Analysis by Compression Algorithms. International IEEE Symposium INBS (Intelligence in Neural and Biological Systems), Herndon, Washington, 29-31 mai 1995, IEEE Computer Society Press (N. Bourbakis, T.Head ed.), 1995, pp. 233-239.

E. Rivals. Études d'algorithmes de compression de données et applications à la recherche de motifs dans les séquences génétiques. Thèse, janvier 1996, Université de Lille.

INRA, Institut National de la Recherche Agronomique Groupe "Analyse de séquences"

Christine. Gaspin INRA, unité de biométrie et intelligence artificielle Chemin de Borde-Rouge, Auzeville 31326 Castanet-Tolosan Tél: 61 28 52 82 email:

Christine.Gaspin@toulouse.inra.fr

L'analyse systématique de génomes chez les plantes comme chez les animaux à l'INRA nécessite la mise en commun de moyens humains et matériels. La complémentarité des compétences en présence sur le centre toulousain s'organise localement autour de quatre unités comportant chacune une composante bio-informatique. Plus particulièrement en bio-informatique dans le cadre de programmes nationaux (GREG, ACC-SV) ou internationaux (Union Européenne), les études réalisées ont progressivement mobilisé un nombre croissant de chercheurs en génétique, biologie moléculaire, biométrie et intelligence artificielle avec l'appui de la Direction nationale informatique de l'INRA.

Les thématiques d'intérêt commun sont la cartographie génétique de diverses espèces (tomate, luzerne, poulet), la modélisation moléculaire (ARN et pro-

### I.A. ET BIOLOGIE MOLECULAIRE

génomiques issues des projets de séquençage d'organismes complets. Les recherches méthodologiques autour de l'analyse génétique et l'analyse informatisée de séquences sont très actives depuis plusieurs années dans les laboratoires de génétique et biologie moléculaire. Plus récemment, l'unité de biométrie et intelligence artificielle s'est impliquée, apportant des compétences plus pointues en statistiques et IA. Les résultats acquis (par exemple [1], [2], [3], [4]) ont donné lieu à des logiciels utilisés dans les laboratoires de biologie. L'unité informatique offre des moyens (matériels, réseaux, logiciels appliqués à la biologie, assistance) constituant le premier centre de ressources bio-informatiques de la région. Dans ce groupe, les compétences de la composante intelligence artificielle (C. Gaspin et T. Schiex) s'articulent essentiellement autour des problèmes d'optimisation et de satisfaction de contraintes (CSP, [5]). Les problèmes de biologie moléculaire déjà approchés via des méthodes de l'intelligencé artificielle, et plus particulièrement via le formalisme CSP, ont suscité son intérêt ([6]). L'utilisation du cadre CSP pour la visualisation et la détermination de structures secondaires d'ARN constitue ses premiers résultats ([7], [8]). Ce travail, réalisé en collaboration étroite avec le CNRS/IBMC à Strasbourg et le CERT à Toulouse dans le cadre d'un projet d'aide à la détermination de structures d'ARN (soutenu par le GREG), débouche aujourd'hui sur la distribution dans le domaine public de logiciels tels que RNAsearch (représentation de structures secondaires d'ARN), ESSA (édition et analyse de structures secondaires d'ARN) SAPS-SARN (aide à la détermination de structures secondaires) et DRAWNA (visualisation de structures tertiaires). La complexité liée à l'élaboration d'un modèle tridimensionnel (3D) d'une molécule d'ARN a pour conséquence une hiérarchisation dans le processus de détermination de ce modèle. A partir de la séquence, le repliement en structure secondaire (2D) est d'abord recherché. La détermination d'une structure 3D s'appuie ensuite en grande partie sur la connaissance des motifs structuraux

téine), la cartographie comparée des

génomes et l'intégration des données

#### Bibliographie

[1] J.-M. Elsen, B. Mangin, B. Goffinet, C. Chevalet, Protocol designs for building genetics linkage maps in livestocks, Theoretical and applied genetics, Vol. 88, 1994.

[2] F. Corpet, Multiple sequence alignment with hierarchical clustering, Nucl. Acids Res., Vol. 16, pp. 108811-10890, 1988.

[3] E. Sonnhammer and D. Kahn, Modular rearrangement of proteins as inferred from analysis of homology, Protein Science, Vol. 3, pp. 482-492,

[4] T. Schiex, J.-C. Régin, C. Gaspin, G. Verfaillie, Lazy arc-consistency, AAAI'96.

[5] C. Gaspin, C. Bessière, A. Moisan,

2D. Le problème de la détermination d'une structure 2D se traite différemment selon que l'on dispose d'une seule séquence ou de plusieurs séquences homologues. Le cas où une seule séquence est disponible a été étudié. Un inconvénient des approches traditionnellement utilisées est la difficulté à prendre en compte les contraintes de repliement issues de données expérimentales (pseudo-nœuds en particulier). L'approche CSP proposée pour traiter ce problème permet de spécifier pas à pas une structure 2D par ajout/retrait de contraintes structurales. La prise en compte locale des pseudonœuds devient alors possible dans ce cadre. Nous nous intéressons maintenant à la modélisation 3D de ces molécules, domaine dans lequel peu de résultats existent.

Un projet plus récent faisant appel à des techniques IA concerne l'ordonnancement de marqueurs moléculaires. Disposant d'un échantillon d'une population comportant un certain nombre de marqueurs, le problème est de trouver un ordre entre ces marqueurs qui optimise sa vraisemblance maximum. Ce type de problème est abordé via des méthodes d'optimisation connues dans le cas de données complètes et nos premiers résultats se situent dans ce cadre. Le cas de données incomplètes est plus difficile à traiter. Une approche mêlant algorithmes génétiques et algorithme EM est en cours d'étude [9].

T. Schiex, Satisfaction de contraintes et biologie moléculaire, RIA, Vol. 9, N°3, pp. 355-381, 1995.

[6] C. Gaspin, P. Bourret, E. Westhof, Détermination assistée par ordinateur de structures secondaires d'ARN, TSI, Vol. 14, n°2, pp. 141-160, 1995.

[7] C. Gaspin, E. Westhof, An Interactive framework for RNA secondary structure prediction with a dynamic treatment of constraints, J. Mol. Biol., Vol. 254, pp. 163-174;

[8] C. Massire, C. Gaspin, E. Westhof, DRAWNA: A program for drawing schematic views of nucleic acids, J. Mol. Graphics, Vol. 12, 1994.

[9] O. Gach, Algorithmes génétiques et ordonnancement de marqueurs moléculaires, Rapport de DEA d'intelligence artificielle, IRIT, 1996.

### LIRMM, Montpellier – Systèmes et Modèles à base d'Objets et d'Agents

LIRMM 161 Rue Ada 34392 Montpellier Cedex 5 Tel 67 41 85 85 Fax 67 41 85 00 E-mail: nom@lirmm.fr

Responsable: J. Quinqueton Membres: D. Bardou, C. Dony, B. Esfandiari, Y. Hamadi, P. Jambaud, E. Lemoine, M. Liquière, L. Maillet-Contoz, D. Merceron, F. R. Monclar, D. Pierre, J. Quinqueton, P. Reitz, H. Ripoche, J. Sallantin, R. Terrat, L. Vignal, P. Vismara, L. Yriarte

### AMELIE: Approche Multiagents pour la Localisation et l'Explication des Introns-Exons

L. Vignal, J. Quinqueton

Ce travail est réalisé en collaboration avec Y. d'Aubenton-Carafa, F. Lisacek et C. Thermes du Centre de Génétique Moléculaire de Gif-sur-Yvette pour la partie biologie et a fait l'objet d'un contrat GREG [1]. Il s'appuie également sur les travaux de E. Mephu N'Guifo [2]. Les problèmes d'interprétation des séquences génétiques ont donné lieu à de nombreux travaux de recherche au sein desquels les méthodes issues de l'analyse de données et de l'apprentissage symbolique occupent une place importante.

Le domaine de la biologie moléculaire ne fournit que des théories faibles et des modèles incomplets de phénomènes tel que l'épissage des introns (nous nous intéressons uniquement au cas des gènes humains). Dans ce cadre, la priorité du système AMELIE est d'améliorer la compréhension des mécanismes impliqués dans l'épissage, et de ce fait, d'améliorer la reconnaissance des intronsexons (c'est à dire la distinction des parties non-codantes et codantes des gènes).

La plupart des méthodes de prédiction des exons utilisent les propriétés codantes des séquences exoniques, on pourra se référer à [3] pour un état de l'art complet. Cela n'est pas suffisant; c'est pour cette raison que nous nous intéressons aux différents éléments de la séquence reconnus par les facteurs d'épissage (les régions autour des sites accepteur et donneur, le point de branchement du lasso, la suite de pyrimidines située en amont du site accepteur), à la confrontation de ces éléments et à la sélection de ceux qui constituent les vrais sites d'épissage. Nous cherchons à modéliser l'agencement de ces divers éléments en utilisant des propriétés et en testant des hypothèses issues de l'expérimentation. En effet, AMELIE permet pour cela de simuler des modèles biologiques qui concernent les aspects dynamiques et coopératifs des interactions participant au processus d'épissage comme le scanning model, modèle de balayage [4] ou le modèle exon definition [5].

Le système AMELIE est un système multi-agents dont l'architecture simule les phénomènes connus du vivant. Ce système se présente à la fois comme relevant des systèmes dits de vie artificielle et des systèmes de résolution de problèmes. Le système dans son ensemble se présente actuellement comme un réseau d'agents à forte composante réactive, mais ayant aussi des capacités de raisonnement. Chaque agent produit une hypothèse sur une partie du problème (correspondant à une partie de la séquence analysée). Selon des critères définis par le modélisateur, ces agents négocient pour s'assembler et produire un

nouvel agent avec une hypothèse enrichie et ainsi de suite jusqu'à l'obtention de solutions (hypothèses complètes sur toute la séquence). Cette faculté de mimer le processus vivant étudié confère des capacités d'expli-

Cette approche a été présentée dans [6]. On trouvera des résultats partiels de ces travaux dans [7], concernant en particulier l'amélioration de la prédiction des sites accepteurs par l'utilisation des points de branchement et des suites de pyrimidines. Enfin, une thèse sur ce sujet est en cours de rédaction, la soutenance étant prévue en décembre 1996.

### Bibliographie

- [1] Projet AMELIE F. Lisacek, J. Quinqueton, L. Vignal, Rapport de fin de contrat GREG N°106/94, janvier
- [2] Prediction of primate splice junction gene sequences with a cooperative knowledge acquisition system, E. Mephu N'Guifo, J. Sallantin, First International Conference on Intelligent Systems for Molecular Biology, p. 292-300, 1993.
- [3] Assessment of protein coding measures, J. Fickett, C. S. Tung, Nucleic Acids Research 20(24), p. 6441-6450, 1992.
- [4] Scanning and competition between AGs involved in 3' splice site selection in mammalian introns, C. W. J. Smith, T. T. Chu, B. Nadal-Girard, Molecular and Cellular Biology 13(8), p. 4939-4952, 1993.
- [5] Exon recognition in vertebrate splicing, S. Berget, J. Biol. Chem. 270, p. 2411-2414, 1995.
- [6] AMELIE: a multi-agent architecture for the localisation and explanation of introns-exons, L. Vignal, F. Lisacek, E. Mephu N'Guifo, J. Quinqueton. Workshop on Gene-Finding and Gene Structure Prediction, Philadelphia, October 13-14, 1995.
- [7] Exon prediction in eucaryotic genomes, L. Vignal, Y. d'Aubenton-Carafa, F. Lisacek, E. Mephu N'Guifo, P. Rouzé, J. Quinqueton, C. Thermes, Biochimie, A paraître.

### Annotation et Révision de Motifs Protéiques

Jérome Gracy\*, Hugues Ripoche, Jean Sallantin \* Jerome.Gracy@EMBL-Heidelberg.de

Les motifs biologiques servent à reconnaître des sous-séquences biologiques ayant une fonction intervenant dans des réactions biochimiques. Ils correspondent à des régions stables sur le plan évolutif.

A l'EMBL, Jérôme Gracy a construit des méthodes qui construisent automatiquement des motifs à partir de sousséquences biologiques.

Ces méthodes inductives dirigées par les données constituent une suite d'étapes pour produire des motifs. Dans l'état actuel, un ensemble restreint de ces motifs 10% (200) sont déjà connus et ils correspondent à 90% des motifs de PROSITE. La base de données PROSITE est construite selon une démarche opposée ; elle est le résultat de la collation de nombreuses expertises humaines.

Il existe donc deux manières de produire les motifs : une construction par classification automatique et une élicitation menée directement par les biologistes. Une confrontation de ces démarches semble extrêmement intéressante, car la qualité des résultats est suffisante pour justifier le travail suscité par la confrontation.

Dans les travaux menés dans sa thèse, Jérôme Gracy a proposé des modules d'alignements multiples tenant compte d'informations structurelles. Il s'est avéré qu'en additionnant des contraintes géométriques, aux séquences protéiques, il était possible d'interpréter le regroupement par leur fonction de séquences de faibles ressemblances primaires. Il est ainsi possible de décharger le biologiste du travail d'homologie pour lui permettre de se concentrer plus directement sur le travail de découverte.

Dans cet esprit, le travail de thèse d'Hugues Ripoche a porté sur l'aide à la découverte. Il a montré comment la structure de treillis de Galois permet de construire une navigation locale sur un ensemble de séquences protéiques regroupées par leurs propriétés dans un

### I.A. ET BIOLOGIE MOLECULAIRE

SGBDOO. Les informations d'annotations supplémentaires permettent de spécialiser/généraliser des classes pour mieux en comprendre les articulations. Suite à sa thèse, Hugues Ripoche a mis au point au LIRMM dans le contexte d'un contrat GREG (94 Framentec Euroclid Lirmm) un serveur WWW intégrant des outils de J. Gracy d'homologie de séquences en accord avec sa problématique de thèse. Ce serveur va compléter le serveur de motifs qui est mis au point à l'EMBL par Jérôme Gracy dans le cadre de l'équipe d'AR-GOS.

### Bibliographie

J. Gracy, Conception d'un environnement d'apprentissage appliquée à l'étude structurale des protéines. Thèse, Université Montpellier II 1993 H Ripoche, une construction interactive d'interprétation de données (application aux bases de données de séquences génétiques. Thèse, Université Montpellier II 1995.

> LIRMM, Montpellier Systèmes d'Information Coopératifs et Bases de Données

LIRMM, Montpellier-Systèmes d'Information Coopératifs et Bases de Données Responsable : D. Herin

Membres: Z. Bellahsène, C. Boksenbaum, D. Herin, T. Libourel, P. Pompidor, M. Teisseire, I. Bedou, S. Billet, L. Domenech, F. Koriche, I. Mougenot, M. Sala. LIGM-DB: système d'information en immunogénétique I. Mougenot, T. Libourel

#### Introduction

Intégrer au sein d'un même système, l'information relative aux données brutes expérimentales et celle relative à l'expertise est au centre de la conception de LIGM-DB [1].

En effet, les mécanismes génétiques complexes et tout à fait uniques (combinatoire entre les gènes) [2] qui régissent la synthèse des récepteurs d'antigènes constituent un contexte particulièrement attrayant et novateur pour ce genre d'entreprise.

#### Notre approche

Le défi auquel nous étions confrontés était donc de gérer de manière uniforme, les données de base et les interprétations ou annotations sur ces données.

La méthodologie suivie se base sur les quatre étapes constituant le parcours de l'information depuis le séquençage jusqu'à la gestion des séquences dans la base de données [3]:

- acquisition des séquences résultats expérimentaux du biologiste,
- exploitation des séquences qui correspond à la notion d'annotation [4], c'està-dire à l'extraction et à l'interprétation (expertise) des informations contenues dans ces séquences,
- modélisation intégrée uniforme des données expérimentales et des annota-
- gestion et restitution de ces informations à la communauté.

L'originalité de l'approche est double : nous avons pris conscience à la fois de l'importance d'une modélisation globale (données expérimentales et expertisées) et de l'importance de l'étape d'annotation qui constitue, non plus un travail autonome, mais un élément de contrôle de la pertinence et de la consistance des informations stockées.

Le système créé s'appuie sur les modèles les plus répandus dans le cadre des systèmes d'information (modèle entitéassociation, modèle objet) et est implémenté sur les standards de la communauté (SGBD relationnels, interface OSF/Motif et W3).

L'étape d'annotation a mis en lumière outre la complexité de structure, l'émergence de contraintes spécifiques que nous ne pouvions traiter que par des modèles plus riches et plus flexibles [5]. Une grammaire descriptive des séquences et de leurs sous-régions a été proposée. Cette grammaire a deux résultats pertinents:

- un ensemble de termes consensuels décrivant les concepts et les relations des séquences du domaine de l'immunogénétique validé par les experts du réseau international LIGM-DB,
- une proposition d'outil de contrôle et d'aide semi-automatique pour l'annotation implémenté sous Prolog.

#### Conclusion

Outre le fait que ce travail a permis une collaboration étroite entre divers laboratoires de communautés scientifiques diverses (informatique et biologie), nous pouvons aussi affirmer que les retombées scientifiques de l'expérience sont loin d'être terminées.

La meilleure connaissance des sousrégions constitutives des séquences que cette expérience a révélée peut par exemple, dans le cas des régions variables, aboutir à la modélisation des "boucles" des récepteurs d'antigènes au contact de l'antigène et avoir une implication en immunothérapie (vaccins,

### Bibliographie

[1] "LIGM-DB un système de gestion intégré des séquences nucléiques et de leurs annotations", I. Mougenot, Thèse de doctorat, Université Montpellier II, Décembre 1995.

- [2] "Multiple mechanisms participate in the genration of diversity of human H chain CDR3 regions", J. Sanz, The Journal of immunology, Vol 147, pp 1720-1729, 1991
- [3] "Colibri, une base de données spécialisée dédiée à l'analyse du génome d'escherichia coli ", C. Médigue, Les Cahiers Imabio, CNRS, Vol 2, pp 5-13, 1992
- [4] "Traitement de l'information des séquences biologiques", A. Viari, Les Cahiers Imabio, CNRS, Vol 1, pp 31-39, 1991
- [5] "The linguistics of DNA", DB. Searls, American Scientist, Vol 80, pp 579-591,1992.

Sigale, un environnement d'aide à la découverte scientifique en immunogénétique

#### Michel Sala

### Introduction

Notre travail se situe dans le cadre d'une aide à la découverte scientifique pour un chercheur en immunogénétique. Dans ce contexte, le chercheur peut concevoir, construire et réviser son modèle du domaine étudié. L'originalité de notre approche réside dans la coopération et l'interaction entre des outils informatiques existants, des bases de données du domaine et d'une base de connaissances que nous appellerons "modules". Le

processus de contrôle et de synchronisation entre ces modules est réalisé par un planificateur.

### Le contexte du travail du chercheur en immunogénétique

Suite aux différents programmes de séquençage massif du génome, le chercheur se trouve confronté à l'analyse, la gestion, la représentation et le traitement d'un flot de données très important. Pour l'aider dans sa démarche, il a à sa disposition une pléthore de logiciels de calcul informatique sur la recherche de motifs, l'alignement ou la visualisation en deux ou trois dimensions. De nombreuses bases de données stockent ces différentes données en utilisant des modèles relationnels, objets et même déductifs. La création d'outils de navigation assistée lui permettent même de consulter ces différentes bases généralistes ou spécialisées en utilisant par exemple des références croisées. Toutes ces informations sont disponibles sur le réseau Internet et le chercheur peut rapatrier sur sa propre machine le résultat de ces requêtes sur les bases ou même les outils de traitement qu'il pourra ensuite utiliser en local.

Concernant le chercheur qui souhaite faire évoluer ses connaissances, l'ensemble de ces outils lui sont fort utiles, mais il doit construire lui même son propre modèle, tant au niveau de la donnée d'expérimentation que des résultats de calcul qu'il souhaite obtenir. Malgré la volonté de standardisation des entrées/sorties des outils de calcul ou des bases de données, il doit sans cesse modifier ses formats de données ou interpréter dans un format canonique les différents résultats des outils de calcul.

#### L'objectif de notre travail

Notre travail porte sur la mise en place d'un environnement qui va permettre au chercheur de disposer de divers modules. Nous distinguerons:

- le module qui permet de représenter son modèle du domaine sur les données d'expérimentation et sur les résultats de calcul en créant une base de connaissances,
- le module interface entre les outils de calcul et le chercheur qui lui permet de transformer de manière automatique les formats entrée et sortie pour fournir des résultats dans le même formalisme (format "harmonisé").

- le module de recherche automatique des informations complémentaires dans les bases de données du domaine qui permet de présenter ces résultats dans un certain formalisme,
- le planificateur qui contrôle et qui synchronise les interactions inter-modules (base de connaissances, outils de calcul, bases de données),
- le module explications qui permet au chercheur de réviser son modèle du domaine dans le cas où les données ou les résultats des outils sont non conformes à la représentation du modèle du chercheur stocké dans la base de connaissances.

### Notre proposition: l'environnement Sigale

Sur les idées précédentes, nous avons développé une méthodologie et une architecture informatique pour créer un environnement d'aide à la découverte scientifique en immunogénétique. Ce développement a donné naissance au système SIGALE (System-IG ALignement Explanation, système d'alignement et d'explications pour les séquences d'immunoglobulines (IG)), que nous avons testé et validé avec le LIGM (Laboratoire d'ImmunoGénétique Moléculaire) de Montpellier dirigé par les professeurs Marie-Paule et Gérard Lefranc. Nous avons plus particulièrement travaillé sur la classification automatique des régions variables des immunoglobulines et des récepteurs de cellules T (TcR).

Pour obtenir cette classification dynamique, nous disposons d'un échantillon initial composé d'un ensemble de séquences protéiques préalablement classées. En utilisant un algorithme de classification, nous obtenons une classification en sous-familles. A partir de cette classification initiale, nous ajoutons cycliquement une nouvelle séquence protéique et nous analysons le résultat de la nouvelle classification obtenue.

- Dans cette problématique :
- les données d'expérimentations sont des séquences protéiques d'immunoglobulines.
- les outils de calcul sont des logiciels d'alignement et de classification. Ils permettent d'aligner et de classifier des séquences protéiques d'immunoglobulines. Historiquement, nous avons choisi comme algorithme de classification

- "align8.x" [GRAC 93], qui procède en deux temps, en premier lieu il réalise un alignement par paire, puis un alignement multiple. Nous avons ensuite utilisé les logiciels QR2 [GAS 96] et Clustal W [THO 94].
- la base de connaissances sur les données en entrée représente les contraintes sur ces données, spécifiées par des points d'ancrage particuliers par espèces et des distances inter points d'ancrage. Ces contraintes sont implémentées sous forme de règles de production,
- la base de connaissances des résultats des outils, qui indique le nombre de sous-familles autorisés, pour un seuil de similarité fixé.
- concernant le module de recherche dans les bases de données, nous avons utilisé le logiciel SRS [ETZ 93] développé par Thure Etzold, chercheur à l'EMBL de Heidelberg en Allemagne. Seules des requêtes simples à partir de trois points d'entrée : le numéro de la séquence (accession number), nom d'auteur et la longueur de séquence, sont possibles. Elles sont toutefois très rapides et peuvent être combinées par des opérateurs logiques.
- le module de création des explications générées par les différents modules présentés ci-dessus.

### La phase de révision

Le système SIGALE s'adresse à un chercheur en génétique et son objectif est d'aider le chercheur à réviser ses hypothèses à la fois sur la donnée d'expérimentation et sur les résultats d'un outil. Pour cela il procède en trois

- à partir de nouvelles séquences expérimentées par le chercheur, l'environnement vérifie que l'ensemble des contraintes sur la donnée sont respec-
- en cas de violation de ces contraintes, il déclenche d'autres outils de classification afin de vérifier la validité de ses propres résultats.
- · il utilise des bases de données du domaine pour rechercher des informations complémentaires sur la donnée d'expérimentation pour l'aider à prendre une décision.

Afin d'aider le chercheur à réviser ses hypothèses, le rôle des explications est fondamental dans chacune de ces trois

### I.A. ET BIOLOGIE MOLECULAIRE

- dans la première étape, les explications sont construites à partir d'une confrontation entre la donnée d'expérimentation et les hypothèses du chercheur,
- · dans la deuxième étape, les explications sont construites à partir du résultat de différents outils disponibles sur les réseaux et de leur confrontation avec les hypothèses du chercheur,
- · dans la troisième étape, les explications sont construites à partir de la concaténation et analyse des informations issues de différentes bases de données a priori hétérogènes.

La forme des explications doit respecter la forme usuelle des outils de travail du chercheur. Par exemple, le chercheur en biologie a l'habitude de manipuler un formalisme propre à un outil et il est donc nécessaire de transformer le formalisme spécifique des outils dans un formalisme canonique.

#### Conclusion

La méthodologie et l'architecture proposée ont pu être validées en immunogénétique en montrant une particularité sur un point d'ancrage du requin et en proposant un ensemble de motifs communs pour la super-famille des immunoglobulines. Ces résultats constituent une première étape vers une extension de cette approche à tout autre domaine en sciences expérimentales.

### Bibliographie

[ETZ 93] "SRS - an indexing and

retrieval tool for flat file data librairie"

- Etzold T., Argos P. - CABIOS, Vol 9

n° 1, pp 49-57, 1993

[GAS 96] "A reduction algorithm for approximating a (non-metric) dissimilarity by a tree distance" - Journal of Classification, vol 13, 1996 [GRA 93] "Conception d'un environnement d'apprentissage appliqué à l'étude structurale des protéines" -Gracy J. - Thèse de troisième cycle, univ de Montpellier II, décembre 93 [THO 94] "Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties weight matrix choise" - Thompson J.D., Higgins D.G., Gibson T.J. -

Nucleic Acids Reseach, Vol 22, pp

4673-4680, 1994

Laboratoire de Biologie Informatique et Théorique, Montréal, Québec, Canada

Département d'Informatique et de Recherche Opérationnelle, Université de Montréal, C.P. 6128, succ. centre-ville, Montréal, Québec H3C 3J7, Canada http://www.iro.umontreal.ca/labs/ lbit/HTML/LABO/index fr.htm

Reconstruction de structures tridimensionnelles d'ARN François Major major@iro.umontreal.ca http://www.iro.umontreal.ca/ userinfo/user?major)

La connaissance de la structure d'une macromolécule est essentielle pour en déduire son fonctionnement. Cependant, il est bien connu que le problème qui consiste à prédire la structure tridimensionnelle d'une protéine à partir de l'information contenue dans sa séquence de résidus (PPSTP) n'est toujours pas résolu et ce malgré un nombre grandissant de structures déterminées par cristallographie aux rayons X et par résonance magnétique nucléaire (RMN). Depuis plus de vingt ans, c'est en vain que de nombreux chercheurs ont tenté d'identifier les règles qui permettraient de reconstruire les structures tridimensionnelles de ces protéines. Sans ces règles, le PPSTP est un problème de recherche combinatoire à travers un nombre insaisissable de conformations possibles déterminées par plusieurs degrés de liberté dans les angles de torsions de chaque résidu.

Plus récemment, la découverte de propriétés catalytiques chez les acides ribonucléiques (ARN) et un intérêt grandissant pour les utiliser dans des thérapies géniques nous ont amenés à poser le problème de prédiction de la structure tridimensionnelle d'un ARN à partir de l'information contenue dans sa séquence de nucléotides (PPSTAN). Bien que moins étudié que son analogue des protéines, le PPSTAN possède la même problématique. Afin de contrer la combinatoire, nous l'avons posé comme un problème de satisfaction de contraintes. Pour le résoudre, nous recherchons de l'information structurale par des analyses minutieuses des séquences d'ARN et nous effectuons des expériences en laboratoire dont les résultats sont toutefois moins précis que ceux obtenus par cristallographie aux rayons X ou RMN. Pour réaliser notre objectif de réduire au minimum l'intervention humaine dans le processus de modélisation, nous employons des techniques empruntées au domaine de l'intelligence artificielle. C'est ce dernier aspect qui distingue notre approche de celles préconisant l'interactivité graphique.

Nous avons développé au cours des cinq dernières années le logiciel MC-SYM (1) (FTP anonymous: ftp.iro.umontreal.ca dans le répertoire pub/lbit ou au site WWW http://www.IRO.UMontreal.CA/peopl e/major/mcsym.html). Une syntaxe spécifique à MC-SYM permet de condenser l'information structurale représentée par un graphe de relations. MC-SYM interprète le graphe et génère systématiquement des structures partielles qui satisfont à toutes les contraintes géométriques déduites de l'information structurale en appliquant un algorithme de retour-arrière (2,3). L'extension des structures partielles jusqu'à des structures qui contiennent tous les nucléotides constitue les solutions du problème. Nos structures tridimensionnelles sont construites en appliquant des relations spatiales binaires et en assignant des conformations de nucléotides provenant de la banque de données de structures déjà déterminées ou, lorsqu'aucun cas s'applique, de la génération de nouveaux cas par des méthodes de mécanique quantique. Une des étapes les plus pénibles est d'extraire les relations et les conformations de la banque de données en visualisant les structures et en annotant chaque cas manuellement. L'automatisation de cette tâche est actuellement l'une des préoccupations principales de notre équipe de recherche, en plus d'avoir un tampon toujours rempli de nouvelles molécules à modéliser. D'autres projets du laboratoire incluent la recherche de motifs dans les structures secondaires des ARNs et le transfert de la technologie MC-SYM vers les protéines.

Il faut mentionner que mes recherches sur la modélisation des ARNs ont débuté en 1988 alors que j'amorçais mes études de doctorat avec le groupe du Professeur Robert Cedergren au département de Biochimie de l'Université de Montréal. Par la suite, j'ai pu poursuivre ce projet en stage postdoctoral avec l'équipe du Dr David Lipman au National Centre for Biotechnology Information au National Institutes of Health à Bethesda au Maryland. Je suis revenu à l'Université de Montréal au cours de l'été 1994 pour y fonder le Laboratoire de Biologie Informatique et Théorique au Département d'Informatique et Recherche Opérationnelle. Notre équipe est actuellement composée de huit chercheurs et étudiants dont plus de la moitié sont informaticiens. On y compte aussi un chimiste, un biologiste et une physicienne. Un autre aspect intéressant de notre groupe est qu'il est situé dans un département d'informatique, alors que la plupart des groupes similaires se retrouvent dans des départements de sciences biologiques.

La méthode MC-SYM fut publiée en 1991 dans un article de la revue "Science" (1) qui la qualifiait dans son éditorial de "cutting edge" en informatique dans le domaine de la biologie moléculaire. Bien que le logiciel soit maintenant utilisé dans plus d'une centaine d'institutions d'enseignement et de recherche ainsi que dans quelques compagnies pharmaceutiques, il reste beaucoup à faire pour améliorer la méthode et réduire davantage l'intervention humaine. Parmi les résultats les plus importants générés par MC-SYM, on compte la reproduction, à une précision jamais égalée depuis, de la molécule de l'ARN de transfert du Phenylalanine chez la levure (4), la prédiction de la structure de la molécule de "REV binding", et d'une jonction de trois hélices dans l'ARN de la Ribonuclease P (5), ainsi qu'une validation du modèle S du positionnement de deux ARN de transfert dans les sites A et P du ribosome (6). Nous croyons que MC-SYM constitue actuellement une référence internationale pour la modélisation tridimensionnelle des acides ribonucléiques par ordinateur.

### Références 1. F. Major et als. (1991) Science, 253 pp. 1255-1260. 2. F. Major et als. 1991) Journal of Functional Programming, 1, pp. 213-3. F. Major et als. (1992) Actes des

journées JFLA92: Avancées applicatives, Bigre 76-77, pp. 202-219, Rennes, France. 4. F. Major et als. (1993) Proceedings

of the National Academy of Sciences, 90, pp. 9408-9412. 5. J.W. Brown et als. (1996) Proceedings of the National Academy of

Sciences, 93, pp. 3001-3006. 6. T. Easterwood et als. (1994) Nucleic Acids Research, 22, pp. 3779-

### IRISA/INRIA Rennes, projet REPCO "Représentation des connaissances"

Campus de Beaulieu, 35042 Rennes cedex Contact: Jacques Nicolas, tél. 99 84 73 12, jnicolas@irisa.fr Autres participants: J.Y. Giordano, R. Gras, I.C. Lerman, C. Sinoquet, B. Tallur

Le projet REPCO de l'IRISA s'intéresse d'une façon générale à la gestion de données et connaissances, en se concentrant sur les aspects acquisition et raisonnement. En ce qui concerne la partie acquisition, notre projet de recherches vise le développement d'outils automatiques permettant de structurer une base de données et de connaissances, c'est-à-dire à la fois effectuer des regroupements et établir des relations entre les objets de ces groupes. Les disciplines scientifiques concernées sont essentiellement l'apprentissage automatique et l'analyse de données.

Plus particulièrement, un aspect important de ces recherches consiste à concevoir et à réaliser des outils d'analyse de séquences. Nous avons travaillé sur les séquences d'ADN et sur les protéines.

En ce qui concerne les séquences d'ADN, nous nous intéressons à la conception d'outils d'analyse lexicale

et syntaxique. Nous développons d'une part, FOREST (FOuineur de RÉpétitions dans les Séquences Titanesques) un nouvel outil d'analyse exploratoire de grandes séquences biologiques. Il vise à faciliter la recherche de portions de séquences remarquables par le biologiste et est basé sur deux idées principales. La première consiste à travailler sur une représentation sous forme d'arbre des suffixes de la séquence. On a ainsi accès à tous les sous-mots répétés dans la séquence. On associe à chaque nœud de l'arbre un ensemble d'attributs synthétisés à partir de ses sous-nœuds. Cela permet au biologiste de "fouiner" dans la séquence avec une vue abstraite de ce qu'il peut espérer trouver dans la section de l'arbre qu'il est en train d'explorer. La deuxième idée consiste à résumer l'information avec des vecteurs booléens associés à la séquence. Ces vecteurs peuvent ensuite être visualisés sous forme de cartes linéaires mais permettent également de proposer une puissante capacité de filtrage des résultats. Ces deux mécanismes permettent au biologiste de focaliser sa recherche sur les aspects les plus intéressants de la structure lexicale de ses séquences. Nous avons ainsi pu traiter les séquences connues de Bacillus subtilis et Escherichia coli (plus de 3.5M de bases). Il n'y a pas de travaux équivalents à notre connaissance à part ceux menés dans l'équipe de M. Crochemore, au LITP, qui s'intéressent plus à la recherche de motifs donnés (pattern matching) qu'à un outil exploratoire.

D'autre part, nous commençons à étudier comment analyser ces séquences à l'aide de grammaires formelles. Nous utilisons la programmation logique comme support de représentation et d'analyse. Le but est d'adapter le formalisme des DCGs à l'étude des structures naturelles rencontrées dans un génome. Les travaux de D. Searls (University of Pennsylvania) et de J. Collado-vides (Universidad Nacinal Autónoma de México) se rattachent à cette problématique. Nous poursuivons également des travaux sur l'inférence grammaticale afin de produire automatiquement une partie des grammaires nécessaires. En particulier, un travail est en cours sur l'analyse de régions

### I.A. ET BIOLOGIE MOLECULAIRE

promotrices de gènes. D'autres approches inductives ont été essayées sur ce type de problème de prédiction avec un certain succès (notamment en programmation logique inductive par S. Muggleton, Oxford University), mais il n'y a pas encore de travaux conséquents en inférence grammati-

En ce qui concerne les protéines, nous

travaillons principalement sur le problème de la prédiction de la structure dans l'espace de ces protéines à partir de la connaissance de leur chaîne d'acides aminés. Une première étape s'intéresse aux structures secondaires (hélice, feuillets, boucles). Nous avons développé une méthode de prédiction par construction d'arbres de décision et une seconde par recherche optimale de segmentation à partir d'un ensemble de règles. Les résultats sont comparables aux meilleures méthodes actuelles, avec l'avantage de règles de décision plus explicites. Les perspectives concernent la reconnaissance de structures tridimensionnelles, où l'aspect syntaxique de liaisons à distance jouera un plus grand rôle. Nous avons également travaillé sur la classification de protéines, que ce soit à partir de données alignées ou non alignées. L'étude comportait des aspects à la fois statistiques et combinatoires.

### Références [nic94] 'Classification et caractérisa-

tion de séquences. Le cas des aminoacides-ARNt ligases' J. Nicolas, I.C. Lerman, B. Tallur, J.Y. Giordano, J. Lebbe, R. Vignes Rapport de fin de contrat - décision no 92H0910 du MRE- Février 1994 [nic95] 'Classification et caractérisation de séquences. Définition d'un outil d'analyse pour les grandes séquences. Application à l'étude du Bacillus subtilis.' J. Nicolas, J. Lebbe, R. Ngouenet, B. Tallur, R. Vignes Rapport de fin de contrat -Décision no 59 du GIP GREG - Mars 1995 [cost96] 'Coefficients d'association et

binarisation par la classification hiérarchique dans les arbres de décision. Application à l'identification de la structure secondaire des protéines' J. Pinto da Costa Thèse de l'université de Rennes 1, 1996

Laboratoire de Génétique et Physiologie du Développe ment, Marseille

UMR CNRS 9943 Parc Scientifique de Luminy, Case 907 13288 MARSEILLE CEDEX 9 tel 91 26 90 55, Fax 91 82 06 82 e-mail: jacq@lgpd.univ-mrs.fr

> Christophe Chemla Florence Horn **Violaine Pillet Bernard Jacq**

Les progrès accomplis ces dernières années en biologie et génétique moléculaire, notamment à travers les programmes de cartographie et séquençage de différents génomes, ont amené aux biologistes de colossales quantités d'informations nouvelles dont la grande majorité sont de nature essentiellement descriptive et structurale. Le prochain défi auquel sera confronté la biologie est d'acquérir les informations fonctionnelles correspondantes sur tous les gènes identifiés et d'intégrer ces connaissances. En particulier, il sera essentiel de décrire et de comprendre comment tous les gènes connus interagissent entre eux. En effet, quel que soit le processus biologique considéré, la caractérisation des interactions géniques qui le contrôlent, leur représentation et leur assemblage en un réseau d'interactions et la modélisation du fonctionnement de ce réseau sont des étapes essentielles à sa compréhension. La bioinformatique peut jouer un rôle essentiel, jusqu'ici peu exploré, dans la réalisation de ces différentes étapes. En utilisant comme modèle expérimental et source de données le contrôle moléculaire de l'embryogenèse précoce de la drosophile, nous avons initié un projet (soutenu par une ACC-SV bio-informatique du ministère) visant à :

• collecter des données sur les interactions géniques et de les organiser en un fichier cohérent au format standardisé et accessible sur réseau Internet.

• implémenter ces données d'interactions et d'autres données sur le développement en utilisant Tropes, un modèle de connaissances à objets et points de vue multiples, développé à

Grenoble dans le projet Sherpa, dirigé par François Rechenmann à l'Inria Rhône-Alpes

· réaliser un ensemble de modules algorithmiques travaillant sur la représentation interne du graphe d'interactions et capable de déterminer les conséquences, au niveau de l'activation ou de la répression génique, de variations apportées sur les éléments du réseau en fonction de conditions biologiques (stade développemental, région de l'embryon ou encore situation mutante particulière),

• réaliser un éditeur graphique d'interactions, utilisable pour la saisie des données et l'affichage des résultats des algorithmes d'analyse du comportement dynamique.

Les deux premières phases de ce projet sont actuellement bien avancées puisque d'une part un serveur WEB contenant une banque de données décrivant des interactions moléculaires directes (pour le moment entre protéines et gènes) impliquées dans les processus développementaux chez la drosophile est maintenant accessible sur Internet (http://wwwbiol.univ-mrs.fr/GIFTS\_home\_page.h tml) et d'autre part un premier prototype de bases de connaissances écrit en Tropes et dédié aux interactions géniques a été réalisé en collaboration étroite avec Jérôme Euzenat du projet Sherpa (Inria Rhône-Alpes).

Le prototype de bases de connaissances intègre 18 concepts différents, 49 points de vue, 459 classes. Du côté des données, il rassemble actuellement des informations concernant 62 interactions intervenant entre 25 gènes. Une interface utilisateur a été développée en HTML permettant le dialogue avec la base de connaissances. Des fonctions ont été créées permettant:

- · d'effectuer des requêtes sur les différents champs de la base,
- de connaître tous les gènes cibles ou tous les régulateurs d'un gène donné,
- · de parcourir le graphe du réseau d'interactions pour connaître les chemins de régulation possibles entre deux gènes quelconques.

Dans le futur, nous prendrons également en compte les contraintes spatiales et temporelles au niveau de l'expression des gènes.

Laboratoire de biométrie, génétique et biologie des populations Université Claude Bernard, Lyon

Laboratoire de Biométrie, Génétique et Biologie des Populations Université Claude Bernard Lyon 1 43, Boulevard du 11 novembre 1918 69622 Villeurbanne Cedex Tél.: 72 44 81 96

Fax: 78 89 27 19 Email: cgautier@biomserv.univ-lyon1.fr

Les activités dans le domaine de l'informatique et des génomes de l'équipe "Biométrie Moléculaire, Evolution et Structure des Génomes" s'organisent autour de trois axes principaux : une activité de service, une activité de développement d'outils, et une activité de modélisation.

#### Activité de service

Pour l'essentiel, nos activités de service consistent en le développement et la maintenance du système de gestion de banques de séquences ACNUC (Gouy et al., 1985). Ce système permet de consulter des banques généralistes (telles que GenBank, EMBL, NBRF/PIR) ainsi que des banques dédiées (Hovergen, NRSub). Depuis un an maintenant, le serveur WWW-Query permet d'accéder à ces banques au travers du World Wide Web (Perrière et Thioulouse, 1996; Perrière et Gouy, 1996). Nous développons et diffusons aussi des programmes de phylogénie moléculaire.

### Développement d'outils

Un certain nombre de banques de données dédiées ont été développées dans notre équipe. Parmi ces banques, les deux qui sont le plus utilisées à l'heure actuelle sont Hovergen (Duret et al., 1994) et NRSub (Perrière et al., 1996). Hovergen est consacrée aux séquences et à la phylogénie des vertébrés tandis que NRSub est une banque non-redondante de séquences de la bactérie Bacillus subtilis. Toutes deux sont utilisées par de nombreux laboratoires de biologie, ceci de façon courante.

Pour étudier les phylogénies moléculaires, d'importants développements ont également été réalisés dans notre équipe. Ces développements se sont concrétisés sous forme de deux logiciels: SeaView et Phylo\_win. Ces deux programmes constituent en fait de véritables environnements intégrés pour l'aide à la reconstitution phylogénétique. SeaView est un système d'aide à la construction d'alignements multiples (Perrière et Gouy, 1996), tandis que Phylo\_win permet de construire une phylogénie à partir d'un alignement (Galtier et Gouy, en préparation).

Enfin, nous collaborons avec des biométriciens spécialistes de l'analyse multivariée pour appliquer les nouvelles méthodes du domaine aux séquences biologiques. Cette collaboration s'est tout d'abord concrétisée par le développement d'un serveur W3 permettant de coupler interrogation des banques et analyse multivariée des séquences (Perrière et Thioulouse, 1996). Des résultats

biologiques nouveaux ont également été obtenus en utilisant des méthodes comme l'analyse factorielle des correspondances (Lobry et Gautier, 1994), l'analyse de co-inertie (Thioulouse et Lobry, 1995) ou l'analyse discriminante des correspondances (Perrière et al., soumis).

#### Modélisation

Sur l'exemple de la recherche des parties codantes a été explorée la modélisation des stratégies en analyse des séquences au moyen du système de gestion de tâches Scarp développé par le projet Sherpa à l'Inria Rhône-Alpes. La complémentarité des concepts de spécialisation et de décomposition est apparue comme un moyen puissant d'expression de la complexité des stratégies. Le troisième aspect important de cette modélisation est l'intégration du contrôle sur le flux des données entre tâches dans l'expression de la stratégie. L'exemple traité est un problème réputé difficile : le repérage des parties codantes dans les génomes des eucaryotes supérieurs (Perrière et al., 1994).

#### Publications

Duret, L., Mouchiroud, D. et Gouy, M. (1994) Hovergen: a database of homologous vertebrate genes. Nucleic Acids Res., 22, 2360-2365. Gouy, M., Gautier, C., Attimonelli, M., Lanave, C. et di Paola, G. (1985) ACNUC - a portable retrieval system for nucleic acid sequence databases: logical and physical designs and usage. Comput. Applic. Biosci., 1, 167-172. Perrière, G., Chevenet, F., Dorkeld, F., Vermat, T. et Gautier, C. (1994) Building integrated systems for data representation and analysis in molecular biology, In Proceedings of the 27th Hawai International Conference on Systems Science Vol. 5, Hunter, L. (ed.), IEEE/ACM, pp. 89-96. Lobry, J. et Gautier, C. (1994) Hydrophobicity, expressivity and aromaticity are the major trends of amino-acid usage in 999 Escherichia coli chromosome-encoded genes: Nucleic Acids Res., 22, 3174-3180. Pérrière, G. et Gouy, M. (1996) WWW-Query: an on-line retrieval system for biological sequence banks. Biochimie, 78, 0-0. Perrière, G., Moszer, I. et Gojobori, T. (1996) NRSub: a non-redundant database for Bacillus subtilis. Nucleic Acids Res., 24, 41-45. Perrière, G. et Thioulouse, J. (1996) On-line tools for sequence retrieval and multivariate analysis in molecular biology. Comput. Applic. Biosci., 12, 63-69. Thioulouse, J. et Lobry, J. (1995) Co-inertia analysis of amino-acid physico-chemical properties and protein composition with the ADE package. Comput. Applic. Biosci., 11, 321-329.

### I.A. ET BIOLOGIE MOLECULAIRE

Inria Rhône-Alpes, Grenoble Projet Sherpa "Modèles et bases de connaissances à objets'

Unité de Recherche INRIA Rhône-Alpes ZIRST 655 avenue de l'Europe 38330 Montbonnot Saint Martin Tél. (+33) 76 61 53 65 Fax (+33) 76 61 52 52

## Francois.Rechenmann@inria.fr

Les travaux de recherche du projet Sherpa au sein de l'Unité de Recherche de l'Inria Rhône-Alpes portent sur les modèles de connaissances à objets et sur leurs mécanismes d'exploitation, tels que la classification. Depuis une dizaine d'années, la biologie moléculaire est le domaine d'expérimentation privilégié des résultats de ces travaux, et en particulier des logiciels développés [Rechenmann 95]. Dans une première phase, les outils de modélisation des connaissances ont été utilisés par les laboratoires de génétique moléculaire partenaires du projet pour la constitution de grandes bases décrivant les différentes classes d'objets biologiques [Schmeltzer& 93]. La phase actuelle concerne la modélisation des connaissances méthodologiques relatives à l'analyse et à l'interprétation des séquences, pour y découvrir et caractériser les entités significatives, telles que les gènes et les signaux de régulation de leur expression.

### ColiGene: une base de connaissances à objets sur la régulation de l'expression génétique d'Escherichia coli

La base ColiGene a été développée au laboratoire de biométrie, génétique et biologie des populations de l'université Claude Bernard de Lyon. Elle utilise le modèle de connaissances à objets Shirka concu par le projet Sherpa à partir de 1985. ColiGene rassemble des connaissances sur la régulation de l'expression génétique dans le génome de la bactérie Escherichia coli. Ces connaissances sont exprimées d'une part sous forme d'objets

Shirka et d'autre part via un réseau d'hypertextes connecté à ces objets [Perrière 92] [Perrière 93].

### MultiMap et la modélisation des cartes génomiques

La base de connaissances MultiMap [Dorkeld 94] a été également développée au sein du laboratoire de biométrie, génétique et biologie des populations de l'université Claude Bernard de Lyon et repose sur le modèle et le système Shirka. Elle est destinée à rendre compte des différents niveaux de cartographie des génomes des mammifères.

Le développement de cette base a conduit à étudier, au sein du projet Sherpa, les problèmes liés à la formalisation et la représentation de cartes génomiques [Schmeltzer 95]. Cellesci décrivent un même chromosome selon plusieurs perspectives, chacune issue d'expériences spécifiques. Diverses entités biologiques (gènes, signaux, séquences répétées, etc.) se positionnent sur ces cartes selon un ordre parfois incomplet et se décomposent éventuellement en sous-éléments. La modélisation de ces cartes génomiques met en évidence des difficultés de représentation. Ainsi se posent les problèmes, non seulement de représenter des entités identiques selon des perspectives multiples et des décompositions différentes, mais aussi d'ordonner ces entités à partir de relations qui s'apparentent aux relations temporelles, de gérer les incohérences liées à cet ordonnancement et de pouvoir exprimer des relations sémantiques particulières comme celle associée à la notion d'homologie de gènes d'espèces différentes.

### Apic une interface cartographique générique

L'interface cartographique Apic [Bisson& 95] a pour objet de visualiser et de comparer les résultats produits par les méthodes d'analyse de séquences génomiques (recherche de parties codantes, recherche d'homologies...). Elle permet plus généralement d'obtenir une représentation graphique des informations disponibles sur les séquences nucléotidiques. Apic a été développé dans le cadre d'un projet

soutenu par le GIP GREG (cf. Power-Gene, ci-dessous).

La principale originalité d'Apic réside dans sa généricité. Le terme "générique" recouvre ici trois notions différentes. Premièrement, cette interface est multi-organismes, c'est-à-dire qu'elle n'est pas dédiée à l'étude d'un génome particulier. Deuxièmement, elle permet de représenter de manière homogène les différents types de cartes : cytogénétique, génétique, physique, etc. Enfin, cette interface ne se limite pas au simple affichage d'entités biologiques, elle permet également de visualiser les résultats issus de méthodes d'analyse de séquences comme par exemple des courbes ou des images.

Du fait de cette généricité, Apic peut être utilisée dans des domaines autres que celui de la biologie, en particulier dans tous les problèmes où l'on manipule des objets représentables sous la forme de segments orientés (par exemple, des informations temporelles).

### PowerGene: un environnement coopératif d'aide à l'analyse de séquences génomiques

Le projet Sherpa est à l'origine d'un projet soutenu en 1994 et 1995 par le GIP GREG (Groupement de Recherches et d'Études sur les Génomes). Outre le projet Sherpa, la société Ilog et deux laboratoires de biologie sont impliqués :

- · le laboratoire de biométrie, génétique et biologie des populations, équipe "Biométrie Moléculaire" de l'université Claude Bernard de Lyon (CNRS URA 243);
- l'unité "Régulation de l'expression génétique" de l'Institut Pasteur à Paris (CNRS URA 1129).

L'objectif du projet était la réalisation d'un système coopératif d'aide à l'analyse de séquences, qui permette :

- · de mettre en œuvre aisément des méthodes d'analyse de séquences;
- d'aider un utilisateur à choisir la ou les méthodes adéquates pour une tâche donnée et à enchaîner des méthodes entre elles dans le cas de tâches plus complexes;
- · de mémoriser, de gérer et de visualiser à la fois les données de l'analyse et

les résultats produits par l'application de méthodes;

• d'étendre le système en intégrant de nouvelles méthodes et leurs modes

Au terme de ces deux années de travail, un environnement complet, baptisé PowerGene, a été réalisé par le projet Sherpa et la société Ilog. Il est expérimenté sur des bases d'entités, de méthodes et de tâches d'analyse de séquences, construites par les équipes de l'Institut Pasteur [Médigue& 95] et de l'université Claude Bernard.

Cet environnement constitue ainsi une valorisation concrète des travaux de recherche menés au sein du projet Sherpa depuis une demi-douzaine d'années sur la modélisation des connaissances méthodologiques à l'aide de tâches et la résolution coopérative de problèmes [Willamowski

Publications

[Bisson& 95] Gilles Bisson, Alain

Garreau, APIC: a generic interface

for sequencing project, Actes 3rd

ISMB, Cambridge (GB), pp57-65,

(16-19 juillet) 1995

[Dorkeld 94] Frank Dorkeld, MULTI-

MAP: un modèle objet dédié à la car-

tographie comparée des génomes de

mammifères. Thèse de biométrie, uni-

versité Claude Bernard, Lyon, FR,

[Médigue& 95] Claudine Médigue,

Thierry Vermat, Gilles Bisson, Alain

Viari, Antoine Danchin, Cooperative

computer system for genome

sequence analysis, Actes 3rd ISMB,

Cambridge (GB), pp249-258, (16-19)

[Perrière 92] Guy Perrière, Applica-

tion d'une représentation par objets

des connaissances à la modélisation

de certains aspects de l'expression

des gènes chez Escherichia coli.

Thèse de biologie, université Claude

Bernard, Lyon, FR, 1992.

[Perrière & 93] G. Perrière, F. Dor-

keld, F. Rechenmann, C. Gautier,

Object-oriented knowledge bases for

the analysis of prokaryotic and euka-

1994

juillet) 1995

### Vers un environnement réparti d'aide à l'analyse

Les méthodes et les connaissances susceptibles d'être incorporées dans cet environnement résultent des recherches menées dans des équipes et laboratoires de biologie, de biométrie, et d'informatique appliquée à la génétique moléculaire. De nouvelles méthodes, ainsi que des variantes et améliorations de méthodes existantes, sont régulièrement proposées. De plus, l'accroissement des données et des connaissances sur certains organismes vivants, en particulier les orga-"modèles", d'envisager des méthodes de plus en plus spécialisées, car tirant parti des spécificités des génomes de ces orga-

Afin de maintenir un environnement d'aide à l'analyse de séquences au

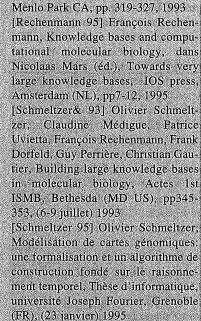
ryotic genomes, actes First Internatio-

nal Conference on Intelligent Systems

for Molecular Biology, AAAI Press,

meilleur niveau de compétence, il faut qu'il puisse exploiter à tout moment les versions les plus à jour des méthodes et des bases de connaissances proposées et développées par ces laboratoires. Il faut donc permettre que des bases méthodologiques spécialisées soient développées et maintenues sur des sites géographiquement distincts, en accord avec leur domaine de compétence, ou avec la disponibilité locale de données ou de ressources logicielles et matérielles spécifiques. Lors de l'analyse d'une séquence particulière, le système peut alors rechercher et exécuter les tâches nécessaires, quelle que soit leur localisation, et de façon complètement transparente pour l'utilisateur. Celuici pourra ainsi suivre et interagir avec le processus d'analyse comme si celui-ci se déroulait localement, tout en bénéficiant des méthodes les plus récentes et les plus efficaces proposées par les différents sites.

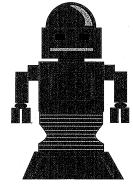
Il s'agit donc de développer un environnement d'analyse de séquences dont les capacités résultent de l'union virtuelle de bases méthodologiques géographiquement réparties, dont la maintenance et l'évolution sont assurées localement. Une première version de l'environnement générique a été obtenue en juin 96. Elle doit être expérimentée avec les bases conçues par les deux laboratoires de biologie et de biométrie qui ont participé au développement de PowerGene dans le



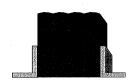
[Willamowski 94] Jutta Willamowski, Modélisation de tâches pour la résolution de problèmes en coopération système-utilisateur, Thèse d'informatique, Université Joseph Fourier, Grenoble (FR), (6 avril) 1994

# de séquences

cadre du projet soutenu par le GIP GREG.



### **LIVRES**



# Livres III

Philippe Breton

À l'image

Seuil

de l'Homme

Du Golem aux créatures virtuelles



### A l'image de l'Homme, du Golem aux créatures virtuelles

**Philippe Breton** Editions du Seuil

# Notes de lecture par Alexis Drogoul (LAFORIA)

Galatée, la statue animée de Pygmalion; le Golem, créé de toutes pièces par des rabbins du Moyen-Age; les automates de la Renaissance (ceux de Vaucanson, par exemple); Olympia, femme mécanique (de Ernst Hoffmann); l'Eve future (de Villiers d'Adam); Pinocchio; la créature du Dr Frankenstein; les robots de Karel Capek; Maria, la femme artificielle de Metropolis; les animaux synthétiques de Grey Walter; l'ordinateur, cerveau artificiel pour les fondateurs de l'IA; les créatures

peuplant les espaces de réalité virtuelle... Pour Philippe Breton, la persistance au cours des siècles de ces êtres créés par l'homme, qui peuplent notre imaginaire ou notre quotidien; n'est sans doute pas incidente, mais révélatrice d'une tendance plus profonde et l'Intelligence Artificielle, pour novatrice qu'elle ait pu paraître en son temps, n'en semble pas moins constituer le dernier avatar d'une longue tradition - celle de la construction de créatures artificielles à l'image de l'homme.

L'auteur s'appuie sur un élément peu exploité jusqu'à présent pour justifier son propos : celui de l'analyse des similitudes (structurelles et conceptuelles) entre les textes " fondateurs " qui mettent en scène ou présentent pour la première fois ces créatures. Il parvient ainsi à isoler un petit nombre d'éléments qui restent constants dans

ces différents récits (le plus savoureux étant le nécessaire " acte magique " qui permet à la créature de prendre son autonomie par rapport à son créateur, et dont je laisse au lecteur le soin de découvrir la forme qu'il prend en I.A.), et à retisser le fil conduisant des textes de l'Antiquité à ceux de nos contemporains (ou presque, puisqu'il s'agit d'Alan Turing, de John von Neumann et de Norbert Wiener), ou du Golem aux créatures virtuelles, pour reprendre le titre de l'ouvrage.

Cette analyse comparée va alors servir de base à ce qui constitue la véritable thèse de l'auteur, à savoir que ces créatures virtuelles, contrefaçons forcément parcellaires, constituent, à chaque époque, dans chaque culture, le reflet de la représenta-

tion que l'homme se fait de lui-même; Galatée exaltait la perfection esthétique, les ordinateurs d'aujourd'hui tentent de simuler l'intelligence et la décision. Autres temps, autres mœurs, mais besoin identique d'exprimer au travers de la créature artificielle les qualités humaines que nous mettons en avant dans la construction de notre civilisation. Comment ne pas s'interroger alors sur le sens de cette quête, et par ricochet sur le véritable projet de l'intelligence artificielle ? La réponse à cette question fournit sans conteste à l'auteur la matière de ses plus belles pages, et de ses plus dérangeantes également. Les avoir lues permet de jeter un regard véritablement neuf sur le "virtuel" dont nous sommes environnés et que nous contribuons, jour après jour, à façonner au travers de nos réalisations. C'est de ce point de vue une réussite étonnante.

Ce livre foisonnant et extrêmement érudit amène, cependant, quelques critiques. Tout d'abord, il s'appuie quasi-exclusive-

ment sur la relation de mythes ou de techniques occidentales. Peut-être cette quête du virtuel constitue-t-elle un trait particulier de notre chère civilisation iudéo-chrétienne, mais il eût été intéressant d'élargir un peu le champ d'investigation, notamment du côté des traditions orientales. Deuxième critique, la seule analyse des textes ou des actes fondateurs de l'intelligence artificielle ne permet pas, à elle seule, de brosser un tableau complet des réalisations ou des ambitions en ce domaine; les tortues de Grey Walter, par exemple, commencent à dater un peu au regard de ce qui se fait maintenant en robotique. Enfin, l'inscription et la signification des créatures artificielles dans le politique et le social, qui ne sont que survolées dans cet ouvrage, sont argumentées de manière un peu trop faible (notamment pour la période moderne, qui ne s'appuie presque que

sur les visions de Norbert Wiener) pour être réellement convaincante.

Mais ceci ne doit pas faire oublier qu'en jetant un regard d'anthropologue sur notre discipline, Philippe Breton nous délivre une formidable leçon d'épistémologie, et replace de ce fait l'intelligence artificielle dans une tradition historique qui change un peu des sempiternelles machines à calculer de Pascal et de Babagge. Il vient également nous rappeler fort à propos que l'informatique est fille de l'intelligence artificielle, et donc d'une quête plus profonde que celle d'un outillage sophistiqué - et non l'inverse, comme on peut avoir tendance à le penser de nos jours...

### **LIVRES**

# **Introduction to**

### **Knowledge Systems**, Mark Stefik,

Morgan Kaufmann Publishers, Inc., San Francisco, 1995.

### Notes de lecture par Amedeo Napoli (CRIN)

Mensurations: environ 1,7 kg pour 870 pages. Contenu: 9 chapitres pour 3 grandes parties, une bibliographie commentée, des exercices, leurs corrigés, un résumé pour chaque sous-chapitre, une liste de problèmes ouverts pour chaque chapitre ...

C'est un nouveau livre non pas sur l'intelligence artificielle en général, mais sur la conception et l'analyse de systèmes à base de connaissances — peut-être, à l'image de l'auteur, faut-il maintenant dire plus simplement systèmes cognitifs — et il s'agit véritablement du premier grand livre sur le sujet. Le livre est impressionnant du point de vue de la taille comme du point de vue du contenu. La façon de présenter n'est pas classique et suit plutôt une approche pragmatique. Ainsi, une petite histoire, riche d'enseignements, sert à introduire le livre et explique la démarche de l'auteur. La première partie introduit les fondations théoriques et pratiques des systèmes cognitifs. Dans le premier chapitre sont abordés les systèmes formels, les notions de symboles et d'organisation de symboles, les notions de syntaxe et de sémantique — les différentes sémantiques possibles — d'un système formel, ce que doit être une représentation, les diverses formes qu'elle peut prendre, ce que recouvre la notion d'inférence, et les formalismes possibles pouvant être employés pour concevoir des systèmes cognitifs. Le second chapitre introduit le problème de la recherche dans un graphe ou un espace d'états, et plus généralement la problématique de la résolution de problèmes, en proposant des algorithmes classiques et d'autres méthodes moins classiques (recherche hiérarchique). Le troisième chapitre aborde l'ingénierie de la connaissance et propose différentes méthodes d'acquisition de connaissances : comment caractériser les connaissances qu'il faut associer à un système et comment les organiser. Le célèbre système MYCIN est pour l'occasion analysé du point de vue de l'ingénierie des connaissances, et sera repris comme exemple dans les chapitres suivants. De très nombreux exemples, généralement bien choisis, illustrent le propos. Autre particularité, une discussion clôt chaque chapitre et présente les problèmes ouverts, ou, comme pour le chapitre 3, discute des idées reçues et des idées fausses sur les systèmes cognitifs et les problèmes d'ingénierie de la connaissance (est-ce à dire que les idées fausses circulent plus facilement dans ce domaine précis ?).

La seconde partie du livre traite de ce qui est appelé le niveau des symboles, qui est le niveau relatif à la caractérisation des algorithmes et des méthodes de représentation, par opposition au niveau de la connaissance ou niveau cognitif — la troisième partie du livre — où les systèmes cognitifs sont appréhendés du point de vue de la tâche à accomplir et de l'exploitation des connaissances pour arriver à cette fin. Les chapitres 4 et 5 sont courts, une trentaine de pages chacun, et traitent respectivement du temps et de l'espace, de façon relativement parallèle, au niveau de la représentation et des contraintes. Le chapitre 6 est dédié au traitement de l'incertitude et de l'imprécision, introduit brièvement des éléments de probabilités, de théorie de la décision, l'approche de Dempster-Shafer, la théorie des ensembles flous. Là encore, MYCIN et ses coefficients de plausibilité, entres autres, servent d'exemple.

La partie 3 commence par un chapitre sur la classification, et c'est d'ailleurs, je dois l'avouer, ce qui m'a de suite attiré l'œil lors de ma première consultation de l'ouvrage : enfin, un chapitre dédié à la classification dans un ouvrage sur les systèmes cognitifs (notons au passage que le livre de Stefik a été devancé, sur le sujet, par un ouvrage francophone sur le raisonnement, paru il y a maintenant 5 ans ...). Les structures du chapitre 7 sur la classification et celle du chapitre 8 sur la configuration sont identiques : la classification (la configuration) et ses différentes formes sont analysées dans un premier temps, puis est décrit un certain nombre de systèmes dont le fonctionnement repose essentiellement sur la classification (la configuration), et ensuite sont donnés des algorithmes génériques de classification (de configuration), qui expliquent les comportements des systèmes précédemment présentés. La classification est vue comme un processus permettant d'identifier la classe d'un élément dans un espace de classes donné à l'avance, dans une perspective de résolution de problèmes. La configuration quant à elle consiste à trouver l'arrangement d'un certain nombre de composants, toujours pour résoudre un problème, et, cette fois, l'espace des configurations — par opposition à celui des classes possibles - n'est pas connu a priori et peut être infini. Classification et configuration peuvent apparaître comme des méthodes de résolution de problèmes duales et la structure des chapitres permet de mener assez facilement une étude comparative. Le neuvième et dernier chapitre est dédié au diagnostic et à la détection de pannes dans les systèmes physiques. Sont abordées les notions d'entropie et de raisonnement à base de modèles, ainsi que la causalité. Là encore, des systèmes opérationnels sont passés en revue, détaillés et comparés. En particulier, des figures très explicites permettent de comprendre le comportement de ces systèmes et de comparer les différentes démarches, classification, configuration et diagnostic.

### **LIVRES**

Sur un plan global, le livre est très bien fait et il est impressionnant sur bien des aspects : la nouveauté du discours, la clarté du propos, le souci pédagogique constant — mis en valeur par les nombreux exemples, figures, et exercices ainsi que la rigueur scientifique font partie des caractéristiques de ce livre. Le livre renferme une importante masse d'informations et a du représenter un énorme travail pour son auteur. A côté des points positifs, mentionnons quelques regrets plutôt que points négatifs : le livre passe sous silence les systèmes d'apprentissage et toute la problématique qui en relève ou qui en est proche, en particulier le raisonnement à partir de cas, les systèmes de découverte de

connaissances dans les bases de données, qui sont pourtant des travaux d'actualité (au passage, le lecteur frustré par le manque de référence à l'apprentissage pourra consulter avec profit et plaisir le récent ouvrage de P. Langley, Elements of Machine Learning, chez Morgan Kaufmann Publishers, 1996). Pour terminer, je ne saurai que trop conseiller à tous les enseignants d'intelligence artificielle d'avoir la curiosité d'aller voir comment sont traités certains sujets dans ce livre : force est de constater qu'un tournant a été pris et qu'est né le premier d'une nouvelle génération d'ouvrages sur l'intelligence artificielle en général et les systèmes cognitifs en particulier.



### Du côté des'éditeurs

Plusieurs éditeurs ont bien voulu nous faire parvenir leurs parutions en Intelligence Artificielle. Ces livres sont à la disposition de ceux de nos lecteurs qui accepteraient d'en faire une critique, contacter Philippe Laublet par E-Mail de préférence, (voir adresses en page 2). Les livres disponibles actuellement sont :

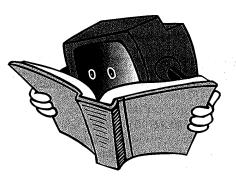
L.Arditti: La programmation une approche fonctionnelle et récursive avec Scheme, Eyrolles 1996

N.Aussenac, P.Laublet, C.Revnaud (eds): Acqusition et ingéniérie des connaissances, Cepadues 1996,

S.Goonatilake and P.Treleaven (eds): Intelligent Systems for Finance and Business. John Wiley.

V.Prince: Vers une informatique cognitive dans les organisations, Le role central du langage, Editions Masson 1996

G.Tisseau: Intelligence artificielle, problèmes et méthodes, PUF 1996 THE PROPERTY OF STREET



### Vous avez aimé un livre

Vous voulez faire partager votre intérêt. Vous pouvez nous faire parvenir une courte note, de quelques lignes, indiquant en quoi il vous semble digne d'attention pour les membres de l'AFIA. Nous la publierons volontiers.



# CONFÉRENCES



# CONFÉRENCES

### JAVA'96

Sète, 8-10 mai 1996

Devenues régulières, les journées JAVA ont eu lieu a Sète cette année, regroupant les Journées d'acquisition des connaissances (JAC) et les journées francophones d'apprentissage. (JFA). Elles étaient organisées par Joël Quinqueton et les membres du LIRMM (Montpellier).



### **JAC'96** 7° Journées d'Acquisition des Connaissances

Sète, 8-10 mai 1996

# K. Causse, A. Condamines, J. Charlet, G. Kassel, M. Zacklad

Les Journées d'Acquisition des Connaissances 96, septièmes du genre, ont réuni les 9 et 10 mai à Sète une trentaine de participants. Elles ont permis d'approfondir des questions déjà abordées lors de précédentes éditions, comme l'acquisition de connaissances ontologiques, ou l'acquisition de connaissances à partir de textes. Elles ont également été mises à profit pour faire le point sur les acquis et les évolutions du domaine.

En ouverture des journées, B. Bachimont a abordé au cours de sa conférence invitée les problèmes d'ordre méthodologique et théorique que pose l'utilisation d'ontologies en IA. S'appuyant sur une expérience de constitution d'une ontologie médicale dans le domaine de la coronarographie cardiaque, B. Bachimont propose de distinguer soigneusement deux étapes : une étape de construction des primitives de modélisation d'un domaine, aboutissant à une « ontologie régionale », suivie d'une étape de codage dans un langage formel conduisant à une ontologie formelle; cette dernière sera finalement opérationnalisée avec un langage de programmation. Ce découpage se justifie selon B. Bachimont par un changement de type de sémantique. L'ontologie régionale se situe au niveau du linguistique. Elle se présente sous la forme d'un arbre d'unités linguistiques reposant sur une sémantique différentielle; l'ontologie formelle repose sur une sémantique logique de type référentielle. L'approche proposée souligne l'importance du niveau linguistique dans l'acquisition des connaissances. La discussion qui a suivi a permis d'éclaircir

le sens de certains termes comme la notion « d'engagement ontologique », présente dans les travaux de Guarino, et a tâché de replacer l'acquisition d'ontologies de domaine dans le processus de développement de SBC.

La matinée s'est terminée par un exposé de R. Loisel sur la conception architecturale, un article co-signé avec C. Gréboval-Barry. L'article présente la constitution d'une expertise partielle dans le but de construire un outil-partenaire pour l'architecte en phase de pré-conception et de conception d'un projet architectural. Les auteurs ont mis l'accent sur la difficulté de modéliser une activité par essence de nature créative : la conception est dite innovative. La modélisation proposée distingue trois modèles : le modèle créationnel représente les connaissances mises directement en jeu dans le projet en cours de création (notion de modèle du cas) ; le modèle décisionnel comporte un certain nombre de connaissances de nature stratégique permettant de diriger le processus de création; le module instrumental comporte les outils et décrit les procédés permettant la mise en œuvre de la conception. L'ensemble de ces connaissances permet à l'architecte de construire progressivement des images conceptuelles, représentation de la façon dont l'architecte pense le projet. L'outilpartenaire doit pouvoir s'enrichir de l'interaction avec r l'architecte en étendant ses connaissances en cours de conception. Pour permettre à des connaissances complexes, parfois contradictoires, d'être exprimées, les auteurs ont introduit la notion « d'archigranule dialogique », composant élémentaire de connaissance qui s'organise en un réseau dynamique évoluant en cours de conception.

En début d'après-midi une table ronde organisée par P. Laublet fut l'occasion de réfléchir à l'avenir de l'acquisition des connaissances. Ce domaine peut se prévaloir d'un certain nombre de résultats, à commencer par les descriptions de connaissances et du comportement de SBC dites « au niveau connaissance » ; il a contribué à faire progresser la théorie de la résolution de problèmes, en recensant notamment de nombreuses méthodes de résolution dédiées à différents types de problèmes. Aujourd'hui l'acquisition des connaissances demeure un domaine actif de recherches autour de questions comme le partage des connaissances et des ontologies, ou la définition de nouveaux langages de modélisation et de représentation des connaissances. Cependant son champ applicatif industriel change: on acquiert moins pour construire des résolveurs de problèmes que pour capitaliser, par exemple, des connaissances d'entreprise. Les supports techniques changent également (cf. le Web). A l'avenir, le domaine devra prendre en compte ces évolutions.

. Une session « Langages, formalisation, opérationnalisation » clôturait la première journée. Une première communication présentée par I. Delouis-Jacob, et co-signée par O. Jehl, a permis de faire le point sur l'évolution industrielle du langage LISA, dans le cadre du projet AUSTRAL à la DER de l'EDF (Clamart). Rappelons que le langage LISA vise à opérationnaliser des modèles conceptuels de type KADS tout en permettant la réflexivité, cette dernière étant exploitée pour concevoir des SBC coopératifs. Les forces, mais également les faiblesses, de LISA étant décortiquées, son évolution en un langage nommé LISA@RT est décrite, mettant en avant : un code compilé, et non plus interprété, un environnement de

### **C**ONFÉRENCES

développement sophistiqué empruntant des idées du Génie Logiciel. Cet environnement est construit au dessus des logiciels de la société ILOG. La seconde communication, présentée par C. Serres et co-signée par J.G. Ganascia, portait également sur l'apport des techniques du Génie Logiciel à l'ingénierie des connaissances. La notion de module, importée du Génie Logiciel, est ici utilisée pour améliorer la structuration de modèles d'expertise de type KADS et constitue, au dire des auteurs, une aide à la modélisation des connaissances; les modules sont formalisés à l'aide de types abstraits algébriques. Ce travail a fait l'objet de la thèse de C. Serres. La première session du vendredi « Acquisition de connaissances ontologiques » comportait deux exposés, l'un, « Construction coopérative d'ontologies dans un cadre multiexpertise : ébauche d'un outil » de C. Garcia, l'autre, « Eléments d'une classification des approches ontologiques » de B. Leroux. Le premier exposé présentait une proposition de méthodologie pour permettre à plusieurs experts de coopérer à la construction d'une ontologie dans le domaine de l'accidentologie. Le second exposé s'intéressait aux approches ontologiques et argumentait le fait que la réutilisabilité des ontologies et la réutilisabilité des Modèles de Résolution de Problèmes étaient deux aspects d'un même problème. Tout cela était le prélude à la poursuite de la discussion initiée au sujet des engagements ontologiques impliqués par le choix

La session « Acquisition de Connaissances à partir de textes » était composée de deux articles qui mettaient bien en évidence les deux approches qui péuvent être suivies dans ce processus d'acquisition.

Le premier article, signé par D. Garcia, N. Aussenac-Gilles et A. Courcelle et présenté par D. Garcia « Exploitation, pour la modélisation, des connaissances causales détectées par COA-TIS » a concerné une méthode d'acquisition à la fois de termes dénommant des actions et de relations de cause, basée sur l'utilisation d'indices linguistiques. La présentation a, en outre, montré comment les résultats fournis par COATIS ont été mis en œuvre dans un projet d'aide à la consultation de documents techniques.

Le second article, signé par B. Habert et A. Nazarenko et présenté par A. Nazarenko « La syntaxe comme marchepied de l'acquisition des connaissances : bilan critique d'une expérience » visait à présenter un outil qui donne une image du texte à travers la constitution d'un réseau linguistique mettant en relation les noms qui apparaissent dans des contextes similaires. Après une étape de normalisation (réduction à des formes N Adj ou N de N), sont repérées les expansions communes à des têtes différentes; lorsque deux têtes partagent plus de trois constructions, elles sont mises

La première présentation a proposé une approche descendante qui va de la langue aux langues spécialisées (avec l'idée qu'un certain nombre de phénomènes linguistiques se retrouve dans des domaines spécialisés). La deuxième présentation a proposé une approche ascendante qui permet de rendre compte de ce qui est spécifique à une langue spécialisée (elle s'apparente ainsi aux travaux de type harrissien sur les sous-langages). La discussion qui a suivi a montré que ces approches pouvaient être complémentaires. Elle a aussi

témoigné de la convergence de plus en plus nette qui se fait jour entre les travaux sur les ontologies et les travaux sur les analyses linguistiques de textes spécialisés, faisant écho en cela à la conférence de B. Bachimont.

Malheureusement réduite à une seule intervention, la session processus d'acquisition a donné lieu à des discussions intéressantes faisant suite à la présentation du papier de B. Le Roux et J. Thomas « Raffinement d'une méthode de résolution de problèmes et engagements ontologiques ». Les travaux présentés sont basés sur le concept de GDM (Generalized Directive Model), introduit initialement dans le projet Acknowledge, qui consiste à analyser le raffinement des structures d'inférence comme un processus de réécriture grammaticale soumis à des conditions relatives aux caractéristiques du domaine modélisé. En s'appuyant sur ces recherches, les auteurs développent sur quelques exemples une analyse originale des caractéristiques qui guident le raffinement des structures d'inférence et justifient les choix de conception de l'outil Tomaok développé par le premier auteur selon un principe GDM.

Pour l'an prochain, les JAC expérimenteront une nouvelle formule basée sur un positionnement scientifique différent de la discipline. L'ambition est de situer l'acquisition au cœur de l'ingénierie des connaissances, une problématique plus vaste tant du point de vue de la diversité des sources<sup>4</sup> de connaissances prises en considération pour construire les modèles que du point de vue des fonctionnalités des systèmes cibles. Mais ne déflorons pas le suspens du prochain appel à communication d'Ingénierie des Connaissances 97 ...



### Conférences



### JFA'96 11° Journées Françaises de l'Apprentissage

Sète, 8-10 mai 1996

### Olivier Gascuel

Cette édition des Journées Françaises de l'Apprentissage ou JFA était la onzième du nom. Organisées pour la première fois par Yves Kodratoff en 1986, les JFA ont évolué progressivement. Elles étaient initialement très liées à L'intelligence Artificielle symbolique. Elles se sont par la suite ouvertes à d'autres approches, notamment numériques, comme les Réseaux de Neurones, mais aussi liées à la Reconnaissance des Formes et en particulier l'Inférence Grammaticale. Des travaux de nature théorique, prenant appui sur les modèles d'apprentissage définis par Valiant et ses successeurs, ont été présentés lors des dernières éditions. Dans le même temps, les JFA ont toujours donné une place importante aux applications. Nous avons essayé cette année de préserver ces acquis, tout en couvrant le plus largement possible les différents domaines de l'Apprentissage. Les membres du comité de programme reflétaient bien cette diversité, étant issus et étant les représentants de communauté scientifiques distinctes: Intelligence Artificielle, Réseaux de Neurones, Reconnaissance des Formes et Informatique Théorique. Les conférences invitées d'Alain Viari sur l'analyse des données biologiques, et d'Eric Martin sur la modélisation de la découverte scientifique, indiquaient bien la volonté des JFA de faire cohabiter des préoccupations théoriques et appliquées. Enfin, on a pu noter le large spectre couvert par les différentes sessions : Inférence Grammaticale, Application à la Biologie, Aspects théoriques, Approches connexionistes et Approches génétiques. Cette volonté d'ouverture se poursuivra très certainement dans les années à venir, afin que soient représentées aux JFA les très nombreuses facettes de ce domaine de recherche vaste mais encore largement inexploré qu'est l'Apprentissage.

numéro 26 / juin 1996

### Journées Francophones "Intelligence Artificielle Distribuée et Systèmes multi-agents"

1er-3 Avril 1996 - Port Camargue

### Compte-rendu et réflexions Alexis Drogoul, LAFORIA

Les JFIADSMA, dont c'était la quatrième édition, ont pour vocation de rassembler les chercheurs francophones en Intelligence Artificielle Distribuée, afin de faire le point sur les avancées dans le domaine. L'organisation, assurée cette année par le LIRMM, était réellement parfaite, mais on pouvait cependant reprocher à ceux qui l'ont prise en charge deux choses : la première étant que rien n'ait été prévu pour le 1er Avril (pas de conférences farfelues, pas de fausses annonces - ou alors, elles sont passées inaperçues), la seconde que le programme n'ait pas été mieux réparti sur les trois jours, la dernière journée n'ayant été consacrée qu'à la conférence invitée de Humberto Maturana, et aux comptes-rendus des groupes de travail, alors que les deux premières ont été très denses (20 présentations).

Les travaux présentés se sont montrés d'une grande diversité, ce qui est la preuve du dynamisme de la communauté multi-agent, mais également d'une certaine difficulté à la circonscrire. Aucun thème commun n'a donc réellement émergé, si ce n'est une sorte de consensus "par défaut" sur le fait qu'un agent ne peut plus s'appréhender que comme coopératif, intentionnel, doué de motivations propres et de capacités de communication très importantes. Ceci conduit tout naturellement à ne s'intéresser qu'à des structures organisationnelles explicites, de type hiérarchique, coopératif, fondé sur des réseaux de dépendances (exposé de Brahim Chaib-Draa, sur la classification des structures relationnelles d'activité) ou autres. L'idée d'auto-organisation n'a été évoquée et défendue que par Jean-Pierre Müller (très intéressant exposé sur l'utilisation d'agents réactifs appliqués à la déduction automatique de théorèmes), Youssef Hamadi (sur la distribution de la procédure GSAT) et, dans une moindre mesure, Alain Boucher (sur la segmentation de séquences d'images) et Bruno Bouzy (application au jeu de Go).

gique et économique qui a dominé, que ce soit en termes d'organisation (exposés d'Olivier Garro, qui présentait un aperçu des processus de conception distribuée dans l'industrie, de Christine Piquemal, avec un système de diagnostic médical proche de l'approche multi-experts, de Yacine Ouzrout, sur la prise de décision dans les organisations productives, de Zahia Guessoum, qui présentait un système de modélisation de l'activité économique, de Marie-Pierre Gleizes, sur la proposition d'implémenter une attitude coopérative parmi des agents résolvant un problème-jouet, de Flavien Balbo, sur une approche multi-critères de l'allocation de tâches, ou d'Emmanuelle Le Strugeon, qui proposait un mécanisme permettant à une société d'agents de passer d'un type d'organisation à un autre suivant le contexte) ou de communication (exposé de James de Almeida et Nicolas Gregori sur la logique interlocutoire, qui se propose de dépasser les limites actuelles de la théorie des actes de lan-

Sinon, c'est surtout la métaphore sociolo-

Peu de travaux se sont cependant préoccupés de la nécessité qu'il y avait, dans les applications proposées, à utiliser précisément ce type de métaphores, qui conduit à effectuer des choix coûteux en termes de complexité d'architecture d'agent. Seuls, peut-être, les deux exposés traitant d'architectures temps-réel (Michel Occello et Thierry Salvant) ont mis en avant des problèmes dont la résolution exigeait effectivement à la fois une forte réactivité et une capacité individuelle de contrôle ou de prévision à moyen terme, justifiant de ce fait l'approche cognitive qu'ils proposaient.

Mais, à l'heure où les sources d'information autorisées nous abreuvent continûment d'"agents intelligents" mis à toutes les sauces, il n'est peut-être pas inutile de s'interroger de manière plus générale que cela sur le statut ontologique de l'agent (intelligent ou non, d'ailleurs) ... Est-ce seulement un terme vague appelé à remplacer le vocable "système" trop connoté "système expert" ? Quels sont les attributs objectifs qui permettent de décerner ce nom à une structure informatique? A quoi cela sert-il d'appeler agent cette structure? A quoi cela sert-il de lui attribuer des intentions ou des motivations? Deux exposés se sont préoccupés, chacun à leur manière, de ces questions. Celui

### **C**ONFÉRENCES

d'Eric Batard, tout d'abord, qui, à partir d'une approche sémiotique de l'IAD, faisait un point assez complet du rôle représentationnel de l'agent et des problèmes qu'il soulève, et celui d'Anne Collinot, qui, dans sa présentation de la méthode de conception Cassiopée, tentait de faire le point sur les rapports existant entre l'agent comme entité d'analyse d'un problème à résoudre, et l'agent comme entité d'implémentation du système qui le résout, en proposant une vision intégratrice de l'ensemble des techniques (de coopération, de coordination, etc.) existantes.

Se poser ce type de questions n'est pas inutile. Je souhaiterai en effet terminer ce compte-rendu succinct par quelques réflexions personnelles, qui n'engagent donc que l'auteur de ces modestes lignes. Tout d'abord, la diversité des présentations ne doit pas masquer le repli de la plupart des travaux sur des valeurs sûres, qui structurent déjà solidement le domaine aux Etats-Unis: formalisation de la coopération, formalisation de la communication (actes de langage), formalisation de la représentation (Croyances -Désirs - Intentions...), applications industrielles. De ce côté-là, il ne semble guère plus y avoir de spécificité européenne, et à fortiori française, au niveau des thèmes dominants. Si, aux Etats-Unis, la plupart des travaux "alternatifs" (agents réactifs, auto-organisation, etc.) ont trouvé leur place dans d'autres disciplines comme la Vie Artificielle, ils sont en France relativement marginalisés au sein de l'IAD, en raison essentiellement de l'absence institutionnelle de ces autres disciplines. Il en résulte un appauvrissement certain des débats (je pense notamment au débat réactif-cognitif, qui a connu ses heures de gloire lors des premiers workshops MAAMAW), puisque tout le monde semble se mettre peu ou prou d'accord sur une "pensée unique" un peu tristounette. La seule tentative de débat dûment organisée pendant ces journées ne portait d'ailleurs même plus sur le projet de recherche de l'Intelligence Artificielle Distribuée, mais sur un point terminologique fort obscur (distinction entre la coopération, réservée aux "agents intelligents" et la coaction, aux autres, initiée par Christian Brassac et Sylvie Pesty, avec deux réponses de Jacques Ferber et Charles Lenay), dont on pouvait se demander ce qu'il apportait, la question

ayant été tranchée depuis une dizaine d'années en Vie Artificielle (voir par exemple l'introduction de Christopher Langton au premier volume Artificial Life, et l'article de Rodney Brooks "Intelligence Without Reason" dans les actes de l'IJCAI'91) ou en philosophie de l'esprit (par Daniel Denett, par exemple, et sa fameuse "posture intentionnelle").

Cette absence de débat que, faute de mieux, l'on pourrait qualifier d'épistémologique est cependant assez normale. La teneur des travaux présentés était là pour le prouver: la discipline est clairement entrée dans une phase opérationnelle, essentiellement à destination de l'industrie et des télécommunications. C'est une bonne chose, qui tient sans doute un peu de l'effet de mode, mais qui ne permet pas, en même temps, d'adopter le recul nécessaire à une justification des choix conceptuels effectués. Seul l'avenir (et les prochaines journées francophones...) dira si les bonnes applications présentées suffisent à justifier l'approche actuelle.



### **JIM 96** Troisièmes Journées d'Informatique Musicale

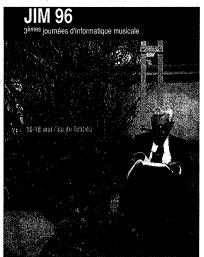
île de Tatihou, 16-18 mai 1996

### Marc Chemillier, François Pachet

Les troisièmes journées d'informatique musicale ont été organisées cette année par le laboratoire d'informatique de l'Université de Caen (GREYC). Elles se sont déroulées du 16 au 18 mai dernier, dans un site tout à fait unique : l'île de Tatihou, dans le département de la Manche, au large (relatif) de Saint-Vaast la Hougue.

L'affluence a atteint des records durant ces journées, le nombre de participants oscillant entre 60 et 90 selon les divers moments du congrès. De nombreux pays étaient représentés, car certains avaient fait le déplacement depuis les États-Unis, le Brésil, la Turquie, la Hollande, l'Allemagne, Israël, etc. pour se rendre à Tatihou. Le comité d'organisation comprenait Marc CHEMILLIER du GREYC (Caen) et Gérard ASSAYAG

de l'IRCAM (Paris) pour la partie scientifique, ainsi que Christian ELOY de l'INA/GRM (Paris) pour la partie musicale. C'est par Internet qu'un appel à communications scientifiques et à œuvres musicales avait été lancé, depuis l'IRCAM, et les propositions avaient été évaluées par un jury d'experts internatio-



Les six demi-journées du colloque ont permis d'entendre une trentaine de communications de haut niveau, présentant toutes les facettes de l'informatique musicale. Le premier tirage des actes du colloque, volumineux document préparé par l'équipe technique du GREYC, est sur le point d'être épuisé.

Le vendredi 17, une conférence de Jean-Claude RISSET (CNRS, Marseille) intitulée "L'informatique musicale a 40 ans. Ouelques remarques sur l'élaboration du son" a permis à la nombreuse assistance présente, chercheurs confirmés mais aussi étudiants de troisième cycle, de parcourir les différentes étapes de l'histoire de l'informatique musicale depuis ses origines. L'orateur, lui-même pionnier de ce domaine en France dans les années soixante, a été salué par des applaudissements enthousiastes.

Deux concerts étaient prévus, le jeudi soir et le vendredi soir. Un "Hommage à Pierre SCHAEFFER" a permis d'entendre, non sans émotion, la voix de ce polytechnicien-musicien inventeur de la musique concrète, décédé le 19 août dernier à l'âge de 85 ans, Parmi les pièces proposées lors de ces concerts, on remarquait une composition de James GIROU-DON interprétée par le contrebassiste virtuose Jean-Pierre ROBERT, dans

### **C**ONFÉRENCES

laquelle la contrebasse dialoguait avec une autre contrebasse, absente au concert, mais dont la partie avait été enregistrée et traitée sur ordinateur. Une autre pièce composée par Georges BLOCH, pour alto et spacialisateur, montrait les possibilités de la machine Next de l'IRCAM de calculer en temps réel un effet de réverbération correspondant à une salle de concert virtuelle. Les concerts se déroulaient dans une atmosphère à la fois chaleureuse, détendue et attentive.

Entre les conférences et les concerts, les participants étaient conviés à des présentations diverses : méta-instrument de Serge de LAUBIER, synthèse par disques photosoniques de Jacques DUDON, etc. Dans le château d'eau, le peintre Yves CHAUDOUET organisait des projections vidéos.

Il y a eu beaucoup de vent et de pluie durant ces trois jours passés à l'île de Tatihou. Mais loin de contrarier les participants, ces conditions météorologiques mettaient en relief le cachet exceptionnel du site, et contribuaient ainsi à la profonde impression que leur a laissé leur séjour dans l'île. Des images vidéos ont été filmées durant cette manifestation, et un projet de film est à l'étude. Une publication des actes est également envisagée. Un mot enfin sur l'île. Une fois arrivés à Saint-Vaast, où l'on se gare au port, on se dirige vers l'embarquement pour prendre un bateau, car l'île de Tatihou est bien une île. Tout va bien, nous sommes plongés dans nos pensées musicales, jusqu'au moment ou l'on s'aperçoit que le bateau en question est amarré sur le sec. On a beau regarder vers le large, point d'eau, que des rochers couverts d'algues, au loin des mouettes et des casiers à huîtres. La réponse à cette énigme nous vient par petits bouts : le bateau est tout simplement pourvu de roues, de taille impressionnante, et équipées de pneus énormes et très dégonflés, qui lui donnent la capacité de se déplacer sur la plage abandonnée (coquillages et crustacés) et sur les cailloux, comme dans un rêve, ou un film de James Bond. Ce bateau visiblement taillé sur mesure pour cette configuration toute particulière nous emmène après un dizaine de minute de "traversée" sur une île battue par les vents sur laquelle se trouve un fort, réhabilité et transformé en centre culturel et équipé de salles tout à fait modernes et d'un hôtel/restaurant. Avis aux organisateurs cherchant un beau coin isolé pour une réunion : la salle de conférence peut accueillir une centaine de personnes, ainsi que le restaurant; 55 personnes maximum peuvent rester dormir sur l'île, et il y a des hôtels dans la région pour l'éventuel surplus, en particulier à Valogne.

Mais pourquoi un bateau, après tout ? C'est qu'une fois la conférence et les concerts terminés, vers minuit, le bateau nous ramène sur la terre. Mais cette fois c'est marée haute. Notre bateau, après avoir roulé quelques mètres sur le sable, retrouve alors l'eau et sa fonction originelle. Le voyage de retour se fait en pleine nuit, sous une mer démontée, et les vagues menaçantes nous aspergent littéralement. Arrivés au port, trempés (ainsi que nos "proceedings" que nous avons gardés sous le bras pendant toute la traversée), nous n'avons plus qu'une seule envie : celle de reprendre 1' " hyper-bateau ", puisque c'est ainsi que nous l'avons baptisé en hommage aux hyper-instruments cités plus haut, pour écouter la suite des JIM.



### **JSJ'96** Premières Journées du **Séminaire Junior**

Villetaneuse, 21-22 Mars 1996

### Serge Bibas, Nathalie **Bouquet, Fabrice Issac**

JSJ'96, les Premières Journées du Séminaire Junior ont été organisées par les doctorants du Laboratoire d'Informatique de Paris-Nord (L.I.P.N.), sous le parrainage de l'AFIA, de l'Université Paris-Nord et de l'École Doctorale de l'Institut Galilée.

Le Séminaire Junior du L.I.P.N. a été créé pour permettre aux doctorants de présenter leurs travaux aux autres membres du laboratoire.

Afin d'établir et de consolider des relations scientifiques entre doctorants travaillant dans des domaines de recherche proches, les organisateurs ont décidé d'étendre ce séminaire aux

autres laboratoires de recherche en Informatique. C'est dans cette optique qu'ont été organisées les premières Journées du Séminaire Junior.

Ces Journées étaient placées sous le signe de la discussion : elles comportaient quatre tables rondes, chacune portant sur un domaine de recherche de l'intelligence artificielle, au cours desquelles une place importante était laissée aux échanges souvent fructueux entre les 42 participants venus des "quatre" coins de l'hexagone. La difficile mission d'animer ces sessions était confiée à des chercheurs confirmés du laboratoire. Afin de situer les différentes interventions proposées, les tables rondes débutaient par une présentation générale du domaine de recherche abordé.

Ces Journées ont été inaugurées par la session Satisfaction de Contraintes, animée par Hachemi Bennaceur. La recherche actuelle sur les CSP se développe essentiellement autour de quatre grands axes: le filtrage, l'exploration ou la recherche de solutions, la décomposition afin de borner la complexité théorique voir pratique, et la détection des classes polynomiales. Deux de ces axes sont abordés dans cette session: le filtrage et la résolution. Parmi les questions soulevées, citons la généralisation des méthodes de résolution des CSP binaires pour des contraintes non bijectives (contraintes de différence, domaines symétriques), la détermination des causes d'échecs lors de la résolution et le choix des méta -heuristiques, ainsi que leurs réglages, pour les problèmes d'optimisation en

La seconde session, conjointement animée par Brigitte Biébow et Jacqueline Castaing, portait sur la Représentation des Connaissances et le Raisonnement. La session a débuté par un tour d'horizon des différents thèmes de ce vaste domaine en une répartition volontairement sujette à discussion. A l'image de l'hétérogénéité du domaine, les interventions étaient très diverses. Le thème le plus représenté était celui des formalismes de représentation avec une prédominance remarquée pour l'utilisation du formalisme des graphes conceptuels. Il a aussi été question d'analyse de donnée et de raisonnement : raisonnement

### **C**ONFÉRENCES

par abstraction sur des représentations par objets et utilisation de logique linéaire pour la démonstration automa-

La fin de la première journée a été marquée par le tutoriel de Daniel Kayser, Professeur à l'université Paris-Nord, sur l'"Histoire de l'Intelligence Artificielle", qui a rappelé que l'I.A. résulte d'une évolution bien antérieure à l'apparition des ordinateurs. Les conceptions dominantes ont été retracées ainsi que certains travaux intéressants, aujourd'hui peut-être un peu oubliés, qu'elles ont pu masquer. Plusieurs leçons à tirer de ce rappel : elles concernent le rôle des métaphores en I.A., la relativement faible variation du domaine qu'elle couvre, et l'actualité du défi auquel elle s'attaque.

Les deux sessions du lendemain portaient respectivement sur le Traitement Automatique du Langage Naturel (TALN) et l'Apprentissage Symbolique et Numérique.

La première table ronde, animée par Françoise Gayral, a débuté par une présentation du large éventail de problèmes soulevés par le TALN. La résolution de ces problèmes est bien sûr liée à l'objectif à atteindre, ces objectifs pouvant se différencier par la nature de l'application à réaliser ou par la nature du phénomène linguistique pris en compte. Le linguiste décrit des mécanismes de langue qu'il peut valider en utilisant des techniques linguistiques telles que les transformations ou les paraphrases. En revanche, l'informaticien doit élaborer des techniques effectives ayant recours à des connaissances implémentables auxquelles il doit gérer l'accès. Cette diversité était présente dans les interventions qui suivirent, dont les préoccupations étaient morpho-syntaxiques, sémantiques et pragmatiques, avec différents objectifs allant de l'étude de phénomènes linguistiques à la traduction automatique, traitant même de différentes langues comme le tahitien ou l'arabe.

Les Journées se sont achevées par la session "Apprentissage Symbolique et Numérique" conjointement animée par Younès Bennani pour l'approche numérique et Marc Champesme pour l'approche symbolique. Si les deux thèmes ont le même souci de généralisation et de prédiction, l'approche symbolique

accorde une grande importance à l'aspect représentation des connaissances. Un seul des exposés s'intéressait au problème de l'apprenabilité, les autres considérant l'apprentissage comme un outil dans un cadre global (analyse de données, data mining, système EAO ou CSP). Quant à la "partie" numérique, le thème principalement discuté fut le contrôle de la complexité des architectures neuronales, ainsi que le lien existant avec les statistiques et notamment la branche concernant l'estimation paramétrique. Il a aussi été rappelé que le connexionnisme n'a pas comme application unique la reconnaissance de carac-

Les objectifs de ces Journées étaient non seulement d'établir des relations entre doctorants, de permettre des discussions, mais aussi de susciter l'intérêt de chacun à de nouveau domaines de l'intelligence artificielle. À chaque table ronde étaient présents des jeunes chercheurs de domaines différents. Ceux du domaine en question ont ainsi pu profiter du point de vue des nonspécialistes et ces derniers ont pu avoir un aperçu de domaines qui leur étaient moins connus.



### **JCSC'96** 2° colloque Jeunes **Chercheurs en Sciences Cognitives**

Presqu'île de Giens, 5,-7 juin 1996

### Cognito Ergo Sum

Le premier colloque Jeunes Chercheurs en Sciences Cognitives avait été organisé à l'initiative de l'ARC, par InCognito et l'ARC. Il s'était tenu les 23, 24 et 25 Mars 1994 à la Motte d'Aveillans (Isère), et avait été, pour la première fois, l'occasion de présenter une vue d'ensemble des préoccupations des jeunes chercheurs en Sciences Cognitives, Le vif succès obtenu par cette première manifestation a conduit à réaliser le deuxième Colloque Jeunes Chercheurs en

Sciences Cognitives, organisé cette fois-ci, toujours à l'initiative de l'ARC, par Cognito Ergo Sum, InCognito et l'ARC. Ce colloque s'est déroulé les 5, 6 et 7 juin 1996 à la Presqu'île de Giens (Var).

Nous avons accueilli à cette occasion 120 personnes, qui ont été hébergées sur le site en pension complète durant trois journées. Ces personnes ont pu assister à la présentation de 24 communications orales, ainsi qu'aux sessions de présentation de 30 communications affichées. Ces exposés avaient été retenus parmi 133 articles proposés suite à l'appel à communication, et ils couvraient principalement les domaines de l'Intelligence Artificielle, de la Psychologie, des Neurosciences et de l'épistémologie. De plus, trois conférenciers invités étaient présents à cette manifestation : les professeurs André Holley (Université de Lyon I), Gérard Ligozat (Université de Paris XI), et Patrick Mendelsohn (Université de Genève). Enfin, dans le but d'augmenter les échanges entre les participants, trois ateliers se sont déroulés sur les thèmes suivants:

- L'Entreprise et les Sciences Cognitives : animé par Nathalie Bonnardel (CREPCO, Aix-en-Provence) et Damien Raczy (LABACOLIL, Villeneuve d'Ascq)
- Cognition Naturelle et Cognition Artificielle : animé par John Stewart (Institut Pasteur, Paris) et Delphine Depy (CNRS-CNRC, Marseille)
- Multidisciplinarité des Méthodes en Sciences Cognitives : animé par Thierry Hasbrouca (CNRS-CNRC. Marseille) et Marina Petitiean (CREPCO, Aix-en-Provence)

Cette manifestation a connu une nouvelle fois un vif succès : elle aura été pour tous les participants l'occasion de cerner les préoccupations actuelles des Jeunes Chercheurs en sciences Cognitives, et de réfléchir sur la notion d'interdisciplinarité. Elle a aussi été l'occasion de créer et/ou de confirmer des liens entre les jeunes chercheurs provenant de diverses universités francaises et de certaines universités étrangères (Québec, Norvège, Suisse...).

Cognito Ergo Sum

### **CALENDRIER**



2es rencontres Terminologie et Intelligence Artificielle 3 & 4 avril 1997, Université Toulouse-le Mirail, France

### **Appel à communications**

Le groupe Terminologie et Intelligence Artificielle (TIA) organise une rencontre de deux jours (3 & 4 avril 1997) sur le thème "Terminologie et Intelligence Artificielle".

Les problématiques communes à la terminologie et à l'IA concernent au moins quatre disciplines : (i) la terminologie, (ii) la linguistique sur corpus, (iii) le traitement automatique des langues, (iv) l'acquisition et la représentation des connaissances en IA. Chacune de ces disciplines peut trouver un apport dans la confrontation et la réflexion commune avec les autres disciplines. Les laboratoires de recherche et les entreprises sont intéressés par des méthodes de constitution de données terminologiques et par les diverses applications qui les mettent en œuvre (Traduction, Aide à la rédaction, Indexation, Capitalisation des connaissances, Mémoires d'entreprises...).

La participation importante aux premières rencontres TIA, en 1995 (plus de 200 personnes), a montré qu'il existe une demande forte dans le sens de cette interdisciplinarité. Afin de mieux répondre à cette attente, TIA-97 sera organisé en deux temps et comprendra : des tutoriels visant à donner les éléments indispensables pour que chaque participant puisse comprendre les objectifs et les méthodes des autres disciplines ; des exposés sélectionnés par le Comité de Programme.

Le présent appel à communication concerne ces exposés. Les articles retenus devront rendre compte de travaux interdisciplinaires ou de l'intérêt des travaux présentés pour une ou plusieurs autres disciplines.

Thèmes concernés par les rencontres : fondements épistémologiques, terminologie et ontologies, analyse de corpus pour la constitution de ressources terminologiques (approches linguistiques et approches semiautomatiques), modèles de données terminologiques, terminologie et modélisation des connaissances, terminologie et acquisition des connaissances, bases de connaissances terminologiques, utilisation des ressources terminologiques (gestion de la documentation, accès à la documentation, mémoire technique).

### Comité de programme :

Président : Didier Bourigault (CAMS et EDF/DER, Clamart)

Bruno Bachimont (DIAM/SIM AP-HP, Paris), Brigitte Biébow (LIPN, Paris XIII), Jacques Bouaud (DIAM/SIM AP-HP, Paris), Jean Charlet (DIAM/SIM AP-HP, Paris), Anne Condamines (ERSS, Toulouse), Chantal Enguehard (IRIN, Nantes), Christian Jacquemin (IRIN, Nantes), Benoît Habert (ELI, ENS Fontenay St Cloud), Daniel Kayser (LIPN, Paris XIII), Pierre Lerat (LLI, Paris XIII), Ingrid Meyer (U. d'Ottawa), Gabriel Otman (LLI, Paris XIII), François Rastier (INALF, Paris), François Rousselot (ERIC, Strasbourg), Jean Royauté (INIST, Nancy), Patrick Saint Dizier (IRIT, Toulouse), Monique Slodzian (CRIL-INALCO, Paris), Sylvie Szulman (LIPN, Paris XIII), Yannick Toussaint (INRIA, CRIN-CNRS, Nancy), Pierre Zweigenbaum (DIAM/SIM AP-HP, Paris)

### Calendrier:

Réception des articles: 11 octobre 1996 • Réponses aux auteurs: 11 décembre 1996 • Réception de la version finale: 31 janvier 1997.

#### Modalités de soumission :

Les articles devront comporter au maximum 5000 mots et 14 pages (figures, bibliographie et annexes comprises) et être envoyés en 5 exemplaires avant le 11 octobre 1996 à :

Anne Condamines - ERSS, Maison de la Recherche 5, Allées Antonio Machado - 31058 Toulouse cedex, France



### Comité d'organisation :

Anne Condamines, présidente

Laurence Lamy, Marie-Paule Pery-Woodley et les membres de l'opération Terminologie et Langues Spécialisées de l'ERSS (UMR 5610 du CNRS) • Courrier-é: erss@cict.fr, téléphone: (33) 61 50 36 08, télécopie: (33) 61 50 46 77 - Page d'accueil www TIA: http://www.ens-fcl.fr/labos/eli/TIA/

### **CALENDRIER**

### CONGRÈS / COLLOQUES / ÉCOLES D'ÉTÉ...

#### **Août 1996**

### 28-31 Août 1996, RJC-IA'96, Troisièmes Rencontres des Jeunes Chercheurs en Intelligence Artificielle, Nantes.

Objectifs: Cette conférence d'ampleur nationale est l'occasion de donner aux auditeurs universitaires ou non universitaires, une vision large des recherches en cours. Le but principal est néanmoins de permettre à des étudiants préparant une thèse en Intelligence Artificielle, ou l'ayant soutenue depuis peu, de se rencontrer et de présenter leurs travaux. Voir appel dans la première partie de ce bulletin.

### Septembre 1996

### September 4-6, 1996, UCIS'96, Using Complex Information Systems, Poitiers, France.

Themes: Texts, documents, electronic databases and other interactive environments play an increasingly important role at home, at school and at the workplace. The cognitive processes by which people use such complex information systems are subject to an increased attention on the part of psychologists, educational scientists and human factors specialists. UCIS'96 will bring together researchers from these different disciplines, and to promote communication and cross-fertilization of theoretical and experimental advances made in each discipline.

Further Information: can be obtained through UCIS'96 Home Page, http://toty.joensuu.fi/ucis96/

### 25-27 septembre 1996, RECITAL'96, Rencontre des Etudiants-Chercheurs en Informatique pour le Traitement Automatique de la Langue, Dourdan.

Objectifs: Cette première rencontre des jeunes chercheurs francophones en TAL est une occasion donnée aux doctorants et jeunes docteurs (ayant soutenu après juin 1994) de se retrouver pour parler de leurs travaux et de l'aboutissement de leurs projets de recherche. Organisé en septembre 1996, RECITAL se déroulera dans le cadre agréable du VVF de Dourdan (Essonne), près de Paris. Le nombre des contributions sera restreint à une quarantaine pour assurer la convivialité du rassemblement en facilitant les échanges formels et informels.

Contact: http://www.limsi.fr/RECITAL96 ftp.limsi.fr/pub/recital

#### Mars 1997

#### 23-26 mars 1997, AIME'97, Artificial Intelligence in Medicine Europe, World Trade Center, Grenoble, France.

Objectifs: La conférence "Artificial Intelligence in Medicine Europe" est une émanation de la Société Européenne d'Intelligence Artificielle en Médecine. Cette société a été créée en 1986, à la suite d'une réunion de travail qui s'était tenue à Pavie un an auparavant. Ses objectifs sont les suivants : (i) soutenir la recherche fondamentale et appliquée sur l'utilisa-

tion des techniques de l'Intelligence Artificielle dans le domaine médical; (ii) favoriser la dissémination des résultats les plus significatifs par des conférences bi-annuelles et (iii) favoriser le transfert industriel des produits médicaux les plus avancés. La conférence comprend des conférences invitées, ainsi que des communications et des posters sélectionnés, des démonstrations, des table-rondes, enfin des tutoriaux permettant la formation aux technologies les plus avancées. Date limite de réception des propositions : 15 Septembre 96. Contributions n'excédant pas 5000 mots. Elles doivent décrire des travaux originaux non publiés par ailleurs.

Notification aux auteurs: 15 novembre 1996 Réception des versions définitives : 15 décembre 1996

Programme Committee Chair: Elpida Keravnou, Department of Computer Science, University of Cyprus, 75 Kallipoleos Strees, P.O. Box 537, CY-1678 Nicosia - Cyprus, Tél: + 357 2 338705, Fax: +357 2 339062, Email: elpida@turing.cs.ucy.ac.cy Organising Committee Chair: Catherine Garbay, Lab.

TIMC - IMAG, Institut Bonniot - Domaine de la Merci, 38706 La Tronche - France, Tél: + 33 76 54 94 85, Fax: + 33 76 54 95 49, Email: Catherine.Garbay@imag.fr

Tutorials Chair :Jeremy Wyatt, Biomedical Informatics Unit, Imperial Cancer Research Fund, Lincoln's Inn Fields, PO Box 123, London WC2A 3PX, United Kingdom, Tél: + 44-71-269 36 37, Fax : + 44-71-269 31 86, Email : j.wyatt@icrf.icnet.uk

Secrétariat: SGV DESTINATION, 73 bis Grande Rue, F-38700 LA TRONCHE, E-mail: dcongres@imaginet.fr, Tél.: (33) 76 01 00 54, Fax: (33) 76 63 16 36, page d'accueil www:http://www-timc.imag.fr/aime97/

#### **Avril 1997**

### 3-4 avril 1997, TIA'97, 2es Rencontres Terminologie et Intelligence Artificielle, Université Toulouse-le Mirail,

Objectifs: Le groupe Terminologie et Intelligence Artificielle (TIA) organise une rencontre de deux jours sur le thème « Terminologie et Intelligence Artificielle ». Les problématiques communes à la terminologie et à l'IA concernent au moins quatre disciplines : la terminologie, la linguistique sur corpus, le traitement automatique des langues et l'acquisition et la représentation des connaissances en IA. Chacune de ces disciplines peut trouver un apport dans la confrontation et la réflexion commune avec les autres disciplines. Les laboratoires de recherche et les entreprises sont intéressés par des méthodes de constitution de données terminologiques et par les diverses applications qui les mettent en œuvre (Traduction, Aide à la rédaction, Indexation, Capitalisation des connaissances, Mémoires d'entreprises...). Voir appel dans ce bulletin.

Nous annonçons dans cette rubrique les manifestations IA (i) se déroulant en France ou (ii) francophones et se déroulant hors de France.

Pour inclure une annonce, envoyer un court texte (disquette Mac ou mail encodé avec accents) à Jean Charlet (voir « l'ours » page 2 de ce bulletin).

### SOMMAIRE DES REVUES

### Bon anniversaire

L'OBJET : logiciel, base de données, réseaux est une revue française et trimestrielle qui existe depuis un an. Pour contacter la rédaction Jean-Claude Royer royer@irin.univ-nantes.fr

Informations complémentaires sur Internet http://www.tools.com/lobjet

ou http://www.sciences.univ-nantes.fr/info/perso/permanents/royer/lobjet

### REVUE D'INTELLIGENCE ARTIFICIELLE

### REVUE D'INTELLIGENCE ARTIFICIELLE VOL 10 N°2-3, 1996

NUMERO SPÉCIAL APPROCHES SYMBOLIQUES ET NUMERIQUES DE L'APPRENTISSAGE

- Editorial, O. Gascuel
- Règles de décision de Bayes et méthodes statistiques de discrimination. Bayes decision rules and discrimination analysis: statistical approaches, G. Caraux, Y. Lechevallier
- Statistique, apprentissage et généralisation. Application aux réseaux de neurones. Statistics, learning and generalization. Illustration with neural nets, P. Gallinari, O. Gascuel
- Algorithmes génétiques et apprentissage. Genetic algorithms and machine learning, G. Venturini
- Contrôle d'un algorithme génétique. Controlling genetic algorithms, M. Sebag, M. Schoenauer

### ARTIFICIAL INTELLIGENCE

# ARTIFICIAL INTELLIGENCE V80, N°2, February

- A metatheory of a mechanized object theory, F. Giunchiglia and P. Traverso
- A logic of time, chance, and action for representing plans, P. Haddaway
- Model simplification by asymptotic order of magnitude reasoning, K.M.-K Yip
- On embedding default logic into Moore's autoepistemic logic(research Note), G. Schwarz

### ARTIFICIAL INTELLIGENCE V81, N°1-2, march

SPECIAL VOLUME ON FRONTIERS IN PROBLEM SOLVING: PHASE TRANSITIONS AND COM-PLEXITY

- Phase transitions and the search problem (Editorial), T. Hogg, B.A. Huberman, and C.P. Williams
- Generating hard satisfiability problems, B. Selman, D.G. Mitchell and H.J. Levesque
- Experimental results on the crossover point in random 3-SAT, J.M. Crawford and L.D. Auton

- The satisfiability constraint gap, I.P. Gent and T. Walsh
- An empirical study of phase transitions in binary constraint satisfaction problems, P. Prosser
- Some pitfalls for experimenters with random SAT, D.G. Mitchell and H. J. Levesque
- Refining the phase transition in combinatorial search, T.
- Locating the phase transition in binary constraint satisfaction problems, B.M. Smith and M.E. Dyer
- Hard random 3-SAT problems and the Davis-Putnam procedure, J.W. Freeman
- Implicates and prime implicates in Random 3-SAT, R. Schrag and J.M. Crawford
- A study of complexity transitions on the asymmetric traveling salesman problem, W. Zhang and R.E. Koff
- A probabilistic analysis of propositional STRIPS planning, T. Bylander
- Critical behavior in the computational cost of satisfiability testing, B. Selman and S. Kirkpatrick
- Epsilon-transformation: exploiting phase transitions to solve combinatorial optimization problems, J.C. Pemberton and W. Zhang
- Problem structure heuristics and scaling behavior for genetic algorithms, S.H. Clearwater and T. Hogg

### ARTIFICIAL INTELLIGENCE V82, N°1-2, april 1996

- Structure-driven algorithms for truth maintenance, R. Dechter and A. Dechter
- Probably approximately optimal satisficing strategies, R. Greiner and P. Orponen
- Knowledge representation and inference in similarity networks and Bayesian multinets, D. Geiger and D. Hecker-
- ALX, an action logic for agents with bounded rationality, Z. Huang, M. Masuch and L. Polos
- On the nature and role of the modal truth criteria in planning, S. Kambhampati and D.S. Nau
- Noise modelling and evaluating learning from examples,
- Optimal composition of real-time systems, S. Zilberstein and S. Russel
- Knowledge-level analysis of belief base operations, S. O. Hansson
- Searching game trees under a partial order, P. Dasgupta, P.P. Chakrabarti and S.C. DeSarkar
- Support set selection for abductive and default reasoning, B. Selman and H.J. Levesque
- On the hardness of approximate reasoning, D. Roth
- Stochastic modelling of Genetic Algorithms, D. Reynolds and J. Gomatam
- A new theorem in particle physics enabled by machine discovery (Research note), R.E. Valdés-Pérez
- A test battery for rational database updating (Research note), S.O. Hansson
- On the relative expressiveness of description logics and predicate logics (Research note), A. Borgida
- Book review of Artificial Intelligence: A Modern

### SOMMAIRE DES REVUES

Approach (S. Russell and P. Norvig), N.J. Nilsson

• Book review of The Computational Brain (P. S. Churchland and T. J. Sejnowski), G.N. Reeke Jr

#### ARTIFICIAL INTELLIGENCE V83, N°1, May 1996

- Measures of uncertainty in expert systems, P. Walley
- The token reification approach to temporal reasoning, L. Vila and H. Reichgelt
- Qualitative system identification: deriving structure from behavior, A.C. Cem Say and S. Kuru
- Abduction to plausible causes: an event-based model of belief update, C. Boutilier
- · Optimization of Pearl's method of conditioning and greedy-like approximation algorithms for the vertex feedback set problem, A. Becker and D. Geiger

### COGNITIVE SCIENCE

### COGNITIVE SCIENCE V20, N°1, Jan.-March 1996

- Generating Facial Expressions for Speech, C. Pelachaud, N. I. Badler, and M. Steedman
- Subsymbolic Case-Role Analysis of Sentences With Embedded Clauses, R. Miikkulainen
- The Impact of Goal Specificity on Strategy Use and the Acquisition of Problem Structure, R. Vollmeyer, B. D. Burns, and K. J. Holyoak
- On the Validity of Simulating Stagewise Development by Means of PDP Networks: Application of Catastrophe Analysis and an Experimental Test of Rule-Like Network Performance, M.E.J. Raijmakers, S. van Koten, and P.C.M. Molenaar

### AI MAGAZINE

### AI MAGAZINE V17, N°1, SPRING 1996

- Woody Bledsoè: His Life and Legacy, M. Ballantyne, R.S. Boyer, and L. Hines.
- CHINOOK: The World Man-Machine Checkers Champion, J. Schaeffer, R. Lake, P. Lu, and M. Bryant
- The 1995 Robot Competition and Exhibition, D. Hinkle, D. Kortenkamp, and D. Miller
- CAIR-2: Intelligent Mobile Robot for Guidance and Delivery, H.S. Yang, J. Chung, B.S. Ryu, and J. Lee
- LOLA: Probabilistic Navigation for Topological Maps, R. Gutierrez-Osuna and R. C. Luo
- LOLA: Object Manipulation in an Unstructured Environment, R. LeGrand and R. C. Luo
- Programming CHIP for the IJCAI-95 Robot Competition, J. Firby, P. Prokopowicz, M. Swain, R. Kahn, and D. Franklin
- REPORTS

The 1995 Fall Symposia Series

IJCAI-95 Workshop on Adaptation and Learning in Multiagent Systems

Thirteenth International Distributed AI Workshop Case-Based Reasoning

IEEE Fourth International Workshop on Enabling Techno-

The Fourth International Conference on User Modeling

### INTERNATIONAL JOURNAL OF **HUMAN COMPUTER STUDIES**

#### INTERNATIONAL JOURNAL OF HUMAN-COMPU-**TER STUDIES (1996) 44(2)**

Special Issue: Verification and Validation

- Editors: Robert Plant and Alun D. Preece • On the validation and verification of production systems:
- a graph reduction approach, Murrell, S. and Plant, R.
- Validating dynamic properties of rule-based systems, Preece, A. D., Grossner, C. and Radhakrishnan, T.
- The relationship between errors and size in knowledgebased systems, O'Leary, D. E.
- Structure-preserving specification languages for knowledge-based systems, Van Harmelen, F. and Aben, M.
- A methodology to incorporate formal methods in hybrid
- KBS verification, Gamble, R. F. and Baughman, D. M. • Refinement complements verification and validation,
- Verification and validation with ripple-down rules, Ho Kang, B., Gambetta, W. and Compton, P.

### INTERNATIONAL JOURNAL OF HUMAN-COMPU-**TER STUDIES (1996) 44(3-4)**

Special Issue: The Sisyphus-VT initiative

Craw, S.

Editors: A. Th. Schrieber and W. P. Birmingham

- Implementing the Sisyphus-93 task using Soar/TAQL, Yost, G. R.
- Reusable ontologies, knowledge-acquisition tools, and performance systems: PROTÉGÉ-II solutions to Sisyphus-2, Rothenfluh, T. E., Gennari, J. H., Eriksson, H., Puerta, A. R., Tu, S. W., and Musen, M. A.
- Solving VT in VITAL: a study in model construction and knowledge reuse, Motta, E., Stutt, A., Zdrahal, Z., O'Hara, K. and Shadbolt, N.
- Sisyphus-VT: a CommonKADS solution, Schreiber, A. Th. and Terpstra, P.
- Solving VT by reuse, Runkel, J. T., Birmingham, W. P. and Balkany, A.
- Combining KARL and CRLM for designing vertical transportation systems, Poek, K., Fensel, D., Landes, D. and Angele, J.
- Modelling an elevator design task in DESIRE: the VT example, Brazier, F.M.T., van Langen, P.H.G., Treur, J., Wijngaards, N.J.E. and Willems, M.
- Configuring elevator systems, Yost, G. R. and Rothenfluh,
- The configuration design ontologies and the VT elevator domain theory, Gruber, T.R., Olsen, G.R. and Runkel, J.

### **SOMMAIRE DES REVUES**

#### INTERNATIONAL JOURNAL OF HUMAN-COMPU-**TER STUDIES (1996) 44(5)**

- Systematic building of conceptual classification systems with C-KAT, Zacklad M. and Fontaine D.
- Causal model-based knowledge acquisition tools: discussion of experiments, Charlet J., Reynaud C. and Krivine
- Inductive effective operator control through ecological interface design, Pawlak W.-S. and Vicente K.-J.
- Roles of design knowledge-based systems, Benaroch

### THE JOURNAL OF LOGIC PROGRAMMING

### THE JOURNAL OF LOGIC PROGRAMMING V26, N°2, 1996 - SPECIAL ISSUE: APPLICATIONS OF LOGIC PROGRAMMING

Guest Editor: L. Sterling

- Computer Support for Protocol-based Treatment of Cancer, P. Hammond and M. Sergot
- The NRL Protocol Analyzer: an Overview, C. Meadows
- Prolog for Structured Character Description and Font Design, M.J. Dürst
- A Tool for Developing Interactive Configuration Applications, T. Axling and S. Haridi
- Intelligent Data Retrieval in Prolog: an Illuminating Idea, C. Ruggieri, M. Sancassani, N. Dore, F. Russo, and U. Manfredi
- Using Logic Programming and Coroutining for Electronic CAD, U. Bieker and A. Neuman
- Constraint Logic Programming for Examination Timetabling, P. Boizumault, Y. Delon, and L. Peridy

### THE JOURNAL OF LOGIC PROGRAMMING V27, N°1, April 1996

- In Memoriam: Paris C. Kanellakis, P. Van Hentenryck
- Stable Models and their Computation for Logic Programming with Inheritance and true Negation, F. Buccafurri, N. Leone, and P. Rullo
- Cuts and Side-Effects in and-or Parallel Prolog, G. Gupta and V. Santos Costa
- Matching of Bounded Set Terms in the Logic Language LDL (technical note), N. Arni, S. Greco, and D. Saccà

### MACHINE LEARNING

### MACHINE LEARNING V22, N°1-3, January / February / March 1996

Special Issue on Reinforcement

Guest Editor: L. Pack Kaelbling

- Introduction, L. Pack Kaelbling
- Efficient Reinforcement Learning through Symbiotic Evolution, D.E. Moriarty and R. Miikkulainen

- · Linear Least-Squares Algorithms for Temporal Difference Learning, S.J. Bradtke and A.G. Barto
- Feature-Based Methods for Large Scale Dynamic Programming, J.N. Tsitsiklis and B. Van Roy
- On the worst-Case Analysis of Temporal-Difference Learning Algorithms, R.E. Schapiro and M.K. Warmuth
- Reinforcement Learning with Replacing Eligibility Traces, S.P. Singh and R.S. Sutton
- Average Reward Reinforcement Learning: Foundations, Algorithms, and Empirical Results, S. Mahadevan
- The Loss from Imperfect Value Functions in Expectations-Based and Minimax-Based Tasks, M. Heger
- The Effect of Representation and Knowledge on Goal-Directed Exploration with Reinforcement Learning Algorithms, S. Koenig and R.G. Simmons
- Creating Advice-Taking Reinforcement Learners, R. Maclin and J.W. Shavlik
- Technical Note: Incremental Multi-Step Q-Learning, J. Peng and R.J. Williams

### JOURNAL OF ARTIFICIAL INTELLIGENCE IN EDUCATION

### JOURNAL OF ARTIFICIAL INTELLIGENCE IN EDUCATION V6, N°1, 1995

- ESSCOTS for Learning: Transforming Commercial Software into Powerful Educational Tools, D. McArthur, M. W. Lewis, and M. Bishay
- Situated Plan Attribution, R. W. Hill, Jr. and W. Lewis Johnson
- QMaPS: Qualitative Reasoning for Simulation Learning Environments, W.R. Van Joolingen
- Using a Distributed Training Architecture to Enhance User Modeling in a Communications Planning Simulation, Ph.D. Gillis

#### JOURNAL OF ARTIFICIAL INTELLIGENCE IN **EDUCATION V6, N°2-3, 1995**

- Modelling Teaching Strategies, N. Major
- Truth Maintenance Techniques for Modelling Student's Behaviour, P. Giangrandi and C. Tasso
- The Construction and Coordination of Complementary Problem Representations in Physics, R. Ploetzner
- Supporting the Use of External Representations in Problem Solving: The Need for Flexible Learning Environments, R. Cox and P. Brna

#### JOURNAL OF ARTIFICIAL INTELLIGENCE IN EDUCATION V6, N°4, 1995

- Monitoring Computer-Based Collaborative Problem Solving, M.M. McManus and R.M. Aiken
- Implementation of Motivational Tactics in Tutoring Systems, T. del Solato and B. du Boulay
- On the Effectiveness of a Neural Network for Adaptive

### SOMMAIRE DES REVUES

External Pacing, A. R. Montazemi and F. Wang

• STEPS: A Simulated, Tutorable Physics Student, S. Ur and K.VanLehn

### JOURNAL OF ARTIFICIAL INTELLIGENCE IN **EDUCATION V7, N°1, 1996**

- Individual Selection of Examples in an Intelligent Learning Environment, G. Weber
- Law Encoding Diagrams for Instructional Systems, P.C.-
- · Refinement-Based Student Modelling and Automated Bug Library Construction, P. Baffes and R. Mooney

### INTERNATIONAL JOURNAL OF APPROXIMATE REASONING

#### INTERNATIONAL JOURNAL OF APPROXIMATE **REASONING V14 N°2/3 1996**

- Modal logic interpretation of Dempster-Shafer theory: An infinite case, D. HARMANEC, G.J. KLIR, Z.Y. WANG
- A theory of Gaussian belief functions, L.P. LIU
- On the normalization of fuzzy belief structures, R.R.
- · Reasoning in evidential networks with conditional belief functions, H. XU, P. SMETS
- A case study in the qualitative verification and debugging of numerical uncertainty, S. PARSONS, A. SAF-FIOTTI
- Texts on evidence theory: Comparative review, A. RAMER

#### INTERNATIONAL JOURNAL OF APPROXIMATE **REASONING V14 N°4 1996**

- · Heuristics and normative models of judgment under uncertainty, P. WANG
- On the use of hierarchical clustering in fuzzy modeling, M. DELGADO, A.F. GÓMEZ-SKARMETA, A. VILA
- A generic ATMS, J.L. CASTRO, J.M. ZURITA
- Development and evaluation of five fuzzy multiattribute decision-making methods, E. TRIANTAPHYLLOU, C.T.

### INTERNATIONAL JOURNAL O INTELLIGENT SYSTEMS

### INTERNATIONAL JOURNAL OF INTELLIGENT **SYSTEMS V11 N°3 1996**

- Knowledge acquisition by random sets, X.T. PENG, P.Z. WANG, A. KANDEL
- On random sets, knowledge acquisition, and pattern recognition, X.T. PENG, P.Z. WANG, A. KANDEL

#### INTERNATIONAL JOURNAL OF INTELLIGENT **SYSTEMS V11 N°4 1996**

- A multivalued logic ATMS, J.L. CASTRO, J.M. ZURITA
- Bucket brigade algorithm for hierarchical censored production rule-based system, N.M. HEWAHI, K.K. BHA-RADWAJ
- Virtual cognition in language, J.A. MOYNE

### INTERNATIONAL JOURNAL OF **UNCERTAINTY, FUZZINESS AND** KNOWLEDGE BASED SYSTEMS

#### INTERNATIONAL JOURNAL OF UNCERTAINTY, FUZZINESS AND KNOWLEDGE-BASED SYSTEMS V4 N°1 1996

- Decision making with belief structures: An application in risk management, K.J. ENGEMANN, H.E. MILLER, R.R. YAGER
- Fuzzy Mealy machines: Homomorphisms, admissible relations and minimal machines, J.N. MORDESON, P.S.
- Qualitative modalities, W. VAN DER HOEK\*
- Modeling fuzzy reasoning using high level fuzzy Petri nets, H. SCARPELLI, F. GOMIDE, W. PEDRYCZ
- Urysohn-closedness on fuzzy lattices, S.L. CHEN
- Book review of "Non-Additive Measure and Integral, by D. DENNEBERG", D. DUBOIS

### JOURNAL OF APPLIED NON-CLASSICAL LOGICS

### JOURNAL OF APPLIED NON-CLASSICAL LOGIC V6 N°2 1996

- · A fugue on the themes of awareness logic and correspondence, E. THIJSSE, H. WANSING
- The universe of discourse of modal logic, A. PRELLER
- A three-valued approach to default logic, A. RADZI-
- On the representation of indeterminate identity via vague concepts, M.K. CHAKRABORTY, A. CHATTERJEE

Henri Prade et Agathe Lorente (IRIT, Equipe "Raisonnements Plausibles, Décision, et Méthodes de Preuves") contribuent régulièrement à cette rubrique, ainsi que Jean Charlet (INSERM, DIAM). Qu'ils en soient ici vivement remerciés.

### RÉSUMES DE THÈSES

### Base de connaissances, Hypermédia, et Contexte

#### Élisabeth Montabord

Thèse de l'Université de Rennes. soutenue le 3 Mai 1996 préparée dans les groupes LIAS à Télécom Brest, et SIM à Rennes

Le projet dans lequel s'insère cette étude a pour objet la conception d'un système interactif d'aide à la décision (SIAD) en situation complexe, fondé sur un système à base de connaissances, pour le traitement de pathologies du cerveau. Le travail de cette thèse s'est alors plus particulièrement porté sur une méthodologie de conception de bases de connaissances objets et d'un hypermédia, dont la principale caractéristique est la sensibilité au contexte.

Nous proposons une organisation de la base de connaissances "cerveau" en différents types de connaissances. Les connaissances sources caractérisent les éléments du cerveau. Elles sont organisées en plusieurs perspectives (une par discipline médicale impliquée). Les connaissances contextuelles comportent deux modèles. Un modèle hiérarchique de tâches permet de représenter les fonctions élémentaires du système, de planifier l'accès aux informations (connaissances et données), et de gérer la coopération entre le système et l'utilisateur. Un modèle hiérarchique de disciplines permet de sélectionner les sous-ensembles de concepts pertinents (éléments, attributs, et tâches) et de restreindre l'ensemble des informations accessibles.

L'hypermédia constitue une interface unifiée du SIAD, support du dialogue entre le système et l'utilisateur. La navigation repose sur une entité appelée point de vue, décrivant le contexte de communication système/utilisateur, et les informations recherchées dans le contexte. Le contexte est constitué du domaine de l'utilisateur, de la tâche courante, et de connaissances relatives à l'exécution des sous-tâches. Le point de vue permet une adaptation continuelle et dynamique aux demandes de l'utilisateur.

Les caractéristiques de l'hypermédia sont les suivantes : contextualité ; dynamisme ; interactivité : dans la planification de l'accès aux informations et dans le processus de résolution des tâches ; déclarativité. Le système dans son ensemble est modulaire : les différents types de connaissances et de données sont partitionnées ; évolutif : d'une part les différents composants du système (base de connaissances, base de données, hypermédia) peuvent être modifiés sans que soit remise en cause l'organisation générale du système, d'autre part l'ensemble des connaissances, des

données, des types d'information, des disciplines, et des tâches représentées, sont facilement extensibles ; générique : par la description abstraite et réutilisable de tous les types de connaissances représentés.

#### Mots-clés:

Représentation à objets, Base de connaissances, Tâches, Hypermédia, Navigation, Contexte, Point de

#### Adresse:

Montabord Elisabeth 12 bis rue saint Yves, 29290 Saint Renan e-mail: montabor@worldnet.fr tel: 98324635

Capitalisation d'un savoir faire industriel pour la réalisation d'un système d'aide à la décision à base de multi-agents. Application dans une chaîne agro-alimentaire d'une malterie : Projet C.O.R.M.O.R.AN.

Françoise. Entzmann-Maurel Thèse de l'Université de Nancy-1 soutenue le 12 Avril 1996 préparée à l'ESSTIN - Doctorat en Automatique

Le procédé de maltage est un procédé difficile à maîtriser en raison d'une connaissance partielle de l'incidence entre les différentes sources d'informations intervenant dans le procédé. Cette recherche de qualité, de maîtrise de procédé au travers d'une formalisation de savoir faire constitue l'objet de la thèse. La première phase a consisté au choix d'un outil de capitalisation de savoir faire parmi les techniques de modélisation de la connaissance.

Sa mise en place a permis de formaliser l'expertise sous forme de conseil d'actions. Pour cela, différentes techniques de modélisation ont été employées et sont présentées dans ce document. Le résultat de l'exploitation de l'expertise en malterie aboutit à un conseil d'actions qui est ensuite validé grâce aux données industrielles de production. Cette phase de validation s'appuie sur les capacités d'apprentissage du neuroagent. De plus, l'analyse de cette validation enrichit l'expertise. Ces trois modes constituent les éléments de base du système global d'aide à la décision : "C.O.R.M.O.R.A.N." Actuellement, cet outil est utilisé sur un des sites de production.

#### Mots-clés:

Expertise - Concept Neuroagent - Procédé de maltage -Apprentissage - Système d'aide à la décision.

Adresse: francoise@mail.azur.fr

### RÉSUMES DE THÈSES

### Sémantique formelle de l'expression du mouvement. De la sémantique lexicale au calcul de la structure du discours en français.

### Pierre Sablayrolles

Thèse de l'Université Paul Sabatier, Toulouse soutenue le 21 Décembre 1995 à l'IRIT, UPS, Toulouse

Le travail que nous présentons dans cette thèse s'inscrit dans le cadre de l'analyse formelle de la sémantique du langage naturel. Plus précisément, il concerne la représentation et le traitement des structures spatio-temporelles de la langue. Développer des formalismes pour représenter la connaissance spatiale et temporelle a toujours été d'une importance considérable en intelligence artificielle et en sciences cognitives. Pour mettre au jour et modéliser les concepts spatio-temporels de base sous-jacents à cette connaissance, le langage naturel constitue une excellente trace de ces concepts, en particulier par la façon même dont on utilise certaines expressions plutôt que certaines autres pour décrire le monde. Or, dans la langue, ce sont les expressions décrivant des mouvements qui expriment le mieux les liens étroits qui unissent l'espace et le temps, puisqu'elles représentent une succession de localisations spatiales évoluant au cours du temps. Notre travail s'appuie par conséquent sur une étude linguistique descriptive de l'ensemble des verbes de mouvement et des prépositions spatiales du français.

D'abord, nous clarifions la notion de "lieu", très ambiguë dans la littérature. Cela nous permet de distinguer quatre grandes classes de verbes de mouvement. Pour l'une de ces classes (les verbes de changement de lieu), nous réalisons une analyse linguistique systématique de l'ensemble des éléments pertinents, en adoptant quatre points de vues différents mais complémentaires : une approche "localiste", une approche "componentielle", une approche "connaissance du monde" et une approche "aspectualité".

En nous appuyant sur une étude localiste des prépositions spatiales du français, nous énonçons, sous la forme d'axiomes logiques, des règles formelles de composition nous permettant de calculer la sémantique spatio-temporelle d'un complexe de mouvement à partir des propriétés spatio-temporelles du verbe et de la préposition qui le composent. Ces règles sont formalisées dans une théorie basée sur la Méréologie. Le calcul fait intervenir des informations de nature lexicales. mais aussi des informations liées à la structure même du discours. Nous utilisons la SDRT — Segmented Discourse Representation Theory — de N. Asher pour modéliser ces interactions et nous faisons appel à l'opérateur d'implication faible des logiques non-monotones pour coder les défauts dans nos entrées lexicales. Nous montrons également comment les informations discursives peuvent être utilisées pour la désambiguïsation lexicale et comment les informations lexicales peuvent aider au calcul de la structure du discours.

#### Mots-clés:

expression du mouvement, sémantique lexicale, analyses linguistiques descriptives, sémantique spatiotemporelle formelle, verbes de mouvement, calcul compositionnel, prépositions spatiales, structure du discours.

#### Adresse:

IRIT, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex, France. e-mail: sablay@irit.fr,

Constribution à la méthode d'exploration contextuelle Applications au résumé automatique et aux représentations temporelles Implémentation informatique du système SERAPHIN

#### Jawad Berri

Thèse de l'Université de Paris-Sorbonne (Paris-4), soutenue le 7 mai 1996 préparée dans l'équipe LALIC au CAMS

Nous présentons dans cette recherche, la méthode d'exploration contextuelle (EC) et son application au résumé automatique de textes et aux représentations temporelles. La méthode d'EC consiste à focaliser l'attention sur des locutions "pivots" présentes dans un texte, par exemple des mots grammaticaux signifiants. Ces locutions correspondent à un savoir linguistique qui est indépendant d'un domaine particulier. On se situe donc, d'emblée, au niveau des connaissances linguistiques pour tout traitement sémantique d'un texte.

Le système SERAPHIN que nous avons réalisé pour répondre au problème du résumé automatique, permet d'extraire les phrases les plus pertinentes d'un texte en vue de constituer un résumé susceptible de donner un aperçu sur le contenu du texte. L'extraction des phrases s'appuie essentiellement sur les indicateurs (qui attirent l'attention d'un lecteur) et les indices linguistiques coprésents dans un contexte donné. SERAPHIN est implémenté sous forme d'un système à base de connaissances. Il s'appuie, dans son analyse, sur un modèle objet générique qui représente la structure logique du texte traité. Le résumé constitué, sert de support pour aiguiller des utilisateurs vers la lecture ou non du texte en entier.

L'application de l'EC aux représentations temporelles a permis la réalisation du système SECAT. Le but du système est de déterminer la valeur sémantique aspectuelle des morphèmes grammaticaux d'un texte. Dans la continuité de ces travaux, nous avons proposé quelques solu-

### RÉSUMES DE THÈSES

tions en vue de construire une représentation temporelle d'un texte. On a notamment, focalisé l'attention sur la notion d'actant et les interactions actancielles qui régissent l'organisation des situations temporelles dans un univers donné. Cette conception devient compatible avec une architecture Multi-agent car le raisonnement est focalisé sur l'actant, sur ses capacités et limitations, sur ses interactions avec d'autres actants, etc.

#### Mots-clés:

Exploration contextuelle, connaissances linguistiques, résumé automatique, représentation temporelle, modèle objet, système à base de connaissances.

#### Adresse:

ISHA - CAMS, 96 bvd Raspail, 75006 Paris

e-mail: berri@cams.msh-paris.fr

### **Du Raisonnement Social** Chez les Agents Une Approche Fondée sur la Théorie de la Dépendance

#### Jaime Simao Sichman

Thèse de l'Institut National Polytechnique de Grenoble, soutenue le 5 Septembre 1995

Cette thèse présente le modèle d'un mécanisme de raisonnement social fondé sur la théorie de la dépendance. Ce modèle permet à un agent de raisonner sur autrui et plus particulièrement de calculer ses relations et situations de dépendance. Un agent est dépendant d'un autre si celui-ci peut l'aider/l'empêcher d'atteindre un de ses buts. Nous considérons notre mécanisme de raisonnement social comme un composant essentiel pour la conception d'agents artificiels réellement autonomes, évoluant dans un univers multi-agents ouvert. La notion d'ouverture désigne la capacité d'ajouter ou de retirer des agents dynamiquement dans le système. Comme dans ces systèmes l'organisation des agents ne peut pas être spécifiée pendant la phase de conception, la résolution coopérative de problèmes est fondée sur la formation dynamique de coalitions. Dans ce contexte, des agents doivent être capables de s'adapter aux changements dynamiques du système, en particulier en évaluant pendant la phase de résolution si leurs buts sont réalisables et si leurs plans sont exécutables. Comme nous ne supposons pas que les agents soient bienveillants, notre modèle fournit un critère pour évaluer les partenaires le plus susceptibles d'accepter une proposition de coalition. Enfin, comme dans ces systèmes des agents n'ont pas généralement une représentation complète et correcte les uns des autres, notre modèle leur permet de détecter une inconsistance au niveau de la société et de choisir un contexte à maintenir. Nous avons implémenté ce mécanisme de raisonnement social en utilisant une programmation orientée objet. Nous

l'avons utilisée pour développer deux applications, le simulateur DEPNET et le système DEPINT, qui illustrent son utilisation selon deux perspectives scientifiques différentes. D'une part, selon une perspective de simulation sociale, notre modèle fournit un outil informatique permettant l'analyse et la prédiction des divers schémas intéressants d'interaction sociale, et l'évaluation du pouvoir social des agents. D'autre part, selon une perspective de résolution de problèmes, notre modèle peut être utilisé pour concevoir dynamiquement l'organisation des agents dans un contexte de systèmes multiagents ouverts.

raisonnement social, raisonnement sur autrui, intelligence artificielle distribuée, systèmes multi-agents, intégration de systèmes, systèmes ouverts.

### Apport à la configuration de connaissances par réingéniérie de systèmes existants. KA7: un système appliqué au domaine des réseaux électriques

### Lionel Perrot

Thèse de l'Université Pierre et Marie Curie (Paris-6) soutenue le 15 décembre 1995

Les travaux de cette thèse sont motivés par le besoin d'adapter des systèmes artificiels à des environnements en perpétuelle évolution. Nous souhaitons rendre cette adaptation la plus automatique possible. L'atelier KA7 (Knowledge Acquisition for SEPT) que nous avons réalisé assiste la gestion du cycle de vie des composants logiciels à intégrer dans ces systèmes, appelés systèmes cibles.

La première partie est consacrée à la présentation de plusieurs approches que nous considérons comme majeures. Les approches que nous avons retenues sont celles de W.J.Clancey, B.Chandrasekaran, J.McDermott, KADS et les travaux connexes ou dérivés, approches basées sur les tâches et méthodes. Nous attachons une grande importance aux formalismes mis en œuvre par chacune d'entre elles.

La seconde partie consiste en la construction incrémentale d'une proposition. Comme en KADS-II, les systèmes cibles que nous construisons contiennent trois niveaux: Tâche, Înférence et Domaine. Cependant, nous utilisons des langages spécialisés pour chacun de ces niveaux. Ces langages sont soit repris (LISA-Runtime et CRL), soit concus et réalisés par nous même (MBMF Modèle de Bon et Mauvais Fonctionnement, TTD Traitement du Temps et de la Durée, SLD Synthèse et Localisation de Défaut). La troisième et dernière partie explique le cycle de vie des composants qui sont produits par l'atelier KA7. Cette partie est essentiellement organisée autour d'images d'écrans et d'exemples de code source commentés. Les étapes d'acquisition, configuration, opéra-

### RÉSUMES DE THÈSES

tionnalisation et maintenance des composants logiciels sont décrites et illustrées.

#### Perspectives.

Deux axes de recherche se dégagent naturellement de cette thèse: le premier concerne la validation statique des composants produits par l'atelier KA7. Le deuxième se rapporte à l'agencement automatique de ces composants à l'intérieur des systèmes cibles.

#### Mots-clés

KA7, Configuration de connaissances, Réingénierie d'artefacts, Cycle de vie des composants logiciels, Surveillance des réseaux électriques, Ingénierie des systèmes à base de connaissances.

### Adresse

E-mail: LionelPerrot@magic.fr

### Un environnement opérationnel de conception et de réalisation de systèmes multi-agents

#### Zahia Guessoum

Thèse de l'Université Pierre et Marie Curie (Paris-6) soutenue le 22 mai 1996

La représentation de systèmes complexes, en vue de leur contrôle ou de leur simulation, nécessite de définir des agents hybrides dotés de capacités substantielles de raisonnement et de représentation du monde dans lequel ils évoluent et de capacités réactives, pour respecter des contraintes temporelles temps-réel.

Nous proposons un modèle d'agents opérationnel qui possède ces caractéristiques. Ces agents combinent des propriétés cognitives et réactives logées dans des modules asynchrones différents, ce qui repose le problème de la coordination à l'intérieur même de l'agent sans pour autant faire disparaître celui de la coordination entre agents. Pour apporter une solution à ce problème, nous proposons un modèle d'agents dont les capacités

cognitives sont fondées sur l'emploi de bases de règles et de métarègles. Leurs facultés réactives sont assurées par des modules asynchrones de communication et de perception. Ce modèle repose sur l'utilisation d'un ATN pour adapter le comportement des modules cognitifs à celui des modules réactifs et pour gérer les interactions entre ces différents modules. Chaque agent gère des graphes de dépendances et d'interférences pour assurer une cohérence et éviter les conflits avec les autres agents. Ce modèle repose sur le concept d'acteurs et nous nous plaçons délibérément dans le cadre de la programmation orientée-objet (Smalltalk-80). Cet environnement nous fournit de puissants outils de base sous la forme de divers "frameworks". Nos agents sont implémentés à l'aide de la dernière version d'Actalk, une plate-forme générique pour implémenter différents modèles d'acteurs. Les bases de règles des modules de raisonnement utilisent NéOpus, un moteur d'inférence d'ordre 1. La principale caractéristique de NéOpus est le contrôle déclaratif du raisonnement. Nous utilisons le framework de simulation à événements discrets pour représenter et gérer l'évolution temporelle de nos systèmes multi-agents. Nous utilisons également RPCtalk qui implémente la technique RPC, pour distribuer nos systèmes multi-agents sur plusieurs machines. Nous avons étendu ces différentes composantes pour les adapter au modèle d'agents proposé et réaliser la plate-forme multi-agents DIMA.

Pour valider DIMA, nous avons réalisé trois applications dont deux sont relativement importantes 1) Un système de monitorage de la ventilation artificielle des patients en soins intensifs et 2) Une modélisation de l'évolution économique et du changement technologique.

### Mots-clés

Agents, acteurs, objets, règles de production, métarègles, ATN, simulation à événements discrets

### Adresse

LAFORIA-IBP, case 169, 4 place Jussieu, F-75252 Paris cedex 05.

e-mail: guessoum@laforia.ibp.fr

### Pour l'envoi d'un résumé de thèse :

Il est conseillé d'envoyer un résumé en français (pour le Bulletin) et en anglais (résumé transmis à AI-Com) • saisir sous traitement de texte compatible avec Word si possible, en times 12, interligne moyen (ou équivalent); présentation (pour un doctorat d'Université)

### <Titre de la thèse>

<auteur> Thèse de l'Université <université>, soutenue le <date> [à lieu de soutenance>  $si \neq de < université > 1$ [préparée à ieu de préparation>  $si \neq de$  <université>]

<Résumé, environ une demie-page> Mots-clés <mots-clés> Adresse <adresse de l'auteur pour contact, demande d'exemplaire...>

• envoyer le fichier sur disquette Mac ou dossier attaché sous EUDORA à Bernard Leroux (voir page 2 de ce Bulletin).

### **A**PPELS A DOSSIERS

### **Dossier Intelligence Artificielle et Robotique**

A paraître dans le Bulletin de l'AFIA numéro 27 en Octobre 1996

### Appel à contributions

La robotique est un excellent support d'application, de développement et d'intégration de différents aspects de l'intelligence artificielle. On demande en effet aux robots, véritables machines intelligentes, d'être capables d'exécuter de manière autonome différentes tâches non répétitives en interaction avec un environnement pouvant subir des variations. Pour atteindre une telle autonomie fonctionnelle et décisionnelle, ils doivent être dotés de capacités de perception et de décision multiples et évoluées, et font pour cela appel à de nombreuses techniques de l'intelligence artificielle. Le but de ce dossier est de faire le point sur les récents progrès effectués dans le domaine de l'intelligence artificielle appliquée aux robots.

Le sujet étant vaste, les auteurs sont invités à situer leur contribution, indépendamment des techniques utilisées, par rapport à un ou plusieurs thèmes de la liste non exclusive suivante :

- Perception : représentations/modélisation/interprétation de l'environnement ;
- Décision : supervision, coordination, coopération multi-robots, planification de tâches, d'actions de déplacement et de perception;
- Raisonnement sur la mobilité et la manipulation ;
- Intégration : conception générale, architecture décisionnelle ;

### **Qui peut soumettre?**

Tout laboratoire ou industriel de la communauté francophone impliqué dans des actions de recherche et développement dans le domaine.

Les contributions devront prendre les formes classiques utilisées dans les dossier du bulletin, suivant le schéma suivant :

- Identification et coordonnées du chercheur ou de l'équipe
- Cadre général : mission de l'équipe, contexte matériel (robots, capteurs...), collaborations principales
- Problématique(s) considérée(s), position par rapport à l'état de l'art
- Panorama des travaux réalisés : techniques employées, résultats obtenus
- Présentation d'un ou deux projets représentatifs
- Courte bibliographie

Les contributions feront une ou deux pages et devront être envoyées avant le 1er Septembre 1996, par disquette ou par courrier électronique, au format Word5.1 Mac de préférence, Bin-Hex ou RTF sinon. Aucune mise en forme particulière ne doit être faite sur les textes sauf l'utilisation du gras ou de l'italique.

**Contact:** Simon LACROIX

LAAS-CNRS

Groupe Robotique et Intelligence Artificielle 7, Ave du Colonel Roche

31077 TOULOUSE Cedex

E-mail: Simon, Lacroix@laas.fr

### - BULLETIN DE L'AFIA -

### APPELS A DOSSIERS

### **Dossier Systèmes multi-agents**

coordonné par Anne Collinot et Alexis Drogoul A paraître en Janvier 1997

### Appel à contributions

Le thème de l'Intelligence Artificielle Distribuée (IAD) et des Systèmes Multi-Agents enrichit l'Intelligence Artificielle en suggérant l'usage de nouvelles métaphores sociologiques — ayant trait aux notions de coopération, de négociation, de groupe ou d'équipe, etc — ou biologiques — ayant trait aux notions d'auto-organisation, d'intelligence en essaim, etc. Les thèmes abordés dans le domaine incluent la modélisation, la conception et la programmation orientée-agent, la simulation multi-agent et vie artificielle, la problématique de l'organisation, l'évaluation de systèmes multi-agents, la définition de modèles de coopération, de coordination, de résolution de conflits et de négociation, la planification et apprentissage multi-agent, la réalisation d'environnements de développement.

#### Qui peut soumettre?

Tout laboratoire ou industriel de la communauté francophone impliqué dans des actions de recherche et développement dans le domaine.

#### Présentation des contributions

Les contributions adopterons le schéma de présentation suivant :

- Identification et coordonnées du chercheur ou de l'équipe
- Cadre général : objectifs, collaborations principales
- Problématiques considérées, position par rapport à l'état de l'art
- Panorama des réalisations
- Présentation d'un ou deux projets représentatifs
- Courte bibliographie: les 5 articles les plus significatifs de l'équipe

#### Format des contributions

Les contributions feront une ou deux pages et devront être envoyées avant le 15 Novembre 1996, par disquette ou par courrier électronique (Eudora, BinHex ou UUencode) au format Word5.1 Mac, Word6 pour Mac ou PC, RTF ou PostScript. Aucune mise en forme particulière ne doit être faite sur les textes sauf l'utilisation du gras ou de l'italique.

Contact: Alexis DROGOUL - LAFORIA-IBP Boite 169

Universite Paris VI - Tour 46-00

4, Place Jussieu - F-75252 PARIS Cedex 05

**E-mail:** drogoul@laforia.ibp.fr

### Dossier "I.A. et Image"

A paraître Juillet 1997

Appel à contributions :

présentation des équipes

#### **COORDINATION:**

Jean-Louis Amat

Institut Polytechnique de Sevenans, Belfort Contact: Jean-Louis.Amat@utbm.fr

**Catherine Garbay** 

Laboratoire. TIMC - IMAG Contact: Catherine.Garbay@imag.fr

# Les dossiers du Bulletin de l'AFIA

I.A. et biologie moléculaire	Bulletin n°26
I.A. et droit	Bulletin n°25
I.A. et fusion de données	Bulletin n°24
I.A. et musique	Bulletin n°23
Apprentissage	Bulletin n°22
Les explications dans les SBC	Bulletin n°20
Pétrole-Chimie	Bulletin n°19
Le raisonnement à partir de cas	Bulletin n°18
I.A. et temps-réel	
Planification et action	Bulletin n°16
Traitement automatique des langues	Bulletin n°15
1 A at 10 folio al 10 a	Rullotin nº1 <i>4</i>
I.A. et médecine	Dulletii 11 14
Diagnostic à base de modèles	
	Bulletin n°13
Diagnostic à base de modèles	Bulletin n°13 Bulletin n°12
Diagnostic à base de modèles  Validation des SBC	Bulletin n°13 Bulletin n°12 Bulletin n°11
Diagnostic à base de modèles	Bulletin n°13 Bulletin n°12 Bulletin n°11 Bulletin n°10
Diagnostic à base de modèles  Validation des SBC  Le connexionnisme  I.A. et jeux  E.I.A.O  I.A. et gestion	Bulletin n°13Bulletin n°12Bulletin n°11Bulletin n°10Bulletin n°9Bulletin n°8
Diagnostic à base de modèles	Bulletin n°13Bulletin n°12Bulletin n°11Bulletin n°10Bulletin n°9Bulletin n°8
Diagnostic à base de modèles  Validation des SBC  Le connexionnisme  I.A. et jeux  E.I.A.O  I.A. et gestion	Bulletin n°13 Bulletin n°12 Bulletin n°11 Bulletin n°10 Bulletin n°9 Bulletin n°8
Diagnostic à base de modèles  Validation des SBC  Le connexionnisme  I.A. et jeux  E.I.A.O.  I.A. et gestion  Conception et I.A.	Bulletin n°13Bulletin n°12Bulletin n°11Bulletin n°10Bulletin n°9Bulletin n°8Bulletin n°8
Diagnostic à base de modèles  Validation des SBC  Le connexionnisme  I.A. et jeux  E.I.A.O  I.A. et gestion  Conception et I.A  Intelligence artificielle distribuée	Bulletin n°13 Bulletin n°12 Bulletin n°11 Bulletin n°10 Bulletin n°9 Bulletin n°8 Bulletin n°7 Bulletin n°7

Adhésion et abonnement	□ Demande	☐ Renouvellement
Nom : Affiliation : Adresse postale : N° de téléphone : Adresse électronique :	Prénom : N° de fax :	
Adhésion simple 1996 (incluant l'a Adhésion étudiant 1996 (incluant Adhésion de soutien 1996 (incluant Adhésion 1996 (incluant l'abonnement à la Revue de l'Interpretation Abonnement au bulletin sans adhesion simple 1996 (incluant le *un formulaire d'adhésion AFCET Veuillez trouver un chèque de Je vous prie de m'envoyer une fa	l'abonnement et sur justific nt l'abonnement au bulletir ment au bulletin) elligence Artificielle:ésion (1996):	ratif):
Adhésion Personne morale Organisme:	☐ Demande cachet de l'organism	u
Nom du représentant : Fonction : Adresse postale :	Prénom :	
N° de téléphone :	N° de fax :	
Adresse électronique :	in de lax.	
1 1	<i>1996</i> 1500 F	
Adresse électronique :  Tarifs Laboratoires universitaires :	1996 1500 F aires : 2965 F	
Adresse électronique :  Tarifs Laboratoires universitaires : Personnes morales non universit  i je joins un bon de command	1996 1500 F aires: 2965 F  le pour une somme de ne facture de ci-dessus apparaissent da ci-dessus soient communichier IA européen.	ns l'annuaire de l'AFIA qués a l'ECCAI dans le cadre

Ce bulletin d'adhésion doit être retourné à : ( chèques à l'ordre de l'AFIA )

René QUINIOU: INRIA / IRISA, Campus de Beaulieu 35042 Rennes Cedex

Compte Crédit Mutuel de Bretagne code banque 15589 code guichet 35130 N° compte 03128396243 clé 23

# SOMMAIRE DU BULLETIN N°26

Editorial 3
La vie de l'A.F.I.A
Congrès A.F.I.A 6
Débat 8
Présentation des laboratoires
Dossier I.A. et biologie moléculaire19
Livres
Conférences
Calendrier42
Sommaire des revues44
Résumés de thèses
Appel à dossiers52

# CALENDRIER DE PARUTION DU BULLETIN DE L'AFIA

Hiver Eté

Réception des contributions:

Réception des contributions:

15 décembre

15 juin

Sortie le 15 janvier

Sortie fin juillet

Printemps

temps Automne

Réception des contributions:

Réception des contributions:

15 mars

30 septembre

Sortie le 15 avril

Sortie le 30 octobre