

*BULLETIN*

*DE*

*L'AFIA*



**ASSOCIATION FRANCAISE POUR  
L'INTELLIGENCE ARTIFICIELLE**

*JANVIER 2002*

*N° 48*

## Présentation du bulletin

Le **Bulletin de l'Association Française pour l'Intelligence Artificielle** 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 comptes rendus de conférences, les notes de lecture et les articles de débat sont très 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. En outre, les articles signés, de même que les contributions aux débats, reflètent le point de vue de leurs auteurs et n'engagent qu'eux-mêmes.

### Pour contacter l'AFIA

Président :  
**Bertrand BRAUNSCHWEIG**  
Institut Français du Pétrole  
Direction Informatique  
Mathématiques Appliquées  
B.P. 311  
92506 Rueil Malmaison CEDEX  
Mél. : Bertrand.Braunschweig@ifp.fr

*Contributions au bulletin*  
**Jean-Paul SANSONNET**  
voir ci-contre

*Serveur WEB*  
<http://www.afia.polytechnique.fr/>

*Adhésions,  
Liens avec les adhérents*  
**Marc AYEL**  
LIA-ESIGEC- Technolac  
Université de Savoie  
73376 Le Bourget du Lac CEDEX  
Mél. : Marc.Ayel@univ-savoie.fr

*Parrainage de manifestations*  
**Sylvie PESTY**  
Laboratoire Leibniz, Institut IMAG  
46, Avenue Félix Viallet  
38031 Grenoble CEDEX  
Mél. : Sylvie.Pesty@imag.fr

### Membres d'honneur

**Jaques Pitrat, Jean-Paul Haton, Marie-Odile Cordier  
Jean-Marc David, Daniel Kayser, Claude Vogel**

### Personnes morales adhérentes à l'AFIA

ADIT, AI\*IA, CNET PARIS-A, ENS Mines, ENST-PARIS, ESIEA, IGN,  
ILOG, INRETS, INRIA, INSTITUT FRANCAIS DU PÉTROLE,  
PEUGEOT S.A., Université de Savoie, Université PARIS 9 DAUPHINE.

### Bureau de l'AFIA

**Bertrand BRAUNSCHWEIG, président**  
**Serge DUPUY, secrétaire**  
**Marc AYEL, trésorier**

**François ARLABOSSE, Jean-Paul BAQUIAST, Gilles BISSON,  
Jérôme EUZENAT, Alain CARDON, Catherine GARBAY,  
Jean-Daniel ZUCKER, Eunika MERCIER-LAURENT, Sylvie PESTY,  
Gérard SABAH, Jean-Paul SANSONNET, Marc SCHOENAUER,  
Michèle SEBAG, Laurent SIKLÓSSY, Fabien TORRE.**

## Comité de rédaction

**Jean-Paul SANSONNET**  
**Rédacteur en chef**  
LIMSI — CNRS  
B.P. 133, 91403 ORSAY Cedex  
jps@limsi.fr

**Jean-Paul BAQUIAST**  
**Rubrique « Groupes de travail et collègues »**  
72 bis, Rue de Paris,  
92190 Meudon  
jean-paul.baquiast@wanadoo.fr

**Brigitte GRAU**  
**Rubriques « Sommaires des revues »  
et « petites annonces »**  
LIMSI — CNRS  
B.P. 133, 91403 ORSAY Cedex  
grau@limsi.fr

**Catherine BARRY-GRÉBOVAL**  
**Rubrique « Présentation de laboratoires »**  
PSI — INSA de Rouen  
place Émile Blondel BP 08  
76131 Mont Saint Aignan Cedex  
Catherine.Greboval@insa-rouen.fr

**Marc-Philippe HUGET**  
**Rubrique « Thèses et habilitations »**  
Agent ART Group, Department of Computer  
Science  
University of Liverpool  
LIVERPOOL L69 7ZF United Kingdom  
M.P.Huget@csc.liv.ac.uk

**Amedeo NAPOLI**  
**Rubrique « livres »**  
LORIA  
B.P. 239  
F-54 506 Vandœuvre lès Nancy  
Amedeo.Napoli@loria.fr

**Sylvie PESTY**  
**Rubriques « Conférences  
et comptes rendus »**  
Laboratoire Leibniz-IMAG  
46 Avenue Félix Viallet  
F-38031 Grenoble Cedex  
Sylvie.Pesty@imag.fr

**Nicole TOURIGNY**  
**Rubrique « IA au Québec »**  
Université Laval - Pavillon Adrien-Pouliot  
Québec, Canada, G1K 7P4  
tourigny@ift.ulaval.ca

**Jean-Daniel ZUCKER**  
**Rubrique « Débats »**  
LIP6, Université Paris 6  
4, Place Jussieu, 75232 Paris Cedex 05  
Jean-Daniel.Zucker@lip6.fr



Réa. D.E.B. Copy – 01 46 31 64 53  
ISSN 1273-1323  
Dépot légal janvier 2002

## Appels à contributions

Avec ce premier numéro de l'année 2002, le Comité de Rédaction de l'AFIA est heureux d'accueillir en son sein Pierre-Marc Huget qui reprend désormais la rubrique « Compte-rendus de thèses et d'habilitations » tenue précédemment par Amédéo Napoli que nous remercions pour son travail. Après un parcours universitaire classique en informatique, Pierre-Marc Huget a obtenu sa thèse en informatique de l'université de Paris 9 en juin 2001, thèse qui a été réalisée au sein de l'équipe MAGMA à Grenoble. Depuis cette année, il est « Research assistant » au sein de l'équipe Agent ART.de l'Université de Liverpool. Son travail de recherche porte sur la spécification et la validation des langages et des protocoles d'interaction dans les systèmes multi-agents sous la direction de Mike Wooldridge et Simon Parsons. Pour sa part, Amédéo Napoli va continuer à s'occuper de la rubrique « Revues de livres » aussi, si vous avez lu un livre qui vous a plu, n'hésitez pas à envoyer à Amédéo un demi-page ou un page de commentaires sur ce livre.

le Comité de Rédaction de l'AFIA commence à mettre en place les idées qui ont été proposées à l'automne dernier par le bureau de l'AFIA. Cela se traduit d'abord par la parution des prix AFIA pour les conférences qui sont parrainées par l'AFIA. Ainsi, on trouvera à la rubrique « Vie de l'AFIA » un premier compte-rendu de ces prix qui visent avant tout à mettre en avant les jeunes chercheurs en IA. Dans ce même esprit, nous présentons un résumé du travail de synthèse qui a été effectué à l'IRIT sur le thème « Qu'est-ce que l'IA » et qui peut être accédé via le web.

Pierre Marquis et Jérôme Lang nous présentent dans la rubrique « Vie de l'AFIA » la nouvelle revue électronique de l'AFIA, appelée « JEDAI », dont le site Internet vient d'être rendu opérationnel et où les premiers articles commencent à arriver dans le « petit bain ». Cette revue se présente sous un format original qui devrait faciliter la proposition d'articles de revue en IA, en particulier pour les jeunes chercheurs. Il s'agit là encore d'une expérimentation qui met en œuvre la politique d'innovation de l'AFIA et qui ne pourra vraiment réussir qu'avec la participation de la communauté : aussi nous faisons un large appel à contributions pour que la revue JEDAI se développe rapidement.

Toujours dans l'esprit d'innovation, le prochain bulletin de l'AFIA, qui va paraître en mars 2002, sera un numéro spécial, lié à la présence en France de la Conférence ECAI, en Juillet 2002 (Voir [http://ecai2002.univ-lyon1.fr/show\\_fr.pl](http://ecai2002.univ-lyon1.fr/show_fr.pl)). Ce numéro spécial contiendra essentiellement une revue des équipes Françaises en IA et ses applications. Dès maintenant, un site est ouvert où toutes les équipes qui souhaitent paraître dans ce numéro peuvent grâce à un formulaire en ligne entrer leur description. Encore une fois, nous appelons la communauté à contribuer largement. Comment faire figurer votre équipe dans ce numéro ? Il suffit de vous connecter au site Web : <http://www-leibniz.imag.fr/AFIA/>, vous y trouverez un court formulaire qui, outre quelques informations administratives, vous demandera de présenter les thèmes de recherche de votre équipe en une quinzaine de lignes (pour plus d'information, le contact est Gilles Bisson : [Gilles.Bisson@imag.fr](mailto:Gilles.Bisson@imag.fr)).

Pour les numéros du Bulletin qui vont paraître à l'été et à l'automne 2002 nous avons besoin de contributions pour alimenter la nouvelle rubrique « Synthèses ». Il s'agit de proposer des textes qui font le point sur l'état de l'art d'une problématique déjà bien développée ou bien qui sont une introduction à de nouvelles problématiques qui émergent en IA ou en connexion avec l'IA. Contactez Jean-Paul Sansonnet ([jps@limsi.fr](mailto:jps@limsi.fr)) pour faire une proposition.

Voilà beaucoup d'appels à des contributions et sous plusieurs formes : descriptions d'équipes, articles de revues, textes de synthèses etc. Pour être au service de la communauté, l'AFIA a besoin de votre participation, en particulier celle des jeunes chercheurs, aussi n'hésitez pas à nous contacter.

**Jean-Paul Sansonnet**

*Rédacteur en chef du bulletin de l'AFIA*

## L'Intelligence Artificielle, mais enfin de quoi s'agit-il?

L'Intelligence Artificielle (IA), en concevant et en réalisant des fonctions cognitives artificielles, et grâce aussi à la puissance calculatoire des ordinateurs, permet de faire effectuer des tâches réputées intelligentes à des machines, généralement de manière à la fois plus fiable et plus rapide que ne le ferait l'homme, tout en exploitant généralement de grandes quantités d'information. Cela donne lieu à des réalisations pouvant être éventuellement spectaculaires, ou perturbantes pour l'idée qu'on se fait de l'intelligence comme attribut spécifique de l'homme. Les recherches en IA, pas toujours bien comprises, alimentent ainsi les légitimes questions du profane.

Aussi à l'automne 2001, un livret (de 64 pages, au format 10,5 - 23 cm), intitulé « L'Intelligence Artificielle, mais enfin de quoi s'agit-il? » a-t-il été préparé par des chercheurs de l'IRIT. L'ambition de ce livret est d'apporter une image un peu structurée, concise quoique précise, des préoccupations de la recherche en Intelligence Artificielle aujourd'hui au travers de la présentation des principaux paradigmes (de base ou applicatifs) qu'elle étudie, en indiquant sa place vis-à-vis des autres disciplines scientifiques, ainsi que l'évolution de sa présence dans les productions culturelles. Ce livret veut offrir un panorama des grandes problématiques de l'IA aujourd'hui, en aidant à mieux la faire comprendre par le non-spécialiste. Son contenu est ainsi agencé :

- Avant-propos (H. Prade)
- Un peu d'histoire (H. Prade)

### *Dix grands paradigmes de recherche :*

- Représenter l'information, acquérir des connaissances (N. Aussenac et J.-L. Soubie)
- Algorithmes généraux de résolution de problèmes (H. Farreny)
- Intelligence artificielle collective (M.-P. Gleizes et P. Glize)
- Formaliser et mécaniser différents types de raisonnement (S. Benferhat, C. Cayrol, D. Dubois, H. Prade)
- Evaluer des situations, décider, planifier (D. Dubois, H. Fargier, J. Lang, H. Prade)
- Raisonner sur le changement, le temps, et l'espace (A. Herzig et L. Vieu)
- Résumer, apprendre, découvrir (G. Richard)

- Langue et IA: des interrelations étroites (M. Bras et P. Saint-Dizier)
- Indexation multimédia et IA (R. André-Obrecht et P. Joly)
- La réalité virtuelle et l'IA (Y. Duthen et J.-P. Jessel)

### *La Place de l'IA dans les Sciences :*

- IA et mathématiques (D. Dubois, L. Farinas del Cerro, H. Prade)
- IA et informatique (L. Farinas del Cerro, L. Sais)
- L'IA et les sciences du traitement de l'information (D. Dubois, H. Prade)
- Sciences de la cognition et IA (M. Borillo)
- L'IA et la théorie des systèmes symboliques (J. Virbel)
- L'IA, un voyage dans notre culture (C. Gadon, Resp. Service Culture de l'Univ. P. Sabatier)
- Pour conclure ... (M. Borillo, L. Farinas del Cerro, H. Prade)

### **Quelques références.**

Comme on peut le voir, le territoire de l'IA est ici circonscrit aux travaux visant à rendre la machine capable d'acquérir de l'information, de raisonner sur une situation statique ou dynamique, de résoudre des problèmes combinatoires, de faire un diagnostic, de proposer une décision, un plan d'action, de communiquer et d'expliquer les conclusions qu'elle obtient, de comprendre un texte ou un dialogue en langage naturelle, de résumer, d'apprendre, de découvrir des informations significatives, et ce en utilisant des méthodes suffisamment génériques pour s'adapter à de larges classes de situations. Des problématiques mettant en jeu d'autres aspects de la cognition, telles que la vision par ordinateur, la reconnaissance de la parole, ou la robotique, qui entretiennent avec l'IA des relations privilégiées, ne sont pas couvertes par ce livret.

Le livret est disponible, sur simple demande auprès du Service Culture de l'Université Paul Sabatier (livrets.culture@adm.ups-tlse.fr). Des versions html et pdf se trouvent sur le site <http://www.irit.fr/Livret-IA/>.

## Annnonce de journal électronique Journal Électronique d'Intelligence Artificielle

JEDAI

<http://jedai.afia-france.org>

JEDAI (*Journal Électronique d'Intelligence Artificielle*) est la revue électronique d'intelligence artificielle de l'Association Française pour l'Intelligence Artificielle (AFIA). JEDAI est dédié à la communication rapide de résultats de recherches en intelligence artificielle, à destination de la communauté des chercheurs francophones.

Tout article de recherche en intelligence artificielle peut être soumis à JEDAI. Dans un souci de compléter les revues traditionnelles, sont particulièrement bienvenus :

- les articles de type « note de recherche » (techniques et non auto-suffisants) qui nécessitent une diffusion rapide.
- les articles centrés sur les applications de techniques d'intelligence artificielle à des problèmes industriels d'envergure.
- les articles de synthèse.

JEDAI vise à étendre le rôle des journaux scientifiques classiques en permettant une diffusion rapide et une discussion ouverte des travaux soumis. Il tire parti des possibilités de la Toile en offrant la tenue de forum de discussions autour des articles, permettant de faciliter la maturation des travaux présentés au moyen de dialogues et débats ouverts. Ainsi, les articles de JEDAI peuvent être lus par tous. En revanche, il est nécessaire de s'identifier (par son adresse électronique) et de s'inscrire pour pouvoir soumettre un article ou en commenter un. Les articles communiqués à JEDAI sont librement accessibles sur le site du journal. Ils peuvent aussi être commentés librement par les personnes enregistrées auprès de JEDAI.

JEDAI a également pour but d'assurer les fonctions plus traditionnelles d'évaluation et d'archivage des publications scientifiques. Ainsi, les articles disponibles sur le site sont séparés en articles soumis (le « petit

bain ») et articles acceptés après une évaluation par les pairs (le « grand bain ») qui constituent, en tant que telles, des « publications dans une revue à comité de lecture ». Un point important ici est l'absence de copyright : les auteurs des publications conservent leurs droits d'auteurs et de reproduction. Ils ne concèdent à JEDAI qu'un droit limité de diffusion de leur article sur le site de JEDAI. Les auteurs peuvent retirer ce droit à tout moment (par exemple parce que leur article est accepté par une autre revue qui réclame les droits exclusifs de diffusion). Dans ce cas, l'article ne sera plus directement accessible à partir du site de JEDAI, mais une page indiquant que l'article a été accepté par JEDAI mais a été retiré par les auteurs subsistera.

Le processus de soumission d'un article à JEDAI est le suivant :

- 1) **Pédiluve** : l'article est soumis à JEDAI et après le contrôle de son adéquation dans un délai rapide (moins d'une semaine), il est plongé dans le petit bain.
- 2) **Petit bain** : l'article est ouvert aux commentaires de la communauté, sous la forme d'un débat animé par un ou plusieurs parrains scientifiques (membres du comité de lecture). La discussion doit permettre aux auteurs de réagir rapidement pour clarifier ou approfondir leur propos.
- 3) **Du petit au grand bain** : le déclencheur du passage dans le grand bain est la proposition de l'un des parrains scientifiques. Au plus tard trois mois après la soumission d'un article dans le petit bain, un parmi les parrains devra l'avoir proposé pour le grand bain, ou un rapport résumant les modifications suggérées et demandées sera fourni à l'auteur. Deux ou trois relecteurs anonymes et indépendants reliront alors l'article et rendront leur avis. Sur la base de ces rapports, le comité éditorial rendra l'un des trois avis suivants :

- acceptation de l'article dans le grand bain ;

- demandes de modifications majeures (possibles une seule fois) ;
- rejet.

4) **Grand bain** : Il s'agit des articles proprement sélectionnés par JEDAI suivant les standards des revues scientifiques classiques.

Pour l'instant, aucun format au sens traditionnel du terme n'est requis pour les articles soumis. Il semble raisonnable de ne pas dépasser les 30 pages. Les formats électroniques acceptés sont précisés dans l'aide en ligne. La langue dans laquelle les articles sont rédigés n'est pas imposée bien que JEDAI ait une vocation naturelle à accueillir des articles en français (veuillez considérer ETAI pour les articles en anglais).

Dans l'avenir, il n'est pas impossible, afin d'assurer la pérennité des articles publiés par JEDAI, que ceux-ci fassent l'objet de publication sur support papier et/ou que JEDAI adopte une politique très libérale en terme de miroir de son grand bain.

Plongez dans le bain !

## Rédacteurs en chef

Jérôme Lang (IRIT)  
Pierre Marquis (CRIL)

## Comité éditorial de JEDAI

Rachid Alami LAAS, Toulouse  
Philippe Balbiani IRIT, Toulouse  
Christian Bessière LIRMM, Montpellier  
Bertrand Braunschweig IFP, Paris  
Jean Charlet Assistance publique, Paris  
Marie-Odile Cordier IRISA, Rennes  
Christophe Dousson France Telecom, Lannion  
Jérôme Euzenat INRIA Rhône-Alpes, Montbonnot  
Frédéric Garcia INRA, Toulouse  
Daniel Kayser LIPN, Villetaneuse  
Yves Lespérance York University, Ontario, Canada  
Jérôme Lang IRIT, Toulouse  
Evelyne Lutton INRIA, Rocquencourt  
Pierre Marquis CRIL, Lens  
Amedeo Napoli LORIA, Nancy  
Patrice Perny LIP6, Paris  
Jacques Pitrat LIP6, Paris  
Chantal Reynaud LRI, Orsay  
Céline Rouveirol LRI, Orsay  
Pascale Sébillot IRISA, Rennes  
Thomas Schiex INRA, Toulouse  
Camilla Schwind LIM, Marseille  
Michèle Sebag École polytechnique, Palaiseau  
Patrick Taillibert Detexis, Thomson CSF, Paris

## Prix AFIA

### Prix décernés pour les meilleurs articles dans le cadre de conférences qui sont parrainées par l'AFIA.

#### Conférence : LFA :

F.T. Bujor, L. Valet, E. Trouvé, G. Mauris et Ph. Bolon

« Détection de changements dans les images RSO par fusion floue interactive d'informations »

LAMII/CESALP (Université de Savoie, Ecole Supérieure d'Ingénieurs d'Annecy)

<http://www.hds.utc.fr/lfa2001>.

[valet@esia.univ-savoie.fr](mailto:valet@esia.univ-savoie.fr)

[mauris@esia.univ-savoie.fr](mailto:mauris@esia.univ-savoie.fr)

#### Conférence : Valgo :

Emilie Poisson et Christian Viard-Gaudin

Réseaux de neurones à convolution : reconnaissance de l'écriture manuscrite non contrainte de.

Ecole polytechnique de l'université de Nantes  
Equipe IVC Rue Christian Pauc - La Chantrerie BP 50609 44306 Nantes Cedex 3

[http://www.supelec-rennes.fr/acth/valgo/Valgo\\_Numero-01-02.html#Poisson](http://www.supelec-rennes.fr/acth/valgo/Valgo_Numero-01-02.html#Poisson)

## Conférence : EA :

Rodolphe Leriche and Frédéric  
Dual Evolutionary Optimisation  
Université de St Etienne,

## Conférence : JFIADSMA :

Olivier Simonin  
Modélisation des satisfactions personnelles et inter-actives d'agents situés coopératifs. Olivier Simonin et Jacques Ferber ,  
LIRMM Université Montpellier II – CNRS  
161, rue Ada – 34392 Montpellier Cedex 5 – France  
*{simonin,ferber}@lirmm.fr*

## Résumé

Cet article aborde le problème de l'évaluation par l'agent de ses actions et interactions. Cette capacité est étudiée pour permettre à l'agent (situé) de déterminer les meilleures actions à effectuer pour poursuivre ses propres objectifs ou pour aider ses voisins (coopération).

Notre approche est basée sur un modèle des satisfactions de l'agent différenciant actions individuelles et interactions avec ses voisins.

La satisfaction personnelle est une valeur calculée incrémentalement dans le temps suivant la perception de la progression de la tâche en cours. Pour cela nous étendons une mesure nommée « progress estimator », introduite par M.J. Mataric, en pénalisant plus fortement l'immobilisation de l'agent que son éloignement du but. Ainsi, une satisfaction positive motive l'agent à

poursuivre sa tâche en cours, au contraire, une valeur négative le force à l'abandonner.

La satisfaction interactive est une évaluation de la présence des agents voisins et de leurs actions: gêne, coopération (potentielle), indifférence. Cette évaluation est transformée en signaux d'intentions, sous formes d'attractions ou de répulsions, agissant comme des champs de potentiels dynamiques entre agents.

Dans le cadre du développement de robots mobiles coopératifs, nous proposons une architecture réactive bâtie sur le principe précédent. Le module de sélection d'actions est basé sur la compétition continue entre la satisfaction personnelle de l'agent et les signaux qu'il perçoit. Ainsi, celui-ci ne répond aux signaux extérieurs que s'ils sont plus intenses que le poids et la progression de sa tâche courante.

La réponse aux signaux, nommée réaction altruiste, est définie par un vecteur de déplacement dont le sens dépend du signe de la valeur communiquée. Ce vecteur est alors combinable à d'autres, comme ceux de l'évitement des obstacles, et assure des comportements coopératifs rapides.

Nous montrons comment l'application de cette architecture permet de résoudre différents problèmes physiquement distribués. Nous présentons en particulier le traitement de conflits spatiaux apparaissant entre des agents réactifs dans des environnements très contraints. Pour cela, les signaux d'insatisfaction sont propagés afin de créer des chaînes répulsives des agents les plus contraints (les plus insatisfaits) vers les agents les plus libres spatialement.

Le modèle proposé est validé expérimentalement par la résolution de ce type de conflits entre robots mobiles autonomes.

## Améliorer ou supprimer les referees ?

Nous avons lancé un débat intitulé « *Évaluation Scientifique : peut-on mieux faire ?* ». Suite à la contribution de Jacques Pitrat dans le n° 45 et à la réponse de F. Pachet dans le n° 46/47 (sous le nom #FP/09.01), nous poursuivons les réactions envoyées à Jean-Daniel Zucker (toutes reçues sous forme de courrier électronique) :

**Francois PACHET** (pachet@csl.sony.fr)

Voici par exemple ce que dit Nature de son processus de sélection: « Note that very few papers are published without significant revisions to the original submission. Some of the papers of which Nature is most proud of publishing are those which were originally weak but intriguing, and have become really striking only by virtue of extensive attention from the referees and editors. »

Francois Pachet SONY CSL-Paris E-mail:  
pachet@csl.sony.fr

**Juliette ROUCHIER** (j.rouchier@mmu.ac.uk)

J'ai reçu le texte de M. Pitrat, et je suis contente de voir que le sujet qu'il aborde est toujours d'actualité. Quelques remarques que cela m'inspire.

1. Il semble très positif que ce sujet soit évoqué de façon publique et pas seulement dans des discussions privées en entre chercheurs qui veulent s'accorder sur des idéaux d'organisation de la recherche. Il est clair que les résultats des processus de sélection font qu'on peut effectivement vouloir les réformer : il est agaçant de voir certains papiers rejettes quand ils semblent bon mais a retravailler, alors que des choses moyennes sont publiées. Souvent, l'éditeur devrait soutenir plus les auteurs dans le travail de calibrage et réécriture en fonction de la ligne éditoriale. Pour cela il faut une ligne éditoriale claire, c'est sur. Plus graves : certaines formes de pensée sont interdites. c'est clair, c'est vrai, même avec peu de recul dans le domaine de la recherche on peut en avoir l'intuition. il y a des aspects ou cela peut être sympathique (thèses révisionnistes), mais ce n'est pas forcément pertinent.

Une note : un collègue du Centre for Policy Modelling - Bruce Edmonds - a écrit un texte disponible sur le oueb, en anglais, traitant du même sujet // <http://www.press.umich.edu/jep/05-04/edmonds.html> et faisant des propositions un peu sauvages, voire terroristes (comme : aller sur les pages des auteurs qui publient des papiers non reviewés pour faire des reviews de ces papiers et les mettre on-line soit même - pas aberrant selon certains points de vue, mais amenant certainement a des débauches un peu inutiles d'agressivité).

2. (Contre l'optimisme) je tends a douter de façon très profonde du fait que le oueb rendrait l'information plus "libre" ou meilleure d'une quelconque façon. certaines contraintes seraient certainement libérées mais celles que l'on récupérerait en conséquence n'ont pas de garantie d'être plus légères, et surtout, elles nous sont inconnues donc moins contrôlables. - déjà, l'accès aux information sur internet nécessite d'avoir beaucoup de temps ou d'appartenir a des listes de diffusion d'information efficaces et pertinentes. En détruisant une forme de pouvoir liée a la situation d'éditeur. on décale les problèmes et on crée en fait un pouvoir parallèle, plus facile a associer a du lobbying : des gros groupes forts en communication détiennent la nouvelle légitimité.

Si la publication sur papier est clairement limitée par des jeux de pouvoir (principalement quelle école domine à un moment donné), malgré tout intérêt a mettre la pensée en avant, je ne crois pas qu'une diffusion trop large, trop facile, puisse garantir de retrouver des textes les plus importants à un moment donné. Dans tous les cas, si la publication ou la non publication d'un texte est affaire de pouvoir, les réseaux d'information aussi le sont.

(À ce niveau, je remarque que j'ai dit que l'injustice existera toujours et que la démocratie n'est pas pour demain... grande nouvelle) la situation de lobbying n'est pas forcément plus attirante, et c'est pourquoi il semble nécessaire de s'assurer la capacité de créer des canaux [ou « portails », très à la mode, non ?] d'information les plus objectifs possibles. a ce niveau, déjà, la question de l'organisation presque spontanée (du oueb) devient moins claire, et les sous-groupes réapparaissent. la lisibilité qu'on pouvait avoir gagnée n'est peut-être plus garantie sur le réseau.

- À ce niveau : je fais une observation directe sur le texte : la comparaison entre « connaître ses bons critiques » pour les textes scientifiques et pour les films ne me semble pas une comparaison pertinente. Même si DE FAIT, les gens agissent comme cela, il y a un risque à se comporter ainsi quand il s'agit de textes scientifiques puisqu'on tombe automatiquement dans l'habitude, on oublie la confrontation ,... tout ce qu'on voulait voir réintroduit.

Le cinéma traite de problèmes de **subjectivité exprimée**, le critique réintroduit de l'objectivité pour rendre lisible la subjectivité. Le reviewer devrait garantir l'objectivité du texte originel. J'ai l'impression que la question est très différente (mais je me trompe certainement).

- Je sais que j'ai appris le travail de chercheur en lisant des textes publiés dans des revues. tout en n'étant peut-être pas complet, cet apprentissage m'a permis de fixer des grilles « ce qu'il est normal de penser », « ce qui n'est pas normal de penser », « ce qui prête à controverse », « ce qui est nouveau ». je pense que pour l'instant il est tout de même possible d'avoir accès à des idées controversées quand on est vraiment impliqué dans un domaine (mais là, c'est peut-être l'inexpérience qui parle), et surtout quand on les cherche (les idées controversées). [notons qu'en terme d'apprentissage, d'ailleurs, il me semble qu'il est bien que de nombreux articles peu innovants mais très clairs, voire « exemplaires » soient publiés. je préfère cela à l'existence de quelques ouvrages seulement sur un sujet - cela revient à avoir en fait des cours distribués entre un grand nombre d'auteurs, et au vu de la complexité des disciplines dans lesquelles nous naviguons, on ne se sent pas étouffé mais bien aidé par cette multiplicité, parfois...]

- Je voulais juste signaler que nous avons essayé de mettre en place l'an dernier. l'idée était d'une revue électronique où les textes ET les reviews seraient publiés, voire des commentaires sur les reviews..... les protocoles d'entrée seraient le plus flexible possible (en fait nous n'étions pas tombés d'accord, dès le départ, sur ce degré de flexibilité, ce qui montre en fait ou est l'enjeu majeur de toute l'organisation, pour l'instant : qu'est-on prêt à voir reconnu comme textes scientifiques, même pour un temps court et même si des critiques sévères de la scientificité sont exposées en parallèle / quel risque à mettre des textes très mauvais dont on signale qu'ils le sont ? (à part la réputation de l'éditeur,

une fois de plus)). Le but de ce système était de forcer les reviewers à faire un travail correct (puisque'il serait public) tout autant que de rendre possible une plus grande publication d'idées. comme je disais, par manque de temps ce projet est tombé à l'eau il y a presque un an, mais il y avait une proposition d'action qui peut certainement toujours tenir le coup si des gens veulent s'y mettre.....

3. Provocation – un peu : doit-on retirer au chercheur le plaisir d'être un chercheur maudit ? s'il sait qu'il a raison, quel sens a la reconnaissance de ses contemporains. Et surtout quel intérêt sans passage à la postérité, qui elle ne sera jamais garantie ?

**Nicolas GLADE (GLADE@dsvsud.cea.fr)**

Je suis bien d'accord avec tout ça, cependant, même si je suis complètement partisan de l'idée (de la publication on-line sans referees. Pour l'instant les cyber-referees me semblent un rien utopiques, mais pourquoi pas ...), et en admettant qu'on résolve le problème de la publication de tout et n'importe quoi et de la fraude (ce que je pense tout à fait faisable effectivement), il reste le dernier point qui, à mon humble avis, est relativement insurmontable dans le contexte scientifique actuel. Je n'approuve, bien entendu, pas du tout ce mode de sélection des chercheurs par la liste des publications, mais la réalité est qu'il faut bien un mode de sélection, le nombre de places étant rarement abondantes. Vous dites « nous savons tous très bien faire la différence entre les chercheurs dont les travaux nous intéressent et ceux qui publient pour publier ». Oui, certes, mais je ne suis pas convaincu que les membres du jury d'un concours d'entrée dans un labo aient très envie de se taper toute la biblio (en détail !) de chacune des personnes qui se présentent pour être bien sûr de prendre le « bon élément ». Mais je me trompe peut-être, il y a sans doute des acharnés de la lecture prêts à dédier leur vie à la sélection de chercheurs... Et en admettant qu'on passe progressivement de ce mode de sélection assez stupide à quelque chose de nouveau, à moins que toute la communauté scientifique (mondiale) ne suive cette voie d'un coup, ce que je pense utopique, il en faudra quelque uns pour tenter leur chance dans ce nouveau concept (qui aura des chances de ne pas être bien reconnu). Deux problèmes alors : qui est prêt à se « sacrifier » en premier pour le bien des futurs chercheurs pleins d'idées ? (à mon avis pas grand monde) et

si quelques personnes se jettent à l'eau, cela suffira t'il pour faire prendre la greffe ? (ça dépend du nombre de personnes, mais c'est pas gagné). Il y a beaucoup de problèmes liés à cette évolution potentielle et je ne suis pas bien certain de les voir résolus un jour ... mais peut être suis-je trop pessimiste.

Nicolas Glade

**Jean-Pierre BACHY (Jean-pierre.bachy@imag.fr)**

J'ai été très intéressé par le texte de Jacques Pitrat « Améliorer ou supprimer les referees ». J'ai été sensible au problème de la publication par un chercheur, en dehors de son domaine d'expertise, d'une idée éventuellement farfelue et en tout cas mal cadrée dans le contexte d'une discipline qu'il ne maîtrise pas. Je suis tout à fait d'accord avec ses propositions, pour en être, de façon pratique, arrivé à la même conclusion (site personnel reprenant un travail présenté en séminaire de modélisation en 93). Cette évolution me paraît de toute façon inéluctable en raison de la facilité d'accès à ce média et il me paraît de la responsabilité du scientifique d'être présent sur ce marché incontrôlé de la nouveauté. Il est significatif que ce soit un chercheur au fait du connexionisme qui attire notre attention dans ce domaine : on passerait d'un contrôle supervisé du réseau à une dynamique propre du réseau avec rétroaction. Attendons-nous à des régimes chaotiques et considérons les comme des facteurs de bonne santé et de capacité d'éveil.

Jean-Pierre BACHY Pédiatre Attaché au PRETA, TIMC IMAG

Pour alimenter le débat :

<http://interactif.lemonde.fr/article/0,5611,2864--126488-0,FF.html>

**Josette PASTOR (josette.pastor@purpan.inserm.fr)**

Si je suis d'accord avec Jacques Pitrat sur certains points : - Le non-sens de l'évaluation fondée sur les « impact factors » dont on sait qu'ils favorisent les revues généralistes, parce qu'elles sont lues par une large communauté, alors que leurs critères d'acceptation semblent quelquefois douteux (notoriété de l'équipe - l'anonymat n'existe pas quand on sait qui fait quoi - ou phénomène de mode). - La frilosité des referees qui

sont l'expression de la pensée scientifique dominante du moment, et qui reprennent là le vieux rôle des "savants officiels" qui barrent la route aux idées novatrices (nous avons fait quelques progrès depuis Galilée, puisque on n'enferme plus les chercheurs non conventionnels en prison, mais on les « met au placard ») - L'absurdité de l'équation « production scientifique = production d'articles » qui conduit à une surenchère de la production et à favoriser « ceux qui disent » au détriment de « ceux qui font ».

Je modulerais cependant certains autres propos qu'il tient : - L'opposition méthode conjonctive-méthode disjonctive me paraît logiquement fondée, mais tous les exemples montrent simplement que la méthode disjonctive est seule utilisée puisque le CRITERE DE CREATIVITE n'est pas utilisé (ex de la lettre de Rob Milne en particulier) ... - Je ne lierais pas le rejet des idées novatrices à leur faiblesse d'expression pour au moins deux raisons :

- l'effort de communiquer pour rendre son idée compréhensible, ne peut qu'être positif. La plupart du temps, les problèmes d'expression viennent du fait que l'idée est encore du domaine de l'intuition et n'est pas assez étayée ou que celui qui en est à l'origine ne la maîtrise pas assez ... Si on n'est pas dans ces deux cas, je pense que la courtoisie veut que celui qui dit fasse l'effort de se faire comprendre et ne laisse pas tout le travail de communication à celui qui lit ou écoute. Et comme on le sait depuis longtemps « ce qui se conçoit bien s'énonce clairement et les mots pour le dire viennent aisément ».
- Les règles de jugement sont surtout fondées sur le « scientifiquement correct » et le texte, politiquement correct, de Milne l'exprimait parfaitement. Si les idées novatrices ont été repérées par les referees, c'est qu'elles étaient suffisamment bien exprimées. Donc, ce n'est pas la faiblesse de forme qui les a fait rejeter... La vraie raison est le fait que novateur entraîne rejet plutôt qu'acceptation. Cette raison est dans quelque chose que connaissent bien les psychologues travaillant sur le raisonnement humain : LE BIAIS DE CONFIRMATION, ou tendance humaine à chercher l'information concordant avec les croyances, théories ou hypothèses et à éviter de recueillir l'information potentiellement falsificatrice

En conclusion, le débat est moins sur les critères d'évaluation utilisés que sur notre capacité à accepter ce

qui n'est pas dans la pensée dominante actuelle. Tout système d'évaluation se heurtera à ce genre d'écueil. La proposition de J. Pitrat de publications libres sur site personnel pointées par d'autres pourrait faire naître d'autres reférees qui auraient certes l'avantage de ne plus être anonymes, mais qui par là même seraient crues en fonction de la confiance qui leur est accordée ... et qui créeraient donc un autre système de pensée dominante .... Je crois qu'elle n'est acceptable que si en plus des « gentils » qui aiment, les « méchants » qui détestent et disent pourquoi ils détestent contrebalancent le jugement.

Pardonnez ce long commentaire. Cordialement  
Dr Josette PASTOR,  
HDR INSERM U455 Fédération de neurologie

**Norbert NOURY** (Norbert.Noury@imag.fr)

C'est particulièrement intéressant, nous sommes trop jugés sur la forme, pas assez sur le fond, du coup ce sont les meilleurs communicants, ou ceux qui crient le plus fort, qui sont le mieux reconnus. Ceux qui bossent restent dans l'ombre.

**Samuel LANDAU** (samuel.landau@lip6.fr)

C'est très intéressant, je suis d'accord sur le fond, surtout car ça présente l'avantage de forcer tout le monde à s'exprimer plus clairement d'entrée de jeu, puisque n'importe qui peut venir lire et commenter les

articles.

Mais sur la forme, je trouve que ça ne prend pas assez en compte des dimensions humaines :

- l'attachement au papier, l'attachement à savoir que circule un papier avec son nom dans le vrai monde réel (et pas juste virtuel). Le papier a traversé les millénaires, alors qu'on ne peut plus lire les supports d'information des débuts de l'informatique. Ces supports informatiques sont très volatiles, à chaque génération on ne reproduit pas toutes les informations disponibles sur celle qui précède... C'est ancré dans les mentalités. Or, laisser quelque chose (de solide, pérenne) à la postérité est précisément un moteur important pour les chercheurs. Il se trouve que ce qu'on laisse de durable, c'est le plus souvent une idée qui portera son nom, et non pas le ou les articles qui les ont rendues publiques, mais bon l'association est là.

- un système qui deviendrait trop « social » par les réseaux informatiques serait à mon avis très instable à terme, il n'y a qu'à voir ce qu'est devenu UseNet (au départ, c'étaient uniquement des communautés de chercheurs). Les seuls forums à peu près lisibles et où il y a de l'information (bon rapport « Signal/Bruit ») sont ceux qui sont modérés. La boucle est bouclée ?

Il faut peut-être trouver quelque chose d'intermédiaire entre le tout papier et le tout réseau, et entre le tout sélection et le pas de sélection du tout.

Samuel Landau

*Si vous avez envie d'intervenir sur ce débat, de donner votre opinion ou de nous faire partager votre point de vue, n'hésitez pas à envoyer une contribution (remarques, articles, commentaires) à Jean-Daniel.Zucker@lip6.fr*



### **N'oubliez pas les petites annonces de l'AFIA !**

*L'AFIA et son bulletin offrent un nouveau service de petites annonces !*

Vous pourrez y soumettre ou consulter des annonces concernant l'intelligence artificielle, des offres ou demandes de stages, d'emploi, de logiciels...

Pour consulter les annonces actuelles, visitez le site de l'AFIA à l'adresse :

*<http://www.afia.polytechnique.fr/bulletin/PetitesAnnonces/PtAnnonces.html>*

Pour soumettre une annonce, envoyez un mél à Brigitte Grau (*Brigitte.Grau@limsi.fr*) qui gère cette rubrique (le champ « objet » (ou « sujet » ou « subject ») de ce mél devra mentionner explicitement « Petite annonce AFIA »)

## Nouvelles du Québec

Nicole Tourigny

Dans ce numéro, vous trouverez une description du laboratoire RALI du département d'informatique et de recherche opérationnelle (<http://www.iro.umontreal.ca/presentation.html>) de l'Université de Montréal (<http://www.umontreal.ca/infogen.html>) et d'une thèse de doctorat récemment soutenue.

En 1878, l'Université de Montréal fut inaugurée comme succursale de l'Université Laval de Québec (<http://www.ulaval.ca>), avec quatre facultés: théologie, droit, médecine et arts. L'Université de Montréal a reçu sa première charte du Parlement de Québec en 1920. L'Université de Montréal compte maintenant treize facultés, plus d'une soixantaine de départements et deux Écoles affiliées, soit l'École Polytechnique et l'École des Hautes Études Commerciales. L'ensemble accueille plus de 47 000 étudiants et comprend près de 2200 professeurs.

L'intelligence artificielle est un axe important de recherche au Département d'informatique et de recherche opérationnelle de l'Université de Montréal, et les activités du laboratoire RALI se situent dans cet axe.

\*\*\*\*\*

---

### Laboratoire RALI (Recherche Appliquée en Linguistique Informatique)

Responsable : Elliott Macklovitch

Professeurs : Philippe Langlais, Guy Lapalme, Jian-Yun Nie

Contact : [macklovi@iro.umontreal.ca](mailto:macklovi@iro.umontreal.ca) - <http://www-rali.iro.umontreal.ca>

---

Le RALI est affilié au Département d'informatique et de recherche opérationnelle de l'Université de Montréal. Il réunit des informaticiens et des linguistes d'expérience dans le traitement automatique de la langue tant par des méthodes symboliques « classiques » que par de nouvelles méthodes probabilistes. Le RALI a déjà publié une description de ses activités dans le bulletin de l'AFIA. Ceci n'est donc qu'une mise à jour sur certains de nos projets. Pour plus d'informations sur toutes nos activités, visitez notre site Web.

- *TSrali.com : le concordancier bilingue*

En juin 2001, le RALI a lancé *TSrali.com*, une version étendue de son concordancier interactif *TransSearch*, qui offre aux internautes une nouvelle base de données juridiques, une mise à jour complète du Hansard (1986 à 2001) et un serveur haute-vitesse dédié. *TSrali* est un service commercial, mais les abonnements sont disponibles à un coût très abordable. Le système connaît déjà un grand succès et le RALI prévoit ajouter des bases de données dans d'autres paires de langues.

- *TransType2 : l'aide à la frappe de traductions*

L'objectif du projet *TransType* est la mise au point d'un nouveau type d'outil interactif à l'intention des traducteurs humains. Le système accélère la saisie du texte traduit en proposant au traducteur des complétions des unités qu'il est en train de taper. Pour le faire, le système exploite deux modèles de langue probabilistes : un modèle de traduction, qui génère des prédictions de tra-

ductions de la phrase courante en langue source, et un modèle de la langue cible, qui privilégie les suites de mots les plus probables, telles qu'observées dans un corpus d'entraînement.

La première phase du projet, qui s'est terminée en novembre 2000, a donné des résultats très encourageants. D'ailleurs, les travaux sur *TransType* vont se poursuivre, car la Commission européenne vient d'approuver *TransType2*, un nouveau projet de trois ans, qui ajoutera à l'ancien prototype deux nouvelles langues cibles (l'espagnol et l'allemand), des complétions d'unités plus longues et une interface vocale. Les partenaires du RALI dans ce projet seront le laboratoire RWTH à l'Université d'Aachen, l'Instituto Tecnologico de Informatica à l'Université de Valence, le centre de recherche européen de Xerox à Grenoble, et la filiale espagnole de la société Schlumberger-Sema.

- *RIPTIDES : l'extraction d'information translinguistique*

Le RALI est le seul laboratoire à l'extérieur des États-Unis à participer au programme TIDES, qui est financé par DARPA, l'agence de recherche de l'Armée américaine. Le but de notre projet est de développer des outils permettant la construction rapide de systèmes d'extraction d'information, systèmes qui fourniront en plus des résumés en anglais de l'information repérée en langue étrangère. Les partenaires du RALI dans le projet *RIPTIDES* sont l'Université Cornell et la société GoGenTex. La contribution de notre laboratoire porte surtout sur les lexiques et les grammaires bilingues.

- *La réponse automatique au courrier électronique*

Ce projet vise le développement d'un système automatique de réponse aux courriers électroniques dans le cadre d'un service d'aide à la clientèle. Il est financé par le CRSNG et par les Laboratoires universitaires Bell. D'ailleurs, le domaine d'application sélectionné porte

sur les courriels reçus par une filiale de la société Bell. L'approche préconisée combine des modules de question-réponse, de classification de textes et de raisonnement à partir de cas. Le projet est dirigé par le professeur Guy Lapalme et mobilise au moins quatre étudiants à la maîtrise ou au doctorat.

---

## **Modélisation d'un système d'aide pour apprendre à résumer dans un environnement de raisonnement à partir de cas**

Auteur de la thèse : Laurence Capus

Date de soutenance : 26 novembre 2001

Rapporteurs : Dr Eugène Chouraqui, DIAM, Université Aix-Marseille III

Dr Richard Gagnon, Département d'études sur l'enseignement et l'apprentissage, Université Laval

Dr Jacques Ladouceur, Département de langues, linguistique et traduction, Université Laval

Directrice de thèse : Dr Nicole Tourigny, Département d'informatique, Université Laval

Laboratoire : ERICAE, Département d'informatique, Université Laval <http://www.ift.ulaval.ca/~ericae>

---

**RÉSUMÉ.** Cette thèse s'inscrit dans le domaine de l'ingénierie des connaissances dans une tâche de construction automatisée de résumés de textes au moyen d'exemples, dans un environnement de raisonnement à partir de cas. À notre connaissance, aucun résumeur automatique n'utilise cette approche. De plus, les méthodes employées dans les résumeurs actuels sont généralement non explicites, ce qui rend leur utilisation moins pertinente pour un environnement éducatif. Aujourd'hui, l'information est de plus en plus disponible et savoir la résumer est une alternative pour gagner du temps. Bien que la rédaction de résumés fasse partie de l'enseignement général, c'est une tâche difficile à accomplir car elle fait appel à des qualités d'abstraction, des facultés logiques et des capacités à s'exprimer. Le résumé d'un document a en effet pour but d'attirer l'attention du lecteur intéressé par le sujet et l'inciter à lire le document dans sa totalité. On peut apprendre à résumer en consultant des exemples de rédaction de résumés pour comprendre comment ils ont été obtenus. Cette façon d'apprendre, à partir d'exemples, est commune est naturelle. Toutefois, les exemples ne doivent pas être pris au hasard et seront plus utiles s'ils sont expliqués. L'objectif de cette thèse est donc de modéliser un système d'aide pour

apprendre à résumer au moyen d'exemples. La méthode de développement KADS / CommonKADS est employée afin de modéliser le système au niveau des connaissances. Un modèle des exemples est ajouté à la méthode afin de structurer les connaissances qu'ils contiennent. Le paradigme du raisonnement à partir de cas, qui permet de résoudre des problèmes en réutilisant des expériences similaires, est exploité pour la mise en œuvre du système envisagé. À partir de cette modélisation, le prototype GARUCAS a été implémenté. Il possède une base de douze exemples, construits à l'aide d'un modèle cognitif de la construction de résumés de textes narratifs. Cette implémentation a permis de montrer la faisabilité de la modélisation proposée. GARUCAS est capable de retrouver des récits similaires, c'est-à-dire ayant la même structure événementielle, et de montrer comment ils ont été résumés. Il est également capable de réutiliser une méthode de résumé par similarité. En conséquence, un nouveau modèle de système d'aide pour apprendre à résumer, exploitant le raisonnement à partir de cas, est proposé. Cette expérimentation permet de poser les bases d'une méthode pour construire des systèmes d'aide pour apprendre au moyen d'exemples.

\*\*\*\*\*

Je profite de l'occasion pour vous souhaiter une bonne et heureuse année. Je rappelle qu'il me fait toujours plaisir de répondre à vos questions et vos commentaires sont grandement appréciés.

Nicole Tourigny, Ph.D., professeur  
Département d'informatique  
Pavillon Adrien-Pouliot

Université Laval  
Québec, CANADA  
G1K 7P4

[tourigny@ift.ulaval.ca](mailto:tourigny@ift.ulaval.ca)  
<http://www.ift.ulaval.ca/~ericae/>

# PRÉSENTATION DE LABORATOIRES

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

LIFIA, Grenoble .....	Bulletin n°1	LIPN, Université de Paris Nord .....	Bulletin n°14
LRI, Orsay .....	Bulletin n°1	Institut EURISCO .....	Bulletin n°15
Service Systèmes Experts, Renault.....	Bulletin n°1	LRDC, Université de Pittsburgh (USA).....	Bulletin n°15
CEDIAG,.....	Bulletin n°2	Société ISOFT .....	Bulletin n°16
CERT, ONERA, Toulouse.....	Bulletin n°2	Dépt. d'Info de l'Université d'Ottawa .....	Bulletin n°16
IRIT, Toulouse.....	Bulletin n°2	Equipe CHM, Université du Colorado (USA) .	Bulletin n°17
LAAS, Toulouse .....	Bulletin n°2	LIRMM, Montpellier .....	Bulletin n°19
HEUDIASYC, UTC.....	Bulletin n°3	Institut autrichien de recherches en I.A. ....	Bulletin n°20
IFP, Rueil Malmaison.....	Bulletin n°3	ENST Bretagne .....	Bulletin n°21
DIAM, INSERM U194 .....	Bulletin n°3	LIA - Université de Savoie.....	Bulletin n°22
Lab. Math. Info., Fac Médecine de Marseille..	Bulletin n°4	INRETS .....	Bulletin n°23
GMD, St. Augustin (RFA) .....	Bulletin n°4	IRIN Nantes.....	Bulletin n°24
ONERA, Chatillon .....	Bulletin n°4	CRIN - INRIA Lorraine.....	Bulletin n°25
KSL, Université de Stanford (USA) .....	Bulletin n°5	DIRO - Université de Montréal.....	Bulletin n°26
Dépt Applications de l'IA au CNET, Lannion.	Bulletin n°5	IRIT - Toulouse (1) .....	Bulletin n°28
LAFORIA, Univ. Pierre et Marie Curie.....	Bulletin n°6	IRIT - Toulouse (2) .....	Bulletin n°29
L'institut FAW, ULM (RFA) .....	Bulletin n°6	LAAS - Toulouse (1).....	Bulletin n°30
Institut IIIA, Compiègne .....	Bulletin n°6	Sony CSL .....	Bulletin n°31
LAIR, OHIO State University (USA).....	Bulletin n°7	LAAS - Toulouse (2).....	Bulletin n°32
ARAMIHS, Labo mixte MATRA-CNRS,.....	Bulletin n°7	LIMSI - Département CHM.....	Bulletin n°33
CEA, Service SERMA, Saclay .....	Bulletin n°8	LAMSADE.....	Bulletin n°34
Société ILOG.....	Bulletin n°8	Institut autrichien de recherches en I.A. ....	Bulletin n°36
LAIAC, Université de Caen.....	Bulletin n°9	LIP6 (Univ. Pierre et Marie Curie).....	Bulletins n°37 & 38
Institut Français du Pétrole .....	Bulletin n°10	GREYC (Univ. de Caen).....	Bulletin n°40
DFKI (Centre allemand de recherches en IA) .	Bulletin n°11	LIFL (Univ. de Lille).....	Bulletin n°41
GRTC, Marseille .....	Bulletin n°11	LRI (équipes IA et IASI) .....	Bulletin n°43
Inst. d'Analyse des Systèmes, Ac. Russe .....	Bulletin n°12	IMAG - Grenoble .....	Bulletin n°44
Georges Mason Univ., Center for AI (USA)....	Bulletin n°13	PSI (Perception, Système, Information - Rouen).....	Bulletin n°45
IRISA, INRIA et Université de Rennes.....	Bulletin n°13	INRIA – Sophia Antipolis .....	Bulletin n°46/47
Société INGENIA .....	Bulletin n°14	LIH (Labo d'Informatique du Havre).....	Bulletin n°46/47

## Dossier « IA & médecine »

Coordonnateurs.

**Vincent Corruble** (LIP6 – Vincent.Corruble@lip6.fr)

**Jean Charlet** (STIM/DPA/DSI/AP-HP & Univ. P 6 – jc@biomath.jussieu.fr)

### ÉDITORIAL

Cela devient presque une tradition pour la communauté AFIA de faire le point périodiquement sur ses activités en IA et médecine. L'IA et la médecine, c'est un mariage qui ne partait pas sur un pied d'égalité. La première est une science (?) très jeune, manipulant des concepts et des techniques, visionnaire par définition, la tête dans les nuages par nécessité. La médecine est un art ancien, a un objectif - soigner les patients - noble mais terre à terre, et ne s'embarasse pas de concepts abstraits quand elle peut s'en passer. Opposées sur beaucoup de points, les deux disciplines se sont rapidement trouvées des complémentarités. La première était accusée de ne rien produire d'utile à la société, la deuxième était à l'affût de nouvelles techniques pour alimenter sa réflexion, améliorer l'efficacité de son propre fonctionnement. Il y a donc déjà près d'un quart de siècle que l'IA (au moins une partie significative de la communauté) s'est posée (certains diraient « casée ») en oubliant ses rêves juvéniles pour faire œuvre utile. Elle y a partiellement réussi en trouvant avec la médecine un terrain d'application rêvé pour le champ naissant des systèmes experts. Partiellement, car les enfants du couple n'ont suscité qu'un intérêt épisodique dans la famille médicale : celle-ci ne s'est que très rarement investie dans le long terme dans l'éducation de ces enfants, qui avaient besoin d'une attention très soutenue (du fait, paraît-il, de risques d'étranglement « au niveau du goulot »), pour des résultats le plus souvent limités dans la pratique médicale. Mais vaillè que vaillè, ce couple a tenu le coup. Et l'IA en éprouvait une fierté

non dissimulée, cette progéniture offrant (malgré ses problèmes) une certaine confirmation du bien-fondé de ses rêves d'adolescence.

L'évolution du couple s'inscrit aussi dans l'évolution d'un contexte, familial et social. La famille de l'IA a changé. Des techniques qui étaient encore balbutiantes il y a 20 ans ont maintenant atteint leur maturité (apprentissage, fouille de données, ingénierie des connaissances, ingénierie documentaire,...). Celles-ci, parfois encouragées par des cousines (statistiques, reconnaissance des formes, ...), donnent de la voix et souhaiteraient influencer à leur manière la direction du couple, et redonner ainsi une nouvelle vie à certains des rêves d'enfance de l'IA, longtemps mis de côté. L'apprentissage ou la fouille de données permettent ainsi d'automatiser tout ou partie de l'extraction de motifs ou de connaissances pour, par exemple, l'analyse d'image, l'aide au diagnostic ou, plus largement, l'aide à la décision. Ces techniques facilitent la coopération dans le couple car elles peuvent s'appliquer potentiellement dans le long terme de la pratique médicale, mais aussi (surtout ?) dans le court ou moyen terme de la recherche médicale, où elles peuvent apporter un soutien essentiel au médecin et sont donc sources de gratifications importantes pour des investissements moindres.

Pour la famille médicale aussi, le contexte a changé. L'art d'Hippocrate a beaucoup évolué. Il ne suffit plus de prendre en compte le bien-être des patients, il faut aussi intégrer des contraintes ou des objectifs de budget et maintenant de traçabilité. Cette évolution qui commence seulement à s'installer en France, mais qui est déjà prépondérante chez les cousins étrangers (en

particulier les anglo-saxons), est susceptible de donner un nouveau rôle à l'IA, car qui parle de budget, d'économie n'est pas loin des problématiques de rationalisation chères à cette discipline depuis la tendre enfance. Cette évolution vers la rationalisation s'accompagne aussi de la volonté de partager de façon efficace le meilleur de la pratique médicale au niveau national voir international, à travers le réseau des hôpitaux et des généralistes. La méthode traditionnelle qui consiste à partager cette connaissance lors des études médicales n'est plus pleinement satisfaisante car elle ne permet pas de mettre à jour la pratique tout au long de la carrière de ces « acteurs de la santé ». Pour pallier ce déficit, l'élaboration et l'utilisation de telles connaissances médicales est théorisée et développée dans une démarche de *médecine fondée sur les données actuelles de la science ou médecine factuelle* (*Evidence-Based Medicine* (EBM) en anglais).

Il est intéressant de noter que de nombreux travaux tiennent compte des nouvelles conditions de diffusion de l'information et de la connaissance médicale à travers le Web et privilégient des modes de représentation textuels. La formalisation, incontournable quand on veut proposer l'aide de machines aux hommes, est alors appliquée à l'indexation de ces ressources ou comme représentation formelle des thésaurus médicaux. Les projets de représentation formelle des connaissances médicales comme GALEN ou SNOMED-CT commencent d'ailleurs à porter leurs fruits en fournissant le support de vocabulaires contrôlés pour l'expression des informations et des connaissances médicales. Même si les contraintes de leur mise en œuvre ralentissent leur acceptation par

la communauté et leur utilisation dans les systèmes d'information de santé, ces approches, *ontologiques*, sont très prometteuses et participent de la problématique de *l'interopérabilité sémantique* indispensable à la communication et les échanges en médecine.

Toutes ces évolutions, récentes, sont pleinement visibles dans ce dossier « IA et Médecine » où l'on voit effectivement apparaître depuis la précédente édition (1993) des travaux utilisant des approches complètement différentes, en particulier l'ingénierie documentaire, les ontologies (déjà citées), la fouille de données pour la découverte de connaissances, ou la simulation multi-agents. On voit alors quasiment disparaître des approches visant à modéliser ou simuler le raisonnement médical. On peut le regretter mais aussi constater que cette approche (pour laquelle les systèmes experts constituaient une voie fondatrice) se réincarne dans des projets visant à définir des normes et des protocoles plutôt qu'à décrire le raisonnement médical, débouchant ainsi sur des *guides de bonne pratique médicale* (GBP) dont l'informatisation suscite à la fois recherches et controverses. Ces guides indiquent le chemin à suivre pour que l'ensemble du corps médical adopte les meilleures pratiques possibles déjà citées ici, dans une démarche de médecine factuelle (ou EBM) et pour le mieux-être du patient.

Pour terminer ces considérations sur des réflexions d'évolution des domaines scientifiques, on peut constater que, de la même façon que l'IA s'immerge au sein des sciences et technologies de l'information et de la communication (STIC), l'IA appliquée à la médecine s'implique rapidement et fortement dans les STIC et les applique à la santé en développant sa dimension de discipline insérée dans un large contexte spatial et socio-professionnel. Les différents domaines composant les STIC sont alors mis à contribution en santé, que ce soit au niveau des protocoles de communication (Web, ebXML, SOAP), de la représentation des connaissances

(document semi-structurés (XML) *versus* ontologies), de l'indexation (*Web sémantique*), de l'interprétation d'images, de l'interprétation des bases de données (fouille de données), etc.

Dans tous les cas – heureusement pour les chercheurs ? –, il reste encore du travail pour donner au médecin les moyens de pratiquer une médecine à la fois *factuelle* et humaine en mettant à sa disposition les outils, interfaces, protocoles de communication, informations, connaissances, etc. en tous lieux (hôpital, clinique, cabinet, domicile). On peut donc dire que le couple IA et médecine, s'il garde sa capacité à s'adapter aux nouveaux besoins et à tirer profit de ses nouveaux potentiels, a encore de beaux jours devant lui. On peut donc lui prédire de nombreuses années de vie commune et beaucoup de recherches fructueuses, au bénéfice de celui qui doit rester au centre des préoccupations de tous, le patient.

- oOo -

#### Organisation du dossier

Les contributions présentées dans ce dossier résultent d'un appel à participation sous la forme de descriptions d'équipes ou, pour ceux qui le désiraient, de propositions de réflexions avancées ou projets. Dix-sept équipes nous ont envoyé des descriptions d'activité, trois des projets (les projets sont les encadrés dispersés au long des contributions). Une équipe ayant envoyé les deux types de contribution, ce sont finalement dix-neuf équipes qui nous ont soumis leurs travaux listés dans l'encadré de cette page. Nous tenons à les remercier car, par leur contribution, elles nous permettent de présenter un panorama représentatif, sinon exhaustif, des recherches en France à l'intersection de l'Intelligence Artificielle et de la Médecine.

#### Liste des équipes ayant envoyé des contributions au dossier

##### Descriptions d'activité

Équipe AFIRM (TIMC-Grenoble)  
Projet AIDA/DREAM (IRISA-Rennes)

Thème APA (LIP6-Paris)  
UPRES EA2672 (Marseille)  
UPRES EA2694 & Grappa (Lille)  
UPRES EA3192 (Rennes)  
Projet Kasimir (LORIA)  
Laboratoire LATIM (Brest)  
Laboratoire LERTIM (Marseille)  
Laboratoire LIM (Rennes)  
Laboratoire LIMBo (Bobigny)  
Laboratoire PSI (Rouen)  
Équipe SIC (TIMC-grenoble)  
Équipe SPIM (Broussais-Hôtel Dieu-Paris)  
Mission de recherche STIM (AP-HP-Paris)  
Unité U455 (INSERM-Toulouse)  
Unité U494 (INSERM-Paris)

#### Articles courts

Projet ICREM (LIP6)  
Équipe SeT  
(UT de Belfort-Montbéliard)  
Mission de recherche STIM (AP-HP)

\*\*\*\*\*

#### Équipe AFIRM : Acquisition, Fusion d'Information et Réseaux pour la Médecine (Laboratoire TIMC)

Laboratoire TIMC-IMAG UMR CNRS 5525.

Adresse : Faculté de Médecine de Grenoble, 38706 La Tronche Cedex.

Tél. 04 76 63 71 11.

Fax. 04 76 51 86 67.

Ad. él. : [Norbert.Noury@imag.fr](mailto:Norbert.Noury@imag.fr);

[Vincent.Rialle@imag.fr](mailto:Vincent.Rialle@imag.fr)

<http://www-timc.imag.fr/AFIRM>

**Composition :** Thierry CREUZET, Norbert NOURY (responsable), Vincent RIALLE. Doctorants : Gilles VIRONE, Lionel BAJOLLE, Florence DUCHE-NE.

#### Cadre général et mission :

Le domaine de recherche de l'équipe se situe au carrefour de deux mondes : celui de la mesure physique au moyen de capteurs de tout type d'une part, celui des sciences et technologies de l'information, de la communication et de la

décision d'autre part. Cette multidisciplinarité fondatrice, incarnée par les fondateurs de l'équipe (N. Noury et V. Rialle), est mise au service d'un troisième monde, celui de la santé et de la médecine, abordé essentiellement à travers sa dimension « télé » : télémédecine, télésurveillance médicalisée, aide au maintien ou au retour de patients sur leur lieu de vie habituel. Centré sur le patient et le médecin, l'objectif de recherche de l'équipe s'inscrit d'emblée dans la perspective actuelle d'évolution vers une société de l'information médicale. Les fondateurs de l'équipe sont sensibles aux questions éthiques et socio-économiques de cette évolution. Issue à son origine du mariage entre le capteur intelligent (acquisition et traitement du signal - traitement et fusion des données) et le système d'information médicale (fusion pour la décision - diffusion et communication), l'activité de recherche inclut dans sa "boucle" une étude approfondie des besoins des patients et des acteurs médicaux et sociaux concernés. Le cœur de cette activité de recherche est constitué par l'application des techniques mathématiques et d'intelligence artificielle à l'acquisition et à l'analyse des données (capteurs intelligents) dans une perspective d'aide à la décision médicale, de détection de situations critiques, de modélisation et d'évaluation fonctionnelle de l'activité du patient.

#### Bref historique des travaux :

Dans ce contexte, les membres de l'équipe contribuent depuis plusieurs années au sein du laboratoire TIMC au développement du concept d'Habitat Intelligent pour la Santé (HIS).

A cet effet, un logement complet a été réalisé et équipé au sein du CHU de Grenoble. Cette réalisation constitue un prototype d'appartement de type T1 (50 m<sup>2</sup> environ), comprenant les zones d'habitat classiques que sont la chambre, le séjour, la cuisine, les toilettes, la douche et un couloir. Une zone technique attenante à l'appartement a été ajoutée afin de recevoir le système informatique d'expérimentation du projet.

Un capteur intelligent de chute a été développé et breveté sous la direction de

N. Noury (n° 01/12046, UJF 28/09/01).

Parallèlement à ces réalisations « en dur », plusieurs logiciels d'analyse et de traitement de l'information ont été mis en place, en particulier : 1) une application en Labview de visualisation et d'analyse des données multi-capteurs avec modélisation de l'occupation des salles et analyse du rythme circadien du patient au domicile, 2) un système distribué multi-utilisateurs selon une architecture orientée-objet couvrant la chaîne complète de d'informations depuis leur perception (capteurs), à leur présentation au médecin (télémonitoring) et à leur analyse en termes décisionnels (fusion de données, détection d'alarmes), médicaux (diagnostic) et statistiques (actimétrie diurne/nocturne, etc.). Ce second logiciel constitue un prototype de système distribué sur Internet selon une architecture n-tiers. Il est fondé sur les techniques de représentation UML et les technologies de développement Java™ (RMI, Java Web Start, etc.). Une représentation en XML des diverses données est privilégiée dans tous les logiciels.

#### Type de problèmes traités :

- Les problèmes essentiels se situent :
- 1) au niveau de l'exploration des données et de l'identification d'événements saillants, tels que des incidents (défaut de prise de médicament, malaise, etc.) ou des urgences (chutes, crises) ;
  - 2) au niveau de l'association d'actions thérapeutiques ou de secours en réponse aux événements identifiés.

En termes d'IA les problèmes se situent au niveau de la représentation des données et des connaissances d'une part, de l'apprentissage et de raffinement des modèles de décision d'autre part. L'IA intervient enfin au niveau de l'apprentissage des formes récurrentes d'événements. Elle opère dans le cadre renouvelé de la « fusion de données » multicapteurs ou multi-sources. Elle fait appel ici à des algorithmes classiques de l'apprentissage supervisé et non supervisé, mais également des algorithmes bayésiens et markoviens. L'IA est également présente mais de manière indirecte, à travers la modélisation objet du

système d'information complet (incluant des aspects de gestion des utilisateurs, de configuration du réseau de capteurs, de configuration des logements, de distribution des traitements entre poste au domicile et poste de contrôle médical, etc.).

#### Panorama des travaux effectués et de projets en cours :

L'équipe AFIRM est actuellement impliquée dans cinq projets financés : TISSAD (MENRT) : Technologies de l'Information Intégrées aux Services de Soins A Domicile.

RESIDE-HIS (Projet IMAG) : Reconnaissance de Signaux de Détresse dans l'Habitat Intelligent Santé.

VTAMN (MENRT) : Vêtement de Télé-Assistance Médicale Nomade.

HIT (Consulat Général de France à Québec, MAE) : Habitat Intelligents pour la Télésanté - coopération franco-québécoise avec l'Université du Québec à Montréal.

EPICT (MENRT) : sélectionné en juin 2001, en cours de montage.

#### Partenaires industriels :

- Société ATMEL-France, Grenoble-site de St Egrève : communications sans fils.

- Société ATRAL, Grenoble - Crolles : systèmes d'alarmes pour la sécurité des biens et des personnes.

- Société MOULINEX, Caen : pèse-personne.

- Société NAÏS, filiale de Matsushita, Japon : tensiomètre.

- Société Mondial Lifeguard Technologies, Montréal, Canada : système d'alarme sociale.

- TAM-Télésanté, Aix-en-Provence, France. Projets VTAMN et EPICT.

#### Partenaires universitaires :

- Consortium TISSAD, Financement MENRT, 1999-2001 : LORIA, Nancy, (laboratoire Lorrain de Recherche en Informatique et ses Applications) UMR 7503 ; TIMC-IMAG UMR 5525, Grenoble, (équipe  $\mu$ ISFV et AFIRM) ; INSERM U 121, Lyon,

(Recherche sur l'activité du cœur) ; INSERM U 518 et LAAS (Laboratoire d'Analyse et d'Architecture des Systèmes), Toulouse.

- Laboratoire CLIPS-IMAG, équipe GEOD : analyse de la parole et du son ; projet RESIDE-HIS, financement : Institut de Mathématiques Appliquées de Grenoble (IMAG), 2000-2003
- Université du Québec à Montréal, Département d'Informatique : base de données, aide à la décision ; projet HIT, financement Ministère des Affaires Étrangères / Consulat de France à Québec, 2001-2002.

#### Partenaires institutionnels :

- Grenoble-Alpes Métropole (« La Métro »), Communauté des Communes de l'Agglomération Grenobloise, Grenoble.
- Centre Communal d'Action Sociale (CCAS), Grenoble.
- Hôpital A. Michallon, Grenoble.
- Association Grenobloise de Lutte contre l'Hypertension Artérielle (AGLHTA). 2001-2003 (Projet EPICT).

#### Bibliographie :

##### 2001

Noury N., Rialle V., Virone G. (2001). The telemedicine home care station: a model and some technical hints. In Patricelli F. & Ray P.K. (éd.) : « Proceedings of HealthCom 2001 : 3rd International Workshop on Enterprise Networking and Computing in Health Care Industry », SSGRR (www.ssgrr.it). 37-40.

Rialle V., Noury N., Hervé T. (2001). An experimental Health Smart Home and its distributed Internet-based Information and Communication System : first steps of a research project. In V. Patel et al. (eds): « MEDINFO 2001, 10th World Congress on Medical Informatics ». Amsterdam: IOPress. 1479-1483.

Rialle V., Noury N., Fayn J., Chan M., Campo E., Bajolle L., Thomesse J.P. (2001). Health smart home information systems : concepts and illustrations. In Patricelli F. & Ray P.K. (éd.) : "Proceedings of HealthCom 2001 : 3rd International Workshop on Enterprise Networking and Computing in Health Care Industry", SSGRR (www.ssgrr.it). 99-103.

Thomesse J.-P., D. Bellot, A. Boyer, E. Campo, M. Chan, F. Charpillat, J. Fayn, C.

Leschi, N. Noury, V. Rialle, L. Romary, P. Rubel, N. Selmaoui, F. Steeneste, G. Virone (2001). Integrated Information Technologies for patients remote follow-up and homecare. In Patricelli F. & Ray P.K. (éd.) : « Proceedings of HealthCom 2001 : 3rd International Workshop on Enterprise Networking and Computing in Health Care Industry », SSGRR (www.ssgrr.it). 3-15.

Castelli E., Istrate D., Rialle V., Noury N. (2001). Information extraction from speech in stress situation : Application to the medical supervision in a smart house. In : C. Cavé, I. Guaitella et S. Santi (eds.), Oralité et gestualité : interactions et comportements multimodaux dans la communication (Actes du colloque ORAGE 2001). Edition l'Harmattan, 362-371.

#### 1999-2000 (avant création de l'équipe)

Rialle V., N. Lauvernay, A. Franco, J.-F. Piquard, P. Couturier (1999), A smart room for hospitalised elderly people : essay of modelling and first steps of an experiment. *Technology and Health Care*, 7(5), 343-57.

Rialle V., Lauvernay N., Noury N. (2000). Besoins et réponses technologiques en télé-gérontologie : un foisonnement de possibilités et quelques interrogations. In A. Franco (éd.) *Télé-médecine en gérontologie*. Serdi Edition 2000, 81-104.

Rialle V., Noury N., Hervé T. (2000). Le concept d'habitat intelligent et son application à la télé-médecine : Etat de l'art et méthodologie de conception. Actes du Congrès Mondial de Télé-médecine « Société globale de l'information », Toulouse - Institut Européen de Télé-médecine -, 22-24 mars. 139.

Hervé T., Noury N., Rialle V. (2000). Microtechnologie ambulatoire en télé-médecine. Actes du Congrès Mondial de Télé-médecine « Société globale de l'information », Toulouse - Institut Européen de Télé-médecine -, 22-24 mars. 125.

Noury N., Rialle V., Hervé T. (2000). Besoins en réseaux de communications pour la télé-médecine : application en soins à domicile. Actes du Congrès Mondial de Télé-médecine « Société globale de l'information », Toulouse - Institut Européen de Télé-médecine -, 22-24 mars. 136.

Noury N., Hervé T., Rialle V., Virone G., Cingala C., Gouze E., Mercier E. (2000). Wireless ambulatory acquisition of high resolution physiological signals. Proc. ETC 2000 : European Telemetry Conference. Garmisch Partenkirchen, Germany, 22-25 mai. Ed. German Society of Telemetering. 199-204.

Noury N., Hervé T., Rialle V., Virone G., Mercier E. (2000). Monitoring behavior in home using a smart fall sensor and position sensors. IEEE-EMBS « Microtechnologies in Medicine & Biology », Lyon-France, Oct 2000, 607-610.

Noury N. and P. Pilichowski, « A telematic system tool for home health care ». Proc. 14th Annual Int. Conf. IEEE-EMBS, Paris, France, 1992. Vol 3/7. 1175-1177.

**Projet AÏDA/DREAM  
(IRISA-INRIA)  
IRISA-INRIA  
Campus de Beaulieu,  
35042 Rennes Cedex**

#### Le projet AÏDA/DREAM :

Les chercheurs du projet AÏDA/DREAM de l'IRISA travaillent sur la modélisation, la surveillance et le diagnostic de systèmes évolutifs, dans le domaine des réseaux de télécommunications, de l'environnement et de la santé.

Il rassemble une dizaine de personnes dont trois travaillent plus particulièrement sur des applications médicales, en particulier sur l'application cardiaque décrite ci-dessous.

#### Contacts :

René Quiniou, Marie-Odile Cordier  
<http://www.irisafraida/>

#### Partenariats

Ces travaux sont menés en collaboration avec le LTSI (Laboratoire de Traitement de Signal) et un cardiologue du CHU de Rennes, ainsi qu'en liaison avec ELA Médical, industriel spécialisé dans le développement et la commercialisation de prothèses cardiaques.

#### Contacts :

Guy Carrault, Feng Wang, LTSI, Équipe Mixte INSERM, Campus de Beaulieu, 35042 Rennes  
Philippe Mabo, Département de Cardiologie, CHRU Rennes, 35000 Rennes  
Marcel Limousin, ELA Recherche Médical Synthé Labo, Plessis Robinson

**Thème de recherche : COOPERATION ENTRE TRAITEMENT DU SIGNAL ET RECONNAISSANCE DE CHRONIQUES POUR LA DÉTECTION D'ARYTHMIES CARDIAQUES**

Les unités de soins intensifs pour coronariens ont conduit au début des années 90 au concept de monitoring intelligent reposant sur une association de méthodes de traitement du signal, de modèles électrophysiologiques et de bases de connaissances. Cependant, très peu de travaux ont réellement réussi cette intégration. Ce travail s'inscrit dans la cette problématique et propose une nouvelle approche de la reconnaissance des arythmies cardiaques. Succinctement, elle met en jeu deux étapes : l'une, hors ligne, a pour objet la constitution de descriptions de haut niveau caractérisant les arythmies cardiaques et l'autre, en ligne, vise à reconnaître l'une des caractérisations répertoriées dans le signal ECG préalablement codé sous forme symbolique.

L'étape hors ligne s'appuie sur des méthodes d'apprentissage automatique et vise à induire les caractéristiques clefs de chaque arythmie et ainsi générer une base de chroniques, chacune constituée d'un ensemble d'événements temporellement contraints. L'apprentissage est réalisé, soit à partir d'observations réelles (des ECG ou électrocardiogrammes étiquetés comme dans la base d'arythmies MIT-BIH), soit à partir de signaux générés par un modèle cardiaque paramétrable selon l'arythmie étudiée. Un échantillon du premier type a l'avantage de correspondre à des observations réelles mais son étendue est limitée tandis qu'un échantillon du second type peut être constitué de manière systématique afin de couvrir la variabilité des arythmies visées. Les chroniques apprises s'expriment sous forme de règles du premier ordre directement interprétables par un expert du domaine. L'apprentissage est paramétrable et permet des ensembles de chroniques lisibles par les experts ou minimales ou encore robustes pour des reconnaissances efficaces.

L'étape en ligne débute par le traitement du signal d'entrée observé sur une ou plusieurs dérivations de l'ECG. Les algorithmes de traitement du signal procèdent à la détection des événements de la révolution cardiaque par une approche multi-source multi-capteur et réalisent la détection des différents types d'ondes suivie d'une étape de classification par réseaux de neurones.

Les battements sont ensuite caractérisés par un vecteur de paramètres issu d'une transformée en ondelettes et discriminés en deux ou plusieurs classes. L'ensemble est ensuite remis en forme pour construire une séquence d'événements symboliques qui constitue l'entrée du reconnaiseur de chroniques.

La reconnaissance de chroniques est une technique qui a prouvé son efficacité dans le diagnostic et la surveillance de systèmes industriels. Elle consiste à comparer, à la volée, l'évolution du processus sous surveillance (décrit par un flux d'événements détectés sur les signaux) avec un ensemble de comportements prédéterminés représentés par les chroniques. L'observation d'un nouvel événement peut provoquer, l'hypothèse d'une nouvelle chronique, la reconnaissance effective d'une chronique, l'infirmité d'une chronique ou, simplement, l'ignorance de cet événement.

Les premiers résultats sont particulièrement intéressants et démontrent la faisabilité de l'approche. Outre son applicabilité en cardiologie, la démarche proposée peut être étendue à d'autres champs du domaine biomédical dès lors que des événements structurés apparaissent sur les signaux.

#### Financement :

Une partie de ces recherches se fait dans le cadre d'une ACI (action concertée incitative) du Ministère de la Recherche, en partenariat avec le LTSI, l'industriel ELA Médical et des cardiologues du CHU de Rennes. Le projet financé, nommé PISE, étudie le développement de prothèses cardiaques (pacemakers, défibrillateurs) qui au lieu d'être programmées une fois pour toutes sont capables de s'adapter au contexte en analysant le signal en ligne et adaptant l'action correctrice.

#### Références :

Carrault G., Cordier M.O., Quiniou R., Garreau M., Bellanger J.J., Bardou A. : A model based approach for learning to identify cardiac arrhythmias, AIMDM'99, LNAI (Lecture Note in Artificial Intelligence), vol 1620, pages 165-174, Springer Verlag, 1999.  
Quiniou R., Carrault G., Cordier M.O., Wang

F. : Application of ILP to cardiac arrhythmia characterization for chronicle recognition, in C. Rouveiroi and M. Sebag editors, ILP'2001, LNAI, vol 2157, pages 193-205, Springer-Verlag, 2001.

G. Carrault, F. Wang, R. Quiniou, M.-O. Cordier : Apprentissage de séquences structurées : exemple en ECG, 18e colloque GRETSI'01 sur le traitement du signal et l'image, Toulouse, France, 2001.

### Thème APA Laboratoire d'Informatique de Paris 6 (Pôle IA)

#### Responsable du thème :

Bernadette Bouchon-Meunier

(Bernadette.Bouchon-Meunier@lip6.fr)

Contact IA & Médecine : Vincent

Corruble (Vincent.Corruble@lip6.fr)

UPMC – CNRS

8, rue du Capitaine Scott

75015 Paris

La médecine est un domaine qui nécessite la manipulation de connaissances complexes et étendues, ainsi que de nombreuses données (cliniques, biologiques, expérimentales,...). Il est donc naturel que le thème APA (Apprentissage et Acquisition des connaissances) du LIP6 s'y intéresse tout particulièrement depuis plusieurs années. Du grand nombre de projets en bio-médecine abordés par les chercheurs du thème APA, on peut dégager un certain nombre d'axes.

L'axe principal est celui de la fouille de données et de la découverte de connaissances. Dans cette problématique, on part des données bio-médicales réunies par des médecins, pour construire grâce à des algorithmes classiques ou dédiés de fouille de données (issus par exemple de l'apprentissage automatique) des modèles qui aident à la compréhension d'un phénomène. Cette approche diffère donc de l'approche classique des systèmes experts car le matériau de base est la base de données d'observation et non un expert du domaine. Cependant celui-ci garde un rôle critique, en particulier pour la formulation du problème et l'évaluation des résultats. Nous verrons ensuite un autre axe de recherche qui n'a pas pour

optique la découverte de nouvelles connaissances mais plutôt la conception de systèmes aidant la pratique médicale (par exemple le diagnostic) et la formation.

**Fouille de données et Découverte de Connaissances Bio-Médicales**

La problématique de la Fouille de données et de la découverte de connaissances médicales intéresse les chercheurs d'APA depuis plusieurs années et a conduit à des travaux dans le domaine de l'histoire de la médecine<sup>1</sup>, de l'évaluation des patients victimes de traumatismes crâniens<sup>2</sup>, ou de la prédiction de l'efficacité d'anti-dépresseurs<sup>3</sup>. Les projets plus récents mêlent médecine et biologie.

**Génétique fonctionnelle de l'obésité**

**Participants :** Jean-Daniel Zucker, Mélanie Courtine, Vincent Corruble

L'obésité est devenue la pathologie nutritionnelle la plus fréquente dans les pays industrialisés. Elle est maintenant considérée comme la pathologie d'un organe, le tissu adipeux, et de ses relations avec les autres structures impliquées dans la mise en réserve et l'utilisation d'énergie : le muscle, le foie et le système nerveux central [Clement et al. 1996 ; Clement 2000]. Nous collaborons avec la jeune équipe CNRS de Broussais-Hotel Dieu co-dirigée par Dr. Karine Clément sur un projet d'étude **génomique fonctionnelle** de l'obésité<sup>4</sup>. Cette étude se base notamment sur l'application de la technologie des puces ADN [DeRisi et al. 1997] dans la recherche des facteurs étiologiques de l'obésité.

1. Vincent Corruble, Jean-Gabriel Ganascia, Induction and the discovery of the causes of scurvy: a computational reconstruction, Artificial Intelligence (91)2 (1997) pp. 205-223.

2. A. McQuatt, P.J.D. Andrews, D. Sleeman, V. Corruble, P.A. Jones. The Analysis of Head Injury Data using Decision Tree Techniques. Actes du Congrès AIMDM'99. Springer Verlag, Lecture Notes in Artificial Intelligence 1620, Artificial Intelligence in Medicine.

3. A. Banto, J.G. Ganascia. Induction dans l'aide à la découverte médicale. Actes de JICAA'97.

4. Ce projet rassemble des compétences dans des domaines complémentaires : cliniques, biologiques et d'analyse d'information à travers un partenariat entre des enseignants chercheurs hospitaliers, des chercheurs de l'IFR056, d'unités INSERM, et implique la génopole parisienne (Montagne Sainte Geneviève).

Les progrès dans le séquençage du génome ouvrent aujourd'hui une nouvelle étape, celle de la génomique fonctionnelle. En effet, au-delà des analyses de profils d'expression des puces ADN, le problème de comprendre *ce qui caractérise* les gènes qui possèdent le même profil d'expression est essentiel. Dans le cadre de notre collaboration, nous travaillons sur trois types de bases de données : ARNOBMOUSE (données de puces ADN sur des souris), ARNOBHUMAN (données de puces ADN sur des humains) et ADNOBHUMAN (données génétiques et phénotypiques sur des humains qui ne proviennent pas de puces ADN mais d'autres protocoles).

Ces tâches de fouille de données illustrent des problèmes dans lesquels il faut manipuler tout à la fois des données *numériques* (des profils d'expression ou des modalités phénotypiques quantitatives) et des données *symboliques* (homo, hétéro, wild, up, down, données phénotypiques qualitatives, etc.). Or, il existe une franche séparation entre les systèmes qui permettent de faire du regroupement numérique et ceux qui traitent des données purement symboliques. Un problème encore peu traité est celui du traitement de données numériques en tant que telle dans les systèmes de *regroupement conceptuel*. Pour mener à bien ces différentes tâches d'analyse, nous avons développé plusieurs systèmes d'apprentissage autour d'une plateforme baptisée ObeLinks dont l'architecture est résumée dans la figure.

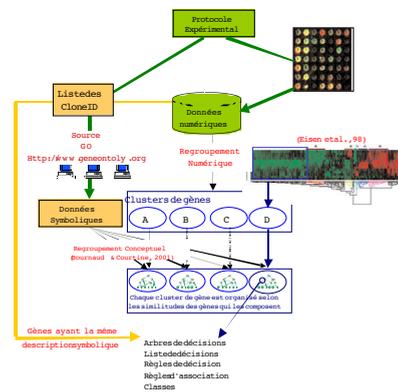


Figure : Schéma récapitulatif de la procédure d'analyse des données issues des puces ADN

**Epidémiologie de l'asthme.**

**Participant :** C. Marsala

L'analyse de données épidémiologiques peut se formuler comme un problème typique de la fouille de données où il s'agit d'apprendre à partir de données numérico-symboliques, éventuellement entachées d'imprécision et d'incertitudes. Différentes études ont été réalisées au sein du thème APA dans ce domaine. L'analyse des comportements de patients asthmatiques vis-à-vis du traitement qui leur est attribué a notamment été effectuée à l'aide d'arbres de décision flous dans le cadre d'une collaboration avec le docteur Lurie de l'hôpital franco-britannique. Ces travaux ont conduit à l'obtention d'un ensemble de règles lisibles qui caractérisent les différences entre profils de patients. Ce travail est actuellement soumis à publication.

**Prévision du risque cardio-vasculaire**

**Participants :**

F. d'Alché-Buc, C. Marsala.

Nous participons aussi au projet « Analyse des problèmes d'apprentissage posés par les données INDANA (INDividual DATA aNalysis of Antihypertensive intervention trials) pour la prévision du risque cardio-vasculaire ». Ce projet, initié par le département d'informatique médicale de l'hôpital Broussais (M.-C. Jaulent) et financé par l'Action concertée incitative « Technologies pour la Santé » du ministère de la Recherche, a débuté en octobre 2001. Il vise à l'élaboration, la comparaison et la combinaison de méthodes d'apprentissage automatique pour la prévision du risque cardio-vasculaire. La base de données disponibles, constituée et maintenue par l'Unité de Pharmacologie Clinique de l'Université Lyon 1 (Dr François Gueyffier), est constituée des résultats de 10 essais thérapeutiques conduits pour évaluer l'efficacité de traitements anti-hypertenseurs. Elle soulève des problèmes de distributions mal équilibrées et de valeurs manquantes que l'équipe étudie actuellement.

**Modélisation de réseaux géniques**

**Participant :** F. d'Alché-Buc.

La biologie moléculaire bénéficie

actuellement des importantes avancées réalisées dans les programmes de séquençage et dans le développement de techniques de criblage d'expressions géniques et protéiques à grande échelle. Parmi ces dernières, les biopuces à ADN ouvrent la voie à l'analyse de l'expression d'un grand nombre d'ARN d'un type cellulaire ou d'un tissu donné et donc à l'étude des interactions entre molécules. Il s'agit d'un progrès majeur en biologie puisqu'il permet d'envisager d'identifier plus précisément les sites sensibles à tel ou tel médicament.

Cependant, cette analyse qui se situe donc au niveau des réseaux moléculaires requiert le développement de nouveaux outils de modélisation de ces systèmes complexes et dynamiques. L'apprentissage automatique offre un cadre théorique et méthodologique pour aborder cette tâche que l'on peut traduire par l'analyse de données complexes, l'extraction de la structure de ces réseaux moléculaires et la modélisation de leur dynamique. Nos activités s'inscrivent dans le cadre d'une collaboration avec le laboratoire L.G.N. (Laboratoire de la Génétique de la Neuro-transmission moléculaire) de l'Université Paris 6, dirigé par le professeur Jacques Mallet. Nos travaux communs portent sur la modélisation dynamique des réseaux géniques à partir de données issues de biopuces à ADN et de connaissances a priori<sup>5</sup>.

#### Aide au diagnostic et formation.

A coté de travaux sur la fouille de données et la découverte de connaissances en médecine, nous nous intéressons aussi au développement de techniques permettant de partager des connaissances, facilitant ainsi la pratique et la formation médicale.

#### Partage de connaissances sur le diagnostic de malformations

**Participants :** M. Rifqi, A. Strauss (UPMC), R. Traphoener (Empolis), M. Hoffet (CHU Montpellier), JP Schaaps (CHU Liège)

L'objectif des projets MUFNET (français) et FM-ULTRANET (euro-

péen) est d'offrir un outil dans le champ du diagnostic prénatal permettant la mise en commun de l'expertise et la mise à disposition des connaissances, en utilisant les nouvelles technologies de communication en particulier Internet et le DVD.

Les malformations fœtales diagnostiquées avant naissance représentent un problème fréquemment évoqué dans un centre pluridisciplinaire de diagnostic prénatal. Le diagnostic précis d'une malformation fœtale est difficile et nécessite une expertise et une expérience importante. Le but est de développer une plate-forme de formation en ligne pour les échographistes pour la surveillance d'une grossesse. Les objectifs techniques sont :

appliquer les techniques CBR aux séquences échographiques pour former les échographistes au diagnostic fœtal  
utiliser un DVD et un site web comme support de validation  
développer deux centres d'expérimentation (Montpellier et Liège).

L'objectif est donc de concevoir un environnement pour l'échographiste afin de :

- construire un cas à partir d'une séquence échographique
- guider l'analyse de la conclusion fournie par le logiciel
- obtenir des recommandations de la part du logiciel.

#### Résultats attendus :

- une base contenant des cas cliniques de référence concernant les malformations fœtales
- des outils aidant l'échographiste à reconnaître son cas d'étude dans la base de référence et à le guider dans son diagnostic
- des outils aidant l'échographiste à analyser l'histoire du cas qu'il étudie et à l'archiver dans sa base personnelle
- un moyen de communication avec des experts extérieurs pour analyser des cas complexes et pour compléter ou mettre à jour la base de cas de référence avec ses propres cas.

#### EA2672 – Projet THOT & logique floue

##### Contact :

Pr. M. Roux

<Michel.Roux@medecine.univ-mrs.fr>

##### Principaux personnels de l'EA2672 impliqués dans ce projet :

Pr. M. Roux (PU-PH, Informatique médicale) ; Pr. A. Bourgeade (PU-PH, Médecine tropicale) ; Dr H. Chaudet (MCU-PH, Informatique médicale) ; Dr B. Giusiano (MCU-PH, Informatique médicale) ; Dr S. Gentile (AHU, Informatique médicale) Mme L. Pellegrin (Post - doctorant, psychologie cognitive), Dr J. Godart (Post internat).

##### THOT : VOLET TALN

##### Le Cadre du projet

Nous proposons des éléments de solution au problème de l'extraction des connaissances décrites dans certains types de comptes rendus médicaux et épidémiologiques. L'objectif est de rendre immédiatement utilisable le sens de ces textes, afin que la connaissance ainsi extraite soit utilisée pour construire un modèle général de suivi des épidémies.

Ce modèle, couplé au système EDISAN [1], a pour but de maintenir à jour, en permanence, l'état des épidémies dans tous les pays du monde.

Or les informations relatives à ces épidémies se trouvent dans les dépêches épidémiologiques envoyées par des observateurs sur le terrain, aux grands organismes internationaux de santé.

Ces dépêches constituent notre « fond documentaire » à analyser.

##### LA RECHERCHE :

C'est l'extraction des connaissances contenues dans les dépêches épidémiologiques.

Classiquement, trois approches sont possibles pour comprendre un texte :

- 1 - Procéder à une analyse syntaxique du texte, phrase après phrase, puis à une analyse sémantique.
- 2 - Procéder à une analyse sémantique directe, sans passer par l'étape syntaxique.
- 3 - Combiner l'analyse syntaxique et sémantique en gardant comme objec-

5. Modélisation dynamique de réseaux géniques par apprentissage automatique, Pierre-Jean Lahaye, Rapport de DEA, DEA IARFA, Université Pierre et Marie Curie, Paris.

tif principal l'élaboration d'une représentation "interne" des objets décrits dans le texte, de leurs propriétés et de leurs relations avec les autres objets.

L'étude préliminaire, aboutissement du projet ARISTOTE [4], nous a montré l'importance majeure d'une intrication étroite entre syntaxe et sémantique. Cette intrication permet de restreindre dans une proportion très importante le domaine de recherche de la sémantique du texte par l'élimination précoce des solutions syntaxiques incorrectes sur le plan sémantique.

Le formalisme des DCG (Definite Clause Grammar) que nous avons adopté, est issu des travaux en Prolog [2] de Pereira et Warren.[3]

L'aboutissement du projet ARISTOTE, nous a montré que l'analyse doit comporter trois modules:

**a. Le lecteur :** qui lit la chaîne de caractères constituant le texte, identifie les syntagmes grâce à une analyse morphologique minimum (genre et nombre), élimine les formes stylistiques qui n'apporte pas d'information, et rend une suite de phrases. Chaque phrase est constituée d'une suite d'unités lexicales correspondant au texte et identifiées grâce au lexique.

Ce dernier comporte pour chaque unité lexicale le signifiant, le signifié syntaxique (et ses propriétés en genre et nombre), le signifié sémantique.

**b. L'analyseur syntactico-sémantique** qui traite le texte phrase à phrase. Chaque syntagme est traité et ses propriétés ou les liens qu'il peut avoir avec les syntagmes voisins (selon l'ordre de la grammaire DCG) sont immédiatement vérifiés. La base de connaissance est interrogée. Sa réponse indique la présence ou l'absence de solution sémantique pour la solution syntaxique en cours. Cet analyseur fournit en sortie pour chaque phrase, la liste de concepts reconnus, les propriétés de chacun d'eux et les relations qui les relient.

**c. Le constructeur :** il reprend la forme de sortie de l'analyseur syntactico-sémantique et procède à l'élaboration d'un graphe (ou plusieurs) conceptuel correspondant à l'ensemble du texte. Il procède par unification. Dans le cadre du projet THOT, cette structure en graphes conceptuels sera remise en cause. En effet cette structure sera adap-

tée aux contraintes de la deuxième grande étape de synthèse de l'état des épidémies.

#### Collaborations

- Avec le LIM, site de Luminy - P. SABATIER et Robert PASERO - pour les problèmes de développement de la grammaire.

- Avec le CREPCO (Université de Provence, site d'Aix en Provence) - Pr. C. BASTIEN - pour l'aspect cognitif.

#### Projet THOT : Volet EDISAN

Ce projet est dominé par une problématique majeure : comment structurer et organiser une documentation hypertexte pour qu'elle corresponde au mieux au mode de travail de l'utilisateur (quelle information, à quel moment du processus de prise de décision). Cette problématique a été abordée en faisant appel conjointement aux méthodes de l'intelligence artificielle, aux systèmes documentaires et à l'ergonomie cognitive.

Le point commun des systèmes hypertextes et hypermédias est la quasi-totale orientation de leur modèle sur des problèmes techniques et une optimisation informatique, marginalisant la prise en compte de la sémantique du domaine d'application. L'existence de l'utilisateur n'y est envisagée que pour définir l'existence d'un dialogue homme - machine et pour souligner les difficultés cognitives associées à l'utilisation de la navigation dans l'hypertexte. Peu de modèles, et seuls les plus récents, proposent une prise en compte de cet aspect, en se limitant toutefois à la sémantique des liens.

Compte tenu de cette situation, et partant du principe que tout comportement humain était motivé par la satisfaction d'un objectif déterminé, nous avons tenté de dégager une organisation des aspects cognitifs liés à l'utilisation d'un hypertexte de façon à mettre en évidence la place de cette motivation. Il est en effet possible de considérer qu'il existe une tâche maîtresse, liée à l'existence d'un objectif général, qui donne un cadre à l'utilisation du système en influençant profondément toutes les autres composantes du modèle, qui sont la tâche de navigation, la connaissance de l'utilisateur, et la connaissance du domaine. Cette dernière composante,

matérialisée par l'hypertexte lui-même, n'est alors utilisée que dans la mesure où elle concerne le problème en cours de traitement. Or, il se trouve que le processus de navigation, qui représente la seule interface entre l'humain et la connaissance - cible, est obligé de suivre les liens établis lors de la réalisation de l'hypertexte pour permettre à l'utilisateur de procéder à l'extraction de l'information qui lui est nécessaire. Le pari pour que, d'une façon générale, l'accès à l'information recherchée soit le plus immédiat possible est très faible si l'hypertexte n'a pas été conçu spécifiquement pour le problème envisagé. Ce choix d'une conception spécifique de l'hypertexte pour une utilisation précise avait déjà été envisagé par Jonassen dans le cadre de l'apprentissage (et les résultats avaient été peu probants, à notre avis du fait d'une énorme difficulté à construire un modèle précis de l'activité qui permette d'obtenir une décomposition fiable des sous-objectifs). Le domaine d'activité que nous avons choisi pour étudier notre hypothèse est la prise de décision dans le cadre d'une hyper - spécialité médicale. Le domaine de la prise de décision est relativement bien connu, a fait l'objet de nombreuses modélisations tant dans le cadre de l'intelligence artificielle qu'en psychologie, et ses objectifs sont clairs, précis. Le choix associé d'une spécialité très restreinte permet de plus de cerner les limites du domaine - cible de connaissance.

La méthode qui a été utilisée se compose de quatre phases :

- Une modélisation de l'activité liée à la tâche - maîtresse à l'aide de l'échelle double de Rasmussen révisée par J.M. Hoc, modèle « rustique » mais qui s'est révélé être bien adapté au besoin.
- Une décomposition des demandes d'accès à l'information sur la base des contextes de problèmes, donnant naissance à ce que nous avons appelé des « palettes de requêtes ».
- L'analyse de l'organisation de la connaissance - cible sur la base de ces contextes.
- La structuration de l'hypertexte à partir de cette organisation lors de sa réalisation, cette structuration concernant aussi bien le contenu des documents que les liens du réseau.

Le système résultant est utilisé sur plus de 130 sites, et une première étude montrerait son adéquation au public - cible initialement prévu.

#### Références

- [1] H. Chaudet, A. Bourgeade, C. Drouvillé, V. Feuillie, Apport du multimédia dans un système documentaire de référence en médecine des voyages (EDISAN). Proceeding of the european conference on tropical medicine. October 1995, (Hambourg, Germany, 1995) 22-26
- [2] F. Giannesini, H. Kanoui, R. Pasero, M. Van Caneghem, Prolog, (InterEdition, Paris, 1985)
- [3] F. Pereira, D.H.D. Warren, Definite Clause Grammars for natural language analysis : a survey of the formalism and a comparison with Augmented Transition Networks, Artificial Intelligence 13 (1980) 231-278.
- [4] M. Roux, V. Ledoray, Understanding of medico-technical reports, Artificial intelligence in medicine 2000 18 149-172.

#### Projet logique floue et applications

**Contact :** Elie SANCHEZ,  
*elie.sanchez@medecine.univ-mrs.fr*

Dans le cadre de la logique floue et Internet, plusieurs projets sont en cours d'étude et développement (collaborations en cours) comme (1) évaluateur de sites Web (personnalité, émotions, couleurs, formes), faisant intervenir la logique floue, (2) télérobotique mobile ("e-motion"), applets floues et contrôle à logique floue (3) construction d'une ontologie floue de termes en relation avec le contexte de requêtes, dans un moteur de recherche et (4) apport de la logique floue au Web Sémantique.

Par rapport à la médecine, l'apport de la logique floue à l'aide à la décision médicale est étudiée : Il s'agit de modélisation et de traitement des connaissances et des données imprécises ou mal définies, à partir de concepts et techniques à base de règles, en logique floue. Les applications portent sur l'élaboration de logiciels d'aide au diagnostic en médecine et biologie, au diagnostic de pannes en ingénierie, à l'aide au bilan et à l'évaluation dans l'orientation professionnelle (coll. DEMETER, Centre de pré-orientation pour personnes handicapées, Marseille).

### Communication avancée et santé publique - LILLE 2 (EA2694) et Grappa - Apprentissage automatique - LIFL & LILLE3

#### Équipe communication avancée et santé publique, EA 2694

**adresse :** CERIM, Centre Etudes et Recherches en Informatique Médicale faculté de médecine Henri Warembourg, pôle recherche, Université de Lille 2 59045 LILLE CEDEX

**web :** <http://www.univ-lille2.fr/cerim/>

**directeur :** Pr Régis BEUSCART,

#### contacts :

Alain DUHAMEL,

*mel : duhamel@univ-lille2.fr*

Nathalie SOUF,

*mel : nsouf@univ-lille2.fr*

**tel :** 03 20 62 69 69

#### Les principaux thèmes de recherche sont :

le développement d'outils innovants issus des Sciences et Technologies de l'Information et de la Communication (STIC) appliqués au domaine de la Santé,

l'analyse des facteurs humains contribuant à l'acceptation de ces technologies (ergonomie cognitive, analyse d'activité, étude d'impact)

les méthodologies d'analyse et d'exploitation des informations recueillies pour une meilleure connaissance de la santé des populations.

#### Thème 1 : STIC et santé

L'équipe travaille sur la mise au point d'outils permettant de mieux coordonner et distribuer des soins. Dans ce cet objectif, le premier axe de recherche concerne la modélisation de l'activité en utilisant des modèles de Workflow adaptatifs et flexibles. Nous travaillons sur la possibilité d'optimiser les stratégies de coordination des processus de soins, d'intégrer des modèles d'activité et d'améliorer la continuité des soins. Le deuxième axe de recherche concerne la prise en compte de données réparties. Nous travaillons sur des modèles de

communication utilisant les concepts de l'intermédiation. Ce travail a débouché sur une application industrielle (réseau ville - hôpital) qui optimise la communication entre les structures hospitalières et les médecins de ville. Cette méthodologie est appliquée actuellement à la problématique de la prise en charge à domicile (projet RNTS en cours).

#### Thème 2 : ergonomie cognitive

L'objectif de cette recherche est de mieux appréhender les facteurs humains concourant au succès ou à l'échec des applications de l'informatique en milieu de soins. Elle est menée en collaboration avec l'équipe PERCOTEC (CNRS Valenciennes). Au plan fondamental, la recherche concerne l'analyse de l'activité des utilisateurs, c'est-à-dire tout ce qui est mis en œuvre par l'utilisateur aux plans mental et comportemental pour accomplir sa tâche. Ces recherches ont pour objectifs de proposer une aide à la conception des prototypes informatiques et l'évaluation de « l'utilisabilité » de l'outil.

#### Thème 3 : fouille de données et entrepôts de données

Actuellement, dans les établissements de santé, grâce aux systèmes d'information hospitaliers et aux multiples recueils de données mis en place dans les différentes disciplines, de grandes quantités de données médicales sont accumulées mais ne sont ni exploitées, ni valorisées. Dans ce contexte, nos travaux concernent deux directions complémentaires :

- Les méthodes d'extraction de connaissances à partir de données (KDD) : nous travaillons en particulier sur la prise en compte des données manquantes tant au niveau de l'alimentation des bases de données décisionnelles (data warehouse) qu'au niveau des algorithmes d'extraction de connaissances (data mining). Ce travail s'effectue en collaboration avec l'équipe Grappa décrite ci-après. L'objectif de ces recherches est de concevoir des outils adaptés aux don-

nées médicales pour la construction d'une plate-forme « fouille de données pour le domaine médical ».

- Le développement de systèmes d'aide à la décision : l'extraction de connaissances à partir de bases de données doit aider à la constitution de systèmes d'information utiles pour les décideurs. Nos travaux portent sur l'apport des techniques de Data Warehouse (DW) et des agents intelligents dans ce contexte : intégration de données distribuées, supervision du processus de data mining, reengineering du DW.

## **Collaborations scientifiques :**

Au niveau régional: CHRU, Cercle E-santé, Eris@, participation au projet TACT du contrat de plan état région Nord-Pas-de-Calais.

Au niveau national : ACI MRT 99 (Projet TIPHAD) ; ACI MRT 2000 (Projet DATADIAB) ; RNTS MRT 2000 (Projet UMVF, projet COQUAS)

Au niveau international: IMIA WG18, Europe, EHTEL

## **Transfert technologique :**

Deux entreprises ont été créées à partir du laboratoire, avec l'aide de l'Anvar et d'Eurasanté: Atalante (1999) et Planet-Healthcare (2000).

## **Equipe Grappa**

Groupe de Recherche sur l'APPrentissage Automatique

**adresse :** équipe Grappa  
domaine universitaire Pont de Bois,  
Université Charles de Gaulle - Lille 3  
59053 LILLE CEDEX

**web :** <http://www.grappa.univ-lille3.fr>

**responsable :** Pr Rémi GILLERON,  
**mel :** [gilleron@univ-lille3.fr](mailto:gilleron@univ-lille3.fr)

**tel :** 03 20 41 61 78

L'équipe Grappa (Groupe de Recherche en Apprentissage Automatique) est une équipe du LIFL (Laboratoire d'Informatique Fondamentale de Lille - Lille 1, UPRESA 8022 CNRS) de l'Université de Lille 3. Ses deux thèmes de recherche sont :

- apprentissage naturel et artificiel de langages naturels et artificiels
- apprentissage à partir de données hétérogènes et « Data Mining »

Dans le second thème, nous étudions l'apprentissage supervisé avec peu d'exemples étiquetés, la combinaison de classifieurs (techniques de boosting en particulier), la fouille de données, la fouille de textes et la combinaison d'algorithmes pour données et textes. Dans le cadre du projet DATADIAB qui concerne le diabète, nous avons développé des algorithmes de prédiction adaptés aux problèmes multi-étiquettes (plusieurs complications possibles pour un même patient). Dans le projet INDANA, nous appliquons les techniques de boosting à un problème de prédiction du risque cardio-vasculaire.

## **Collaborations scientifiques :**

Au niveau régional: participation au projet TACT du contrat de plan état région Nord-Pas-de-Calais.

Au niveau national : ACI MRT 2000 (Projet DATADIAB) ; ACI MRT 2001 et INSERM (Projet INDANA)

## **Intégration de données multimédia en anatomie et physiologie cérébrale pour l'aide à la décision et l'enseignement (EA3192)**

Identification de l'équipe UPRES EA 3192 « Intégration de données multimédia en anatomie et physiologie cérébrale pour l'aide à la décision et l'enseignement »,

**Directeur :** Bernard Gibaud, Chargé de recherche Inserm

Faculté de Médecine, Université de Rennes I, 2 avenue du Pr Léon Bernard, 35043 Rennes Cedex

Téléphone : 02 23 23 45 90

Fax : : 02 23 23 45 86

**mel :** [bernard.gibaud@univ-rennes1.fr](mailto:bernard.gibaud@univ-rennes1.fr)

**site web :** <http://idm.univ-rennes1.fr>

Composition de l'équipe :

3 PU-PH, 1 MCU-PH

1 CR Inserm, 1 MCU, 1 IR

2 PH, 3 ASS-CC, 1PHU, 5 thésards scientifiques

## **Domaine et objectifs généraux de la recherche**

La recherche vise à concevoir et à évaluer des outils pour optimiser le diagnostic, la planification et la réalisation de gestes chirurgicaux en neurochirurgie. Ces améliorations sont censées résulter :

- de la prise en compte d'une imagerie multimodalité (anatomique et fonctionnelle), grâce à des outils de recalage et de comparaison qualitative et quantitative ;
- de la prise en compte de connaissances et de savoir faire a priori (présents localement ou accessibles à distance) sous la forme de modèles (modèles anatomiques ou fonctionnels, modèles de gestes chirurgicaux, etc.) ou de cas antérieurs ;
- de la mise en oeuvre d'une étape de simulation pré-opératoire et de l'optimisation des modes de visualisation et d'interaction avec les données pré-opératoires lors de la réalisation du geste (neuro-navigation et réalité augmentée).

Les principales caractéristiques de la recherche sont les suivantes :

- domaine de l'informatique et de l'imagerie médicale
- recherche méthodologique, dirigée vers la conception, l'intégration et l'évaluation d'outils (logiciels) ; collaboration avec d'autres centres (notamment INRIA/IRISA sur le site rennais)
- recherche appliquée, en liaison étroite avec les utilisateurs médecins (Services de Radiologie, Neurologie, Neurochirurgie, Explorations Fonctionnelles du CHR de Rennes) ; implantation à proximité immédiate du site hospitalier  
collaboration industrielle « en amont » (boursiers CIFRE) et « en aval ».

## **Thèmes de recherche**

Au plan thématique la recherche s'organise autour de deux thèmes. :

- « Fusion coopérative de données anatomiques, physiologiques et positionnelles pour la définition et la réalisation

tion de gestes chirurgicaux ». Il s'agit d'utiliser conjointement différentes modalités d'imagerie anatomique (CT, IRM, angiographie 3D) et fonctionnelle (TEMP, IRMf, MEG, TEP) pour élaborer le diagnostic, déterminer l'approche chirurgicale optimale, et assister la réalisation du geste opératoire lui-même en établissant des liens entre les données pré-opératoires et les données per-opératoires.

- « Partage de données et de connaissances dans un contexte de télé-médecine (aide à la décision, recherche, enseignement) ». Cette recherche s'inscrit dans une double évolution de la pratique médicale, consistant à fonder les décisions médicales sur des connaissances a priori explicites, locales ou distantes (connaissances générales du domaine, cas antérieurs). Notre contribution porte principalement sur la représentation de ces données et de ces connaissances (au niveau sémantique et syntaxique) pour permettre leur partage et leur ré-utilisation dans le cadre d'applications distribuées d'aide à la décision, de recherche et d'enseignement.

### Travaux passés touchant au domaine de l'intelligence artificielle

Depuis le début des années 90 nous nous sommes intéressés à la représentation de connaissances anatomiques et fonctionnelles sur le cerveau. Ces travaux ont été menés en collaboration avec Serge Garlatti du LIASC (Laboratoire Intelligence Artificielle et Systèmes Cognitifs) à l'École Nationale des Télécom Bretagne. La problématique concernait plus spécifiquement la représentation de connaissances relevant de différents domaines (points de vue) et la conception d'hypermédia adaptatifs capables de prendre en compte les domaines et tâches spécifiques de l'utilisateur [Montabard 93, Garlatti 94, Garlatti 96].

Ces travaux ont permis - entre autres - de conceptualiser une démarche de conception d'atlas évolutifs, pour des applications principalement orientées vers l'aide à la décision [Gibaud 98].

### Travaux actuels touchant au domaine de l'intelligence artificielle

#### 1. Modèles de gestes chirurgicaux

Nous cherchons à développer des outils pour aider le chirurgien dans la préparation de l'acte chirurgical. Ceci inclut la sélection des données image pertinentes pour la réalisation de l'acte. Il s'agit en outre de leur associer une sémantique spécifiquement « chirurgicale » (e.g. zones à atteindre, zones à éviter, zones de repère, trajectoires) qui peut dépendre des différentes phases du « scénario chirurgical » (inciser la peau, réaliser le volet osseux, disséquer le cortex etc.)

Ce travail met en jeu la définition d'une ontologie du domaine, mettant en évidence les structures d'informations communes aux différents gestes chirurgicaux rencontrés [Jannin 2001]. Ceci s'avère indispensable pour garantir une sémantique commune entre les utilisateurs (neurochirurgiens), les outils de simulation et les outils de guidage per-opératoire. Au delà de cette première étape, l'analyse approfondie du déroulement des principaux types d'interventions doit faire émerger des modèles de gestes, rassemblant et formalisant les connaissances relatives à ceux-ci. L'utilisation de ces modèles devrait permettre de faciliter la préparation des interventions, de mieux les protocoliser et donc de faciliter leur apprentissage.

On s'intéresse enfin à la création de documents comme des documents de planning multimédia contenant les données relatives au geste chirurgical, sous la forme de données structurées, d'images ou de scènes 3D décrivant les différentes étapes de l'intervention (XML, VRML). Au-delà de leur utilisation première pour guider l'opération et alimenter les outils de guidage par l'image, ils serviront de support pour générer les compte-rendus opératoires. Ces deux types de documents permettront d'alimenter une base de cas cliniques, pouvant servir à la recherche ou à l'enseignement.

#### Position par rapport à d'autres travaux

La modélisation des gestes chirurgicaux a souvent été envisagée sous un point de vue classificatoire (orientation

terminologique) : modèles du CEN TC251, travaux de GALEN. Nous nous inspirons de ces travaux et les complétons pour satisfaire nos besoins « opérationnels » (gestion des données issues des images pré-opératoires).

#### 2. Représentation et partage de connaissances générales et instanciées sur le cortex cérébral

La problématique de ces travaux concerne d'une part la conception d'outils d'aide à la décision dans le domaine de l'imagerie cérébrale anatomique et fonctionnelle pour des applications chirurgicales, et d'autre part la documentation médicale, en fournissant des référentiels sémantiques utilisables dans les documents médicaux. En effet, la représentation formelle et explicite des connaissances constitue un enjeu important pour leur ré-utilisation.

Nos travaux portent à la fois sur des représentations numériques (c'est à dire extraites ou dérivées des images anatomiques et fonctionnelles) et sur des représentations symboliques. Les premières caractérisent les positions, les formes et la variabilité des structures anatomiques cérébrales (cartes de probabilité, modèles statistiques de variations de forme, etc.) Les secondes constituent des représentations conceptuelles des entités du domaine. Elles soulèvent notamment le problème de la cohérence des différentes relations taxonomiques, méréologiques (relation « partie-tout » et topologiques qui lient les structures anatomiques ou fonctionnelles entr'elles (collaboration avec Anita Burgun, équipe PROTIMES Pierre Le Beux, Rennes).

#### Position par rapport à d'autres travaux

Nous envisageons actuellement d'utiliser des outils comme OIL ou Protégé. Nous étudions la compatibilité de nos modèles avec ceux de Digital Anatomist (Université de Washington à Seattle) et GALEN, les deux principaux modèles anatomiques généraux (non spécifiques du cerveau).

#### 3. Fédération de bases d'information en neuroimagerie

Enfin nous participons à un projet national sur la « Gestion de données et

de connaissances réparties en neuro-imagerie » (Projet NEUROBASE, ACI du MENRT). Il vise à étudier comment des bases d'informations de neuro-imagerie conçues indépendamment les unes des autres peuvent être fédérées, de façon à permettre à différentes applications de rechercher des données au travers de requêtes complexes (utilisation de technologies de type médiateurs/adaptateurs).

Ce projet est mené par l'INRIA/IRISA Rennes, et nous associe à 2 unités INSERM (U494 à Paris et U438 à Grenoble), à 1 équipe CNRS (TIMC CNRS UMR 5525 à Grenoble), à 3 unités INRIA (INRIA/IRISA projets Vista et Temics à Rennes, INRIA projet Caravel à Rocquencourt, INRIA projet Epidaure à Sophia-Antipolis) et à l'IFR 49 (qui inclut notamment le CEA SHFJ à Orsay).

*Position par rapport à d'autres travaux*

Human Brain Mapping Project, NIH, USA

### Références bibliographiques

Montabord E., Gibaud B., C. Barillot, S. Garlatti, I. Kanellos, B.S. Wu, A. Biraben, X. Morandi. : An Hypermedia System to Manage Anatomical Knowledge about Brain. in: Lemke HU, Inamura K., Jaffe CC, Felix R. Eds., *Computer Assisted Radiology*, Springer-Verlag Berlin Heidelberg, 1993, pp.414-419.

Garlatti S., Kanellos I., Montabord E., Barillot C., Gibaud B. : YAKA: un Système d'Information Intensionnel sur le Cerveau Humain, Proc. *Xème Congrès AFCET-AFIA INFORSID '94*, Aix en Provence, 18-20 Mai 1994, pp.323-338

Garlatti S., Montabord E., Gibaud B., Barillot C. : Base de connaissances intensionnelle avec perspectives et héritage multiple, *Congrès RFIA 96*, Rennes, 1996, pp 495-503.

Gibaud B., Garlatti S., Barillot C., Faure E. : Computerised brain atlases as decision support systems: a methodological approach. *Artificial Intelligence in Medicine* 14:83-100 (1998).

Jannin P., Raimbault M., Morandi X., and Gibaud B. Modeling surgical procedures for multimodal image-guided neurosurgery. In MICCAI 2001, Utrecht (The Netherlands), *Lecture Notes in Computer Science*, W.J. Niessen and M.A. Viergever (eds) Vol. 2208, pages 565-572, 2001.

Dameron O., Gibaud B., and Morandi X. Numeric and symbolic knowledge represen-

tation on cortex anatomy using web technology. In S. Quaglini, P. Barahona, and S. Andreassen, editors, *Artificial Intelligence in Medicine (AIME 01)*, Cascais (Portugal), *Lecture Notes in Artificial Intelligence*, volume 2101, pages 359-368, Berlin, Heidelberg, 2001. Springer.

### Le projet Kasimir (Loria)

#### Rédacteur :

Jean Lieber, lieber@loria.fr,  
tél. 03 83 59 20 86,  
équipe Orpailleur, Loria BP 239, 54506  
Vandœuvre-lès-Nancy

Le projet Kasimir réunit des chercheurs en IA de l'équipe Orpailleur du Loria (*Vandœuvre-lès-Nancy*, <http://www.loria.fr/equipes/Orpailleur/>), des experts du centre régional de lutte contre le cancer Alexis Vautrin (CAV, Vandœuvre-lès-Nancy) et du réseau lorrain de soins en cancérologie (Oncolor, Vandœuvre-lès-Nancy), et des chercheurs du laboratoire d'ergonomie du CNAM (Paris). Il s'inscrit dans le cadre de la gestion des connaissances pour le traitement des cancers (en particulier, du sein et de la prostate).

Le traitement des cancers au CAV et dans la région Lorraine s'appuie sur des référentiels (similaires à des protocoles de décision ou à des guides de bonnes pratiques cliniques) qui peuvent être considérés comme des bases de connaissances informelles. Un patient atteint d'un cancer est confronté au référentiel qui lui correspond. Le référentiel fournit en général une solution thérapeutique. Si ce n'est pas le cas (cf. incomplétude de la base) ou si la solution est critiquable (contre-indications, etc.), le cas est dit hors référentiel et passe devant un CCP (comité de concertation pluridisciplinaire) constitué d'experts chargés de proposer une solution thérapeutique. Il y a un référentiel et un CCP pour les cancers du sein et de la prostate et, à terme, pour chaque localisation cancéreuse.

Les recherches en ergonomie décrites dans [1] constituent les premiers pas du projet Kasimir. Elles montrent en particulier comment les réunions du CCP facilitent le partage de connaissances

expertes et aboutissent à la création de nouvelles connaissances lors de la prise de décision. Ce deuxième point a été étudié dans [2] sous l'angle de l'apprentissage à partir d'échecs : le cas hors référentiel conduit à un échec de l'application littérale du référentiel qui doit être adapté pour ce cas particulier en tenant compte des explications de cet échec. Cela constitue alors une occasion d'enrichissement du référentiel qui peut être soumise aux experts.

Pour mettre en œuvre ce travail, il était nécessaire de représenter informatiquement le référentiel. Ce travail a conduit au système Kasimir/référentiel [3, 4] qui permet un raisonnement par classification dans une hiérarchie représentant le référentiel du cancer du sein au stage locorégional, avec une représentation des connaissances par objets. Une version pour le diagnostic et le traitement du cancer de la prostate a également été développée. Kasimir/référentiel est en cours de validation et est destiné à être utilisé par des spécialistes lorrains délocalisés qui sont désireux d'appliquer les règles de bonnes pratiques.

Le système Kasimir/hors référentiel est en cours de conception [5]. Il s'inscrit dans le cadre du raisonnement à partir de cas et permettra de faire des adaptations du référentiel pour des cas hors référentiel. L'acquisition et la modélisation des connaissances d'adaptation qu'il nécessite sont décrites dans [6]. Une première version de Kasimir/hors référentiel s'appuyant sur la classification hiérarchique floue est en cours de finition et permet de prendre en compte l'effet de seuil.

À terme, l'intégration de ces deux systèmes dans l'organisation doit être étudiée pour répondre au mieux à un triple objectif : (a) fournir une assistance à l'aide à la décision thérapeutique, (b) favoriser l'enseignement mutuel entre experts et (c) proposer des évolutions des référentiels.

### Références

- [1] Sauvagnac (C.). - La construction de connaissances par l'utilisation et la conception de procédures. Contribution au cadre théorique des activités métafonctionnelles. - Thèse d'Université, Conservatoire National des Arts et Métiers, 2000.
- [2] Boussu (G.). - Apprentissage à partir

d'échecs. - Mémoire de DEA d'informatique, Université Henri Poincaré Nancy 1, 1998.

[3] Bresson (B.) et Lieber (J.). - Classification pour l'aide au traitement du cancer du sein. In : Septième journées de la Société Francophone de Classification, SFC'99, éd. par Le Ber (F.), Mari (J.-F.), Napoli (A.) et Simon (A.). pp. 53-59. - Nancy, septembre 1999.

[4] Bresson (B.). - Un système de classification hiérarchique pour l'aide au traitement du cancer du sein. - Mémoire de thèse CNAM, Conservatoire National des Arts et Métiers, février 2000.

[5] Lieber (J.) et Bresson (B.). - Case-Based Reasoning for Breast Cancer Treatment Decision Helping. In : Advances in Case-Based Reasoning - Proceedings of the fifth European Workshop on Case-Based Reasoning (EWCBR-2k), éd. par Blanzieri (E.) et Portinale (L.), pp. 173-185. - Springer, 2000.

[6] Lieber (J.), Bey (P.), Boisson (F.), Bresson (B.), Falzon (P.), Lesur (A.), Napoli (A.), Rios (M.) et Sauvagnac (C.). - Acquisition et modélisation de connaissances d'adaptation, une étude pour le traitement du cancer du sein. In : Actes des journées ingénierie des connaissances (IC-2001), pp. 409-426. - Grenoble, 2001.

### Raisonnement médical et aide à la décision en endoscopie digestive LATIM (Brest)

J.M. Cauvin, C. Le Guillou, B. Solaiman, M. Robaszkievicz, C. Roux. Laboratoire de Traitement de l'Information Médicale, INSERM ERM 01-02 (LATIM).

Centre Hospitalier Universitaire, 29609 Brest.

*jean-michel.cauvin@chu-brest.fr*

Inventée dans les années 50, la fibroscopie a largement contribué à l'évolution des connaissances en gastroentérologie, et radicalement modifié les modalités du diagnostic et du traitement des pathologies digestives, parce qu'elle autorisait, au moyen d'un geste peu invasif, l'observation directe des lésions et des traitements non chirurgicaux. Apparue vers la fin des années 80, la vidéoendoscopie a amélioré la qualité de l'image, le confort de l'opérateur, et permet de conserver des séquences vidéo qui peuvent être numérisées. Ces évolutions technologiques débouchent

sur la volonté d'intégrer l'image endoscopique dans le système d'information médicale. Exploiter ces masses de données impose de concevoir de nouveaux outils qui classent et retrouvent l'information.

L'endoscopie a plutôt été enseignée par compagnonnage et les ouvrages d'endoscopie digestive sont plus orientés vers la technique (manipulation des appareils, endoscopie interventionnelle...) que vers la sémiologie endoscopique. Ce n'est que récemment que des travaux ont abouti à un recensement et une définition de termes utilisables dans le domaine. Dans tous les cas ces terminologies se situent à un niveau d'agrégation de l'information supérieur au niveau de description élémentaire de l'image.

L'objet de ce projet est la conception d'un système d'aide à la décision en endoscopie digestive. Devant une situation clinique ambiguë pour l'opérateur, comment retrouver dans une base iconographique les situations les plus similaires à la scène observée ? Comment renforcer sa conviction ? Il ne s'agit pas ici de mettre en place des mécanismes automatiques de description de l'image, mais de proposer au médecin endoscopique une autre approche structurée de description de l'endoscopie.

Entre la réalité des éléments observables au cours d'une endoscopie digestive et la formulation d'hypothèses diagnostiques, il était nécessaire d'identifier les différentes étapes qui conditionnent la qualité des conclusions de l'observation. A chaque étape du diagnostic, le raisonnement de l'endoscopiste met en relation des informations issues de l'observation et des connaissances médicales acquises par l'enseignement et/ou l'expérience. La confrontation du contenu de deux terminologies spécialisées, de l'expérience d'endoscopistes et des méthodes utilisées en reconnaissance de forme a permis de concevoir un modèle de raisonnement en deux étapes : le diagnostic d'observation endoscopique qui s'appuie sur l'analyse de l'image et le diagnostic endoscopique, étape finale intégrant le contexte médical de l'examen. Ce modèle a été assimilé à une représentation scène-objet. Les indices propres à chaque étape et leurs relations avec la décision ont été identifiés et rassemblés dans une

base de connaissance comportant 96 classes d'observations et 150 classes de diagnostics endoscopiques.

La méthode de recherche des classes est basée sur une métrique de similarité prenant en compte la nature symbolique de l'information manipulée. Plusieurs mesures de similarité ont été testées sur échantillons et obtiennent des résultats proches : la classe d'observation attendue est la réponse unique ou première plus de 9 fois sur 10 (1).

Des expériences menées avec des endoscopistes confirment la validité du modèle de raisonnement, révèlent les difficultés à définir les relations indice-décision (1). Dans une mise en situation proche de la réalité, la méthode retrouve la classe d'observations attendue dans 41 à 96% des requêtes. Ces résultats encourageants guident les travaux en cours pour améliorer la base de connaissance et la méthode de recherche.

La première réponse attendue de ce système d'aide à la décision est le diagnostic de l'observation dans le contexte du diagnostic endoscopique. Le raisonnement par classification, mis en oeuvre jusqu'à présent, peut être amélioré par la complétude et la définition plus précise des indices, mais aussi par la prise en compte dans la méthode de recherche des relations étroites entre diagnostic d'objet et diagnostic de scène (2).

La seconde réponse attendue du système d'aide à la décision est de proposer à l'utilisateur les images les plus similaires à la situation observée. Pour répondre à cet objectif, il est nécessaire d'identifier parmi l'ensemble des indices recensés, ceux qui permettent à l'endoscopiste d'associer un cas observé avec une expérience antérieure. Cette recherche, en cours de réalisation, constituera la première étape de la mise en oeuvre d'un raisonnement à base de cas.

Le Laboratoire de Traitement de l'Information Médicale poursuit par ailleurs des travaux sur l'indexation de l'image par son contenu numérique (1). Ces deux thèmes de recherche sont traditionnellement menés d'une manière indépendante par deux communautés scientifiques, celle de la compression d'images (du domaine signal/image) et celle des bases de données (du domaine de l'informatique et de l'intelligence

artificielle). A terme, la mise en cohérence des deux approches symbolique plus proche du monde médical, et numérique plus proche du monde de l'ingénierie doit favoriser l'émergence d'outils métier adaptés au domaine de la compression et de l'indexation de l'image médicale.

### Références

1. *Diagnostic reasoning by classification in upper digestive tract endoscopy*  
Cauvin JM, Le Guillou C, Solaiman B, Robaszkiewicz M, Gouerou H, Roux C.  
World congress on medical physics and biomedical engineering. Chicago, juillet 2000.
2. *Diagnostic reasoning model validation in digestive endoscopy.*  
Cauvin JM, Le Guillou C, Solaiman B, Robaszkiewicz M, Gouerou H, Roux C.  
IEEE-EMBS 23rd Annual Conference. Istanbul, 25-28 octobre 2001.
3. *Upper digestive endoscopic scene analyze.*  
Le Guillou C, Cauvin JM, Solaiman B, Robaszkiewicz M, Gouerou H, Roux C.  
IEEE-EMBS 23rd Annual Conference. Istanbul, 25-28 octobre 2001.
4. *Medical image indexing and compression based on vector quantization: image retrieval efficiency evaluation.*  
Ordonez Jr, Cazuguel G, Puentes J, Solaiman B, Cauvin JM, Roux C.  
IEEE-EMBS 23rd Annual Conference. Istanbul, 25-28 octobre 2001.

### Laboratoire d'Enseignement et de Recherche en Traitement de l'Information Médicale (LERTIM)

Faculté de Médecine, Université de la Méditerranée.  
27, bld Jean Moulin,  
13385 Marseille Cédex 5.  
Tél : 04 91 38 70 84  
**Contact :** mfieschi@ap-hm.fr  
**Web :** <http://Cybertim.timone.univ-mrs.fr/Cybertim>

Sous la responsabilité du Professeur Marius FIESCHI, le Laboratoire d'Enseignement et de Recherche en Traitement de l'Information Médicale (LERTIM) a entrepris depuis plusieurs années des recherches visant à élaborer des méthodes et à développer des outils

permettant un couplage entre connaissance médicale et information sur le patient afin d'améliorer la prise en charge du patient. La prise en charge des connaissances concerne la documentation, la bibliographie, l'aide à la décision diagnostique et thérapeutique, les références médicales, les guides de bonne pratique, la formation et le transfert de connaissances. La gestion du patient inclut la gestion de données cliniques, de soins, de recherche, ainsi que son dossier d'évaluation et de management. Le couplage de ces deux aspects – connaissance médicale et information du patient – nécessite la spécification et le développement d'outils de couplage et d'interopérabilité. L'objectif de cette recherche exige, pour déboucher sur des outils utilisables en pratique, une approche intégratrice de différents domaines de recherche classiques (description des concepts médicaux, méthodes d'intelligence artificielle et de psychologie cognitive, élaboration de modèles de raisonnement, modèles cognitifs d'interaction homme-machine) et de développement de composants logiciels de présentation, de traitement et de communication des informations et des connaissances ainsi que la mise en oeuvre des technologies du multimédia. Cette présentation porte sur les aspects liés à l'interfaçage homme-machine dans le domaine de la santé à propos de deux projets menés au sein du laboratoire. Le premier projet, dénommé ARIANE, s'intéresse à l'interopérabilité des différentes sources d'information et de connaissance mises à disposition des professionnels de la santé. Le second projet, Lernet, met en pratique les méthodes et techniques multimédia pour une aide adaptée à la formation. Tous deux mettent en pratique des principes de l'intelligence artificielle et utilisent des technologies de l'information et de communication (TIC).

**ARIANE** [1, 2]. Participent au projet : Michel JOUBERT, Dominique FIESCHI, Marius FIESCHI et des docteurs.

Les bases d'information du domaine biomédical sont des bases de données de patients, des bases de données bibliographiques, de la documentation tech-

nique, etc, dans lesquelles les concepts du domaine biomédical ne sont qu'incomplètement représentés. Un utilisateur est confronté à la difficulté de rapprocher sa propre perception des concepts de la représentation qui en est faite dans chaque base d'information, variable d'une base à l'autre. De plus, ces bases sont nombreuses et délivrées par des systèmes informatiques souvent différents, ce qui rend encore plus difficile leur utilisation. Le but d'ARIANE est de jouer le rôle de médiateur entre les utilisateurs et ces bases d'information. L'objectif premier d'ARIANE est de libérer les utilisateurs des efforts à faire pour interroger les bases d'information à leur disposition en leur proposant des moyens d'accès basés sur une ontologie du domaine qui représente les concepts et les relations entre ces concepts. ARIANE se comporte ainsi comme une interface conceptuelle entre utilisateurs et bases d'information. Pour atteindre totalement son objectif, ARIANE doit être capable de mettre en relation les utilisateurs avec les sources d'information de la manière la plus simple possible et la plus efficace pour eux. Ceci pourra se faire au moyen de requêtes construites automatiquement et dirigées vers les bases de données correspondantes ou bien par la connexion à des sites Intranet ou Internet répondant aux besoins en information des utilisateurs.

Pour atteindre son objectif premier qui est de faciliter l'accès à des bases d'information, ARIANE se base sur une ontologie des concepts, types et relations sémantiques du domaine biomédical : le *Unified Medical Language System* (UMLS), vaste projet de la *National Library of Medicine* des Etats-Unis. La base de connaissances de l'UMLS contient, entre autre, une définition de types de concepts et un (meta-)thesaurus. Aux types de concepts s'appliquent des relations sémantiques, formant ainsi un réseau sémantique. Le meta-thesaurus recense tous les concepts inventoriés dans les nomenclatures intégrées à l'UMLS, ainsi que leurs contextes d'apparition dans les dites nomenclatures. Chaque concept est rattaché à au moins un type de concepts. Sur la base de cette ontologie, nous avons bâti une architecture apte à

atteindre notre objectif. Elle se décline en plusieurs niveaux dont l'organisation est la suivante : une interface conceptuelle capable de dialoguer avec un utilisateur ou d'intercepter dans son activité courante ses besoins en information ; un *broker*, reçoit de l'interface conceptuelle une représentation de la requête d'un utilisateur, l'analyse, la traite et recense les différentes sources d'information dans lesquelles rechercher des données utiles; des médiateurs capables de traiter les demandes, et aptes à établir un dialogue avec les différents serveurs, quels que soient leur langages d'interrogation. Cette architecture rend complètement transparent le fait qu'un utilisateur s'adresse à une base de données localisée dans le réseau de son entreprise ou qu'elle soit délocalisée sur l'Internet. La représentation des requêtes et l'indexation des sources d'information se font au moyen de graphes conceptuels mettant en jeu les concepts et les relations sémantiques présents dans les bases de connaissances de l'UMLS. Les opérations effectuées par le *broker* sont des opérations classiques dans la théorie des graphes conceptuels.

**Learnet** [3,4]. Participent au projet : Gérard SOULA, Marius FESCHI et des doctorants.

La diffusion et l'acquisition du savoir médical conditionnent la qualité des soins et le coût de la santé. De volumineuses bases de connaissances sont aujourd'hui disponibles sur le Web. Des Universités Médicales Virtuelles voient progressivement le jour. La majorité des bases de connaissances disponibles ont été développées sans être associées de manière explicite à un projet pédagogique. Elles constituent des bases de ressources pédagogiques brutes, laissant aux enseignants et aux apprenants l'entière initiative pour en tirer une plus value pédagogique. De ce fait, ces bases de connaissance sont rarement et/ou faiblement utilisées dans les cursus de formation médicale initiale et continue car : les enseignants ne disposent pas d'outils pour les intégrer facilement dans leurs projets pédagogiques ; les apprenants ne disposent pas de moyens pour choisir parmi les innombrables bases de connaissances celles qui sont leur apporteront la plus value pédagogique la

plus efficace en fonction de leur profil ; la navigation dans les bases de connaissances ne tient pas compte des différents profils des apprenants. Le projet Learnet vise à concevoir, développer et évaluer des environnements interactifs pour optimiser la diffusion et l'acquisition du savoir médical dans le contexte des TIC. Il se base sur la prise en compte d'aspects pédagogiques et cognitifs au sein de l'organisation des bases de connaissances et de leurs modes d'accès.

Dans un contexte de savoirs largement diffusés, les utilisateurs sont différents en termes de formation, de motivation, de compétence technique ou de tâche à accomplir. Cette variabilité se traduit par des mécanismes d'acquisition du savoir différents. Des modes d'accès à l'information identiques pour tous les utilisateurs n'intègrent pas cette diversité. Il est donc nécessaire de tenir compte de cette diversité pour offrir aux utilisateurs des modes d'interactions adaptés, garants d'accès pertinents et efficaces pour une meilleure diffusion et acquisition du savoir. Contrairement aux hypermédias classiques, un hypermédia adaptatif est capable de modifier ses aspects visuels et fonctionnels de manière à s'adapter à l'utilisateur. La problématique des hypermédias adaptatifs est double : elle nécessite de se doter de moyens pour qualifier les utilisateurs et de définir les différents types de comportement du système en fonction des utilisateurs. Notre projet est de proposer un modèle d'hypermédia adaptatif caractérisé par les trois points suivants : l'utilisation de stéréotypes pour qualifier les utilisateurs ; une présentation adaptative de l'information grâce au typage des granules d'informations et à l'utilisation de modèles de contenu ; une navigation adaptative exploitant les relations conceptuelles entre les pages. Un stéréotype est défini par une collection d'attributs permettant de caractériser les utilisateurs. Notre objectif est de concevoir des règles dynamiques d'affectation et de ré-affectation d'un utilisateur à un stéréotype qui s'appuient sur une évaluation par le système des utilisateurs et de leurs comportements. L'adaptativité de présentation nécessite un typage sémantique des granules d'informations. La navigation adaptative doit permettre lors

de l'activation d'un lien de diriger l'utilisateur vers l'information la plus pertinente. Notre objectif est d'appliquer les résultats de notre recherche dans un contexte institutionnel en intégrant les réalisations dans le cadre de la formation initiale et continue à la Faculté de Médecine de Marseille. La concrétisation de cette démarche est l'étude et la mise en place d'un serveur pédagogique ouvert aux enseignants et apprenants de la Faculté.

### Bibliographie

1. Aymard S, Joubert M, Fieschi D, Fieschi M. Mediation services with health information sources. Overhage M, ed. *Proc. AMIA Annual Fall Symposium* ; 2000 : 37-41.
2. Joubert M, Aymard S, Fieschi D, Fieschi M. ARIANE: a mediation framework with health information sources. Patel V et al., eds. *Proc. MEDINFO 2001* ; 2001 : 343-7.
3. Pagesy R, Soula G, Fieschi M. DIZADEM : an adaptive hypermedia designed to improve access to relevant medical information. Overhage M, ed. *Proc. AMIA Annual Fall Symposium* ; 2000 : 635-9.
4. Soula G, Pagesy R, Fieschi M. Towards an adaptive medical e-learning environment. Patel V et al., eds. *Proc. MEDINFO 2001* ; 2001 : 1076-80.

### Le Laboratoire d'informatique médicale (LIM)

Faculté de médecine Université de Rennes 1

Av du Pr Léon Bernard,  
35043 Rennes cedex

### Composition de l'équipe

Responsable Pierre le Beux

Anita Burgun, Marc Cuggia, Denis Delamarre, Annie Fresnel, Jerome Grall, Christian Jacquelinet, Franck Le Duff, Pierre Le Beux, Christine Riou

### Cinq thèses en cours

Le **laboratoire d'Informatique Médicale (LIM)** regroupe à la fois des compétences sur les méthodes et techniques informatiques appliquées au domaine biomédical et les techniques d'intelligence artificielle allant de la modélisation bio-mathématique aux

systèmes d'aide à la décision et les ontologies en médecine. (Environ une vingtaine de personnes en incluant les théoriciens). Il a commencé par s'intéresser aux systèmes d'aide à la décision, des systèmes experts en médecine et des problèmes de la représentation du langage médical.

Les sujets de recherche du LIM faisant appel à ces méthodes sont :

**Les systèmes de représentation des connaissances médicales pour l'aide à la décision médicale** parmi lesquels les systèmes ADM (Aide au diagnostic Médical) et ATM (aide à la Thérapeutique) [Riou] constituent une première approche. Ces méthodes font appel aussi bien à des méthodes symboliques (traitement du langage médical et intelligence artificielle) qu'à des techniques probabilistes et statistiques (modèles mathématiques et statistiques). Une thèse est en cours sur l'utilisation de réseaux probabilistes de type bayésiens (système DIAMED). [Cleret]

Par ailleurs dans ce thème de recherche nous avons développé le système CARDIOLAB qui modélise un cœur virtuel et les pathologies associées (troubles du rythme et hémodynamique). Il s'agit ici d'une modélisation multi-échelle qui s'insère dans la thématique scientifique de la *physiologie intégrative*. L'incorporation au niveau objet de modèles mathématiques suppose la gestion du passage symbolique/numérique et numérique /symbolique. L'objectif est de permettre aux modules de raisonnement de conduire eux-mêmes des séances de simulation en modifiant les paramètres des modèles qu'ils soient de nature géométrique ou fonctionnels. Nous avons abordé ce problème en développant un modèle qualitatif de l'activité électrique cardiaque. Son rôle est multiple: qualifier les phénomènes de base en électrophysiologie (Siregar), notamment par l'autodescription (par exemple la possibilité de dire « je fibrille »), et interpréter les signaux réels en proposant des séquences d'activation possibles du cœur (les troubles du rythme) susceptibles d'expliquer les observations. Notre approche, basée sur l'intelligence artificielle, combine des

travaux relevant de la physique qualitative, du raisonnement incertain et hypothétique, et du raisonnement spatio-temporel. Nous avons étendu certains de ces travaux afin de les appliquer à la modélisation multi-échelle (Siregar 98), et à l'analyse du signal électrophysiologique. Il correspond à la modélisation de la physiopathologie cardiaque dans son versant électrique et hémodynamique. Un troisième volet, la régulation du système cardio-vasculaire, a aussi été engagée. Ces modèles doivent rendre compte de toutes les données expérimentales disponibles en rapport avec les objectifs cliniques: conduire une analyse quantitative et qualitative du comportement dynamique du système cardio-vasculaire, des processus générateurs des signaux et de la connaissance des propriétés spécifiques des signaux générés et propagés. De cette analyse corrélative entre comportement dynamique et signaux ainsi que de l'analyse quantitative des signaux eux-mêmes, doivent se dégager des critères de comparaison entre groupes de patients conduisant ainsi à la constitution des *modèles de références* associés à chaque groupe de patients.

#### L'Université Médicale Virtuelle Francophone (UMVF)

Par ailleurs Le Campus numérique d'Imagerie Médicale et l'UMVF sont actuellement en cours de développement au LIM. Dans le cadre de ce projet nous travaillons sur les aspects représentation conceptuelle et indexation des contenus multimedia des ressources pédagogiques [Le Beux]

Les recherches sur les ontologies

**Le développement de modèles de systèmes d'informations et de connaissances médicales**, suppose des systèmes de représentation des connaissances médicales (description de concepts dans les nomenclatures et classifications médicales : les ontologies nécessaires pour la description des maladies et syndromes, des thérapeutiques (actes médico-techniques et médicaments). De même, pour l'analyse des textes en langage naturel, les méthodes peuvent être statistiques (analyse de données) syntaxiques et sémantiques (modèles conceptuels et ontologies).

tiques (modèles conceptuels et ontologies).

Cette thématique de recherche a été développée au LIM dans la suite de sa participation au projet MENELAS, représentant les comptes rendus d'hospitalisation selon le formalisme des graphes conceptuels et de sa participation au projet MAOUSSC, décrivant les procédures chirurgicales par des frames. Par la suite XX champs ont été investis plus particulièrement : le Unified Medical Language System (UMLS), les ontologies de sous-domaines ou applications, et les bases formelles.

1. Une part importante de nos travaux concerne les déterminants ontologiques de l'UMLS. L'UMLS est le Unified Medical Language System, système de représentation des concepts du domaine biomédical créé et maintenu par la US National Library of Medicine (NLM), que nous avons expérimenté pour la description des actes médicaux (projet MAOUSSC). La collaboration du LIM avec le Lister Hill National Center for Biomedical Communications, centre de recherche de la NLM, a contribué à définir les orientations du projet MOR (Medical Ontology Research) initié en 2001 par la NLM, par des études portant sur les co-occurrences en tant que composants des espaces sémantiques, sur l'alignement avec d'autres ontologies (Cyc et WordNet) et sur les aspects formels des ontologies. Les travaux en cours en cours portent sur l'analyse et la réorganisation du Semantic Network de l'UMLS, les définitions des termes médicaux, l'acquisition de relations entre concepts à partir de l'analyse lexicale et le transfert de relations sémantiques entre concepts à partir de la base de connaissances d'aide au diagnostic ADM [leDuff].

2. Les ontologies de domaines ou d'applications occupent une place importante en médecine, puisque l'interopérabilité sémantique doit reposer sur une ontologie [Wierhold]. Un travail réalisé par le LIM pour l'Etablissement Français des Greffes montrait que le thesaurus en usage ne permettait pas, du fait de sa structure, de bien représenter toutes les informations

nécessaires à une bonne connaissance de l'activité de transplantation d'organes, et que de multiples systèmes de codage co-existaient pour représenter les données receveur. En conséquence, une ontologie a été construite pour permettre une interopérabilité sémantique entre les différentes bases ainsi qu'une meilleure exploitation des données [Mc Cray 2001].

De la même façon, l'interopérabilité sémantique entre d'une part les bases de données et de connaissances cliniques et d'autre part les bases de données et de connaissances en génomique et biologie moléculaire est un moyen de potentialiser la recherche et de favoriser ses applications pour les patients. Il apparaît que les ontologies développées pour la biologie moléculaire, e.g., [GeneOnto] sont restreintes aux concepts spécifiques de ce champ, et que le champ de la biologie moléculaire est sous-représenté dans les ontologies médicales générales. Par exemple, seulement 0.1 % des concepts de l'UMLS appartiennent au champ de la biologie moléculaire (alors que ce pourcentage est de 49% pour les médicaments) [McCray]. Notre objectif est d'analyser les conditions de l'interopérabilité entre les ressources en biologie moléculaire et les terminologies médicales, en centrant nos travaux sur la représentation des processus biologiques et de la physiopathologie des maladies et sur la représentation des relations sémantiques entre ces concepts.

L'ontologie est un système de connaissances symboliques complémentaires des modèles numériques. Une ontologie appliquée à l'anatomie sulcogyrale est développée dans la thèse d'O. Dameron dans le cadre d'une collaboration avec l'équipe de B Gibaud.

3. Les méthodes utilisées s'appuient sur des propriétés formelles des ontologies de haut niveau. On notera que l'espace de recherche théorique dans le domaine des ontologies formelles reste ouvert. Cependant nous appuyons nos travaux sur les axes de recherche de l'ontologie formelle, qui, à partir d'une vision théorique de l'ontologie issue de la philosophie, élabore des méthodes permettant de construire des ontologies bien formées. Ainsi, notre analyse de

l'UMLS s'appuie sur le système de règles basées sur des propriétés formelles des objets élaboré N. Guarino et C Welty pour contraindre la relation taxonomique [Guarino]. De la même façon, les travaux en mereotopology comme ceux de l'équipe de B Smith à l'Université de Buffalo ont inspiré le modèle de représentation de l'anatomie cérébrale évoqué ci-dessus.

#### Références

1. *GeneOntology* <http://www.geneontology.org/> The Gene Ontology Consortium. Creating the gene ontology resource : design and implementation. Genome Res 2001 Aug ; 11(8) : 1425-1433
- McCray AT, Burgun A, Bodenreider O. Aggregating UMLS Semantic Types for reducing conceptuel complexity. Medinfo 2001, Patel VL, Rogers R & Haux R (Eds.), 10 t1 : 216-220
- Cléret M., Le Duff F., Fresnel A., Le Beux P. DIAMED :a Probabilistic Diagnostic Aid System on the Web Medinfo. 2001;10(Pt 1):429-33.
- Siregar P., Sinteff JP., Julien N., Le Beux P. An interactive 3D Anisotropic Cellular Automata Model of the Heart. Computer and Biomedical Research 1998 31:323-347
- Riou C., Pouliquen B., Le Beux P. A computer Assisted Drug Prescription System: the model and its implementation in he ATM knowledge base. Methods Inf Med. 1999 Mar;38(1):25-30.
- Le Beux P, Le Duff F, Fresnel A, Berland Y, Beuscart R, Brenetaud JM, Burgun A, Chatellier G, Darmoni S, Duvauferrier R, Fieschi M, Gillois P, Guillé F, Kohler F, Pagonis D, Pouliquen B, Soula G, Weber J. The French Virtual Medical University . MIE 2000 Hasman A (ed) IOS Press :554-62
- Le Duff F, Burgun A, Cleret M, Pouliquen B, Barac'h V, Le Beux P. Knowledge acquisition to quality Unified Medical Language System interconceptual relationships. Proc AMIA Symp. 2000;:482-6.

#### Laboratoire d'Informatique Médicale de Bobigny LIMBO

**Adresse :** UFR de Santé, médecine et biologie humaine - Léonard de Vinci  
74, rue Marcel Cachin  
93017 Bobigny cedex  
**Téléphone :** 01 48 38 76 76  
**Télécopie :** 01 48 38 77 77  
**E-mail :** [alain.venot-limbo@lipn.univ-paris13.fr](mailto:alain.venot-limbo@lipn.univ-paris13.fr)

Ce laboratoire nouvellement créé (23 octobre 2001) résulte de la volonté de l'Université Paris 13 de développer le domaine de l'Informatique Médicale qui fait maintenant partie des axes prioritaires inscrits dans son nouveau plan quadriennal.

Il réunit les compétences de personnels hospitalo-universitaires de faculté de médecine (46ème section – 4ème sous section) et d'enseignants chercheurs d'Informatique (27ème section). Il est dirigé par Alain Venot (PU-PH) en mutation, originellement à la faculté Cochin Port Royal (EA 2494).

L'activité s'articule autour du thème de la *diffusion d'un savoir médical spécialisé par rapport aux caractéristiques de l'utilisateur de ces connaissances*. La démarche se place dans le cadre de l'ingénierie des connaissances médicales.

L'objectif poursuivi est la mise à disposition des citoyens et professionnels de santé d'outils d'assistance décisionnelle, utilisables en routine dans leur activité ou pratique quotidienne. Pour l'atteindre, on vise à étudier les processus de modélisation qui permettent le passage sans perte de sens (ni modification du sens) d'une représentation des connaissances médicales fondamentalement textuelle, c'est-à-dire exprimée en langue naturelle propice à la contextualisation, à une représentation formelle au sein de modèles logiques ou mathématiques, par définition insensibles au contexte. Cette recherche conduit à des résultats tant méthodologiques qu'informatiques.

Cependant, cette recherche se veut appliquée et n'a de sens que si elle s'inscrit dans le cadre des usages professionnels, de ville ou hospitaliers, qu'elle cherche à analyser, à comprendre et à améliorer. Aussi, les modèles et outils proposés doivent-ils être validés par des expérimentations en grandeur réelle où l'appropriation des connaissances par les acteurs et l'impact de ces connaissances sur leurs pratiques à travers une évaluation des conduites et qualité des soins dispensés sont évalués de manière scientifique dans le cadre d'essais contrôlés.

Ces travaux prennent particulièrement pour matériau les guides de bonnes pratiques et recommandations

publiées sur des domaines importants de santé publique tels que par exemple le cancer, l'hypertension, le sida, etc.

Pour développer cette thématique, nous avons individualisé les axes de recherche suivants :

- La synthèse des données médicales
- L'aide à la décision en Santé sous les deux aspects documentaires et critiques
- La gestion des connaissances médicales

#### Principaux membres

(Fabrice Bossaert MCF, Jérôme Nobécourt MCF, Philippe Le Toumelin PH, Alain Venot PU-PH)

#### Principaux projets en cours :

Un des thèmes de l'équipe réside dans la représentation des connaissances sur les médicaments et leur utilisation dans des systèmes d'aide à la décision.

Un travail sur la structuration des indications des médicaments telles qu'écrites dans les Résumés des Caractéristiques Produits (RCP) a été récemment terminé et publié. Il a pu bénéficier d'une analyse lexicale réalisée sur l'ensemble des indications de tous les médicaments commercialisés en France et réunies dans un seul fichier. Ce travail conduit à un ensemble d'applications dont la création de nouveaux outils pour aider le prescripteur dans le choix des médicaments qu'il veut prescrire à un patient donné.

Un travail de même nature est en cours qui porte sur les sections pharmacocinétique et pharmacodynamie des mêmes RCP. Dans ce dernier cas où l'on cherche à modéliser ces connaissances, on est confronté à une situation originale car il existe une abondante littérature sur la modélisation mathématique de la pharmacocinétique des médicaments. Il s'agit alors de modéliser des connaissances à partir d'un ensemble de textes (environ 4000), tout en tenant compte des connaissances implicites contenues dans l'approche de la modélisation mathématique à qui l'on doit un certain nombre de concepts fondamentaux du domaine.

Un autre travail en cours dans le pro-

jet ASTI (<http://cybertim.timone.univ-mrs.fr/asti/>) vise à définir la structure et constituer une base de connaissances sur les meilleures stratégies thérapeutiques à utiliser en face d'un patient donné. Les sources sont alors les guides de bonnes pratiques cliniques édités par les sociétés savantes et relatifs à des domaines médicaux focalisés sur certaines pathologies (traitement de l'hypertension artérielle de l'adulte ou de l'asthme par exemple).

Ce dernier projet poursuit les mêmes objectifs que le projet Prodigy au Royaume Uni mais tente de développer une approche différente.

#### Exemples de publications des membres du laboratoire:

1. Seroussi B, Bouaud J, Dreau H, Falcoff H, Riou C, Joubert M, Simon C, Simon G, Venot A. Asti: a guideline-based drug-ordering system for primary care. *Medinfo* 2001;10:528-32
2. Duclos C, Venot A. Structured representation of drug indications: lexical and semantic analysis and object-oriented modeling. *Methods Inf Med.* 2000, 39:83-7.
3. Bennani Y, Bossaert F. Modular Connectionist Modelling and Classification Approaches for Local Diagnosis in Telecommunication Traffic Management", *International Journal of Computational Intelligence and Applications*, World Scientific Publishing Company(2001).
4. Nobécourt J, Biébow B. MdwV : a modeling language to build a formal ontology in either Description Logics or Conceptual Graphs. In *Proceedings of the 12th International Conference on Knowledge Engineering and Management : EKAW'2000*, R. Dieng & O. Corby (eds), LNAI 1937, Springer, p. 57-64

**Laboratoire PSI**  
**Université et INSA de Rouen**  
<http://www.insa-rouen.fr>

**Directeur :** Jean-Pierre Pécuchet,  
*pecuchet@insa-rouen.fr*  
INSA Rouen – BP 08,  
76131 Mont-Saint-Aignan Cedex  
Tel : 02 35 52 83 43  
**Directeur adjoint :** Jacques Labiche,  
*labiche@univ-rouen.fr*

#### Cadre général

Le laboratoire PSI (Perception, Systèmes, Information) est reconnu par le Ministère à la fois comme Equipe d'Accueil (UPRES EA2120) et Equipe de Recherche Technologique (ERT MIIHS avec EADS-Matra Système et Information). L'équipe regroupe près de 80 personnes, dont 35 permanents et 30 doctorants répartis à peu près uniformément entre les deux établissements de tutelle : l'Université et l'INSA de Rouen.

PSI est un laboratoire pluridisciplinaire dont les compétences se répartissent selon 3 thématiques fondamentales : - Document et Système Interactif; - Apprentissage et Classification; - Vision.

Le laboratoire PSI s'appuie sur des partenariats industriels forts, et loin d'être cloisonnées, les équipes disciplinaires contribuent à de nombreux projets transversaux qui peuvent ainsi bénéficier de compétences et d'approches variées.

Nous avons déjà présenté dans le bulletin AFIA N° 45 un panorama général des activités de PSI relevant de l'IA. Nous nous focaliserons ici sur les activités liées au domaine médical à travers trois projets représentatifs :

- l'aide au diagnostic médical par l'image,
- la modélisation des connaissances médicales,
- le dossier médical informatisé.

#### 1. Aide au diagnostic médical par l'image

Le projet Identification et Diagnostic dans l'Incertain (IDI) analyse depuis 1993 des diapositives fournies par des dermatologues pour le diagnostic de mélanomes malins.

Cette équipe est constituée de Denis de Brucq, Olivier Colot, Patrick Vannoorenbergh, Christèle Lecomte, Philippe Leray.

La clinique dermatologique du CHU de Rouen sous la responsabilité du Professeur Philippe Lauret avait proposé un protocole aux dermatologues de la région rouennaise. Ainsi une base d'apprentissage est constituée de 19 mélanomes et 61 naevi et une base de test est

constitué de 18 mélanomes et 191 naevi. Ces deux bases ont été mises à la disposition de l'équipe.

Sur la seule analyse des diapositives, les diagnostics par le logiciel réalisé rivalisent en performance avec ceux des dermatologues.

Pendant un matériel dédié pour obtenir les images améliorerait les performances. De plus, le dermatologue dispose d'informations médicales sur les patients, informations non disponibles dans le protocole.

Des analyses scientifiques introduisant des outils nouveaux de gestion de l'incertain, placent cette étude dans le domaine de l'Intelligence Artificielle. En effet, la nécessité de raisonnements déductif dans un contexte incertain nécessite l'introduction des réseaux bayésiens.

De même, les marqueurs ABCD (Area, Boundary, Color, Depth) des dermatologues ne sont pas déterminants de la présence d'un cancer et pour gérer la croyance apportée par ces marqueurs seule la théorie de Dempster-Shafer (en abandonnant probabilité et statistique) conduit à des résultats crédibles.

D'un point de vue fondamental, ce projet permet la mise au point de méthodologies scientifiques de gestion d'incertitudes plus adaptées aux sciences humaines qu'à la physique traditionnelle.

### Thèses

- Khaled Taouil, "Faisabilité en détection de mélanomes", 1996
- Eric Lefevre, "Fusion adaptée d'informations conflictuelles dans le cadre de la théorie de l'évidence – Application au diagnostic médical", 2001
- Fatima El Matouat, 'Réseaux Bayésiens et Diagnostic Médical - Application au diagnostic de mélanome', en cours

### Publications

1. E. Lefevre, O. Colot, P. Vannoorenberghe, D. De Brucq « Contribution des mesures d'information à la modélisation crédibiliste de connaissance », Revue Traitement du Signal, vol. 17, n°2, pp. 87-97, 2000.
2. E. Lefevre, O. Colot, P. Vannoorenberghe, D. de Brucq : « Informations et combinai-

son : les liaisons conflictuelles », Revue Traitement du Signal, A paraître, 2001.

3. E. Lefevre, O. Colot, P. Vannoorenberghe, D. de Brucq : « Knowledge modeling methods in the framework of belief theory », IEEE international Conference on Systems, Man and Cybernetics, SMC'2000, pp.2806-2811, 2000.

4. O. Colot, R. Devino, A. Sombo, D. de Brucq : « A colour image processing method for melanoma detection », pp. 562-569, Lecture Notes in Computer Science, 1496, Springer-Verlag, 1998.

### 2. Modélisation des connaissances médicales

Le projet s'articule autour du portail CISMef (Catalogue et Index des Sites Médicaux Francophones <http://www.chu-rouen.fr/cismef/>) et concerne la modélisation de l'information pour décrire et indexer des ressources Internet de qualité dans un catalogue. Il est piloté par

Stéfan Darmoni ([Stefan.Darmoni@chu-rouen.fr](mailto:Stefan.Darmoni@chu-rouen.fr), Faculté de Médecine et PSI), s'appuie sur l'équipe CISMef composée de 4 documentalistes, de 2 médecins informaticiens et d'une doctorante encadrée par deux permanents PSI : Catherine Barry et Habib Abdulrab.

Le modèle de connaissances utilisé actuellement dans CISMef 'encapsule' le thésaurus MeSH (mots-clés et qualificatifs) et s'appuie sur quatre niveaux de classification: méta-terme, mot-clé, qualificatif et type de ressource, méta-terme et type de ressource étant propres à CISMef.

Les méta-termes correspondent à des spécialités biologiques ou médicales qui ont des relations sémantiques avec un ou plusieurs mots-clés, qualificatifs ou types de ressource. Quant aux types de ressources, il indiquent la nature de l'information véhiculée, comme par exemple les associations, les informations au patient ou encore les cours. Utilisation du format de méta-données Dublin Core dans le catalogue CISMef.

La présente étude concerne la problématique de la projection de requêtes d'un utilisateur quelconque sur notre terminologie globale.

Trois approches différentes (lexicale, sémantique et statistique) seront étudiées et évaluées.

### Thèse

Lina Soualmia, « Etude et évaluation d'une approche multiple (lexicale, sémantique et statistique) pour la projection de requêtes sur une terminologie normalisée », en cours.

### Publications

1. S J. Darmoni, JP. Leroy, B. Thirion, F. Baudic, M. Douyere, J. Piot. CISMef: « a structured Health resource guide ». Methods of Information in Medicine 2000; Jan;39(1) 30-35.
2. S J. Darmoni, B. Thirion, JP. Leroy, M. Douyere, B. Lacoste, C. Godard, I. Rigolle, M. Brisou, S. Videau, E. Goupy, J. Piot, M. Quéré, S. Ouazir, H. Abdulrab H. « A search tool based on 'encapsulated' MeSH thesaurus to retrieve quality health resources on the Internet ». Medical Informatics & The Internet in Medicine. 2001;26(3):165-178.
3. P. Zweigenbaum, S J Darmoni, N. Grabar . « The Contribution of Morphological Knowledge to French MeSH Mapping for Information Retrieval ». Proc AMIA Symp. 2001; (in press).

### 3. Dossier médical informatisé

DOPAMINE (Dossier Partagé Médical Informatisé) est le projet d'un dossier patient informatisé destiné à une prise en charge coordonnée des patients par l'ensemble des praticiens médicaux de la Région Haute-Normandie dans le domaine de la cancérologie. Mené en partenariat avec EADS-Matra Système et Information, le projet regroupe des permanents de PSI (Jean-Pierre Pécuchet, Catherine Barry, Jacques Labiche, Eric Trupin) et de Matra (Yann Pollet, Nicolas Chanchevri). L'équipe est renforcée par deux doctorants (dont une CIFRE Matra) et plusieurs stagiaires.

Pour répondre à la nécessité d'un système le plus proche possible des habitudes des praticiens, de manière à susciter une adhésion maximale, DOPAMINE vise à proposer un nouveau modèle de dossier patient "orienté document". Dans cette optique, le système classe, range, accède à des documents produits par les praticiens sous la forme de comptes rendus médicaux usuels produits librement selon les habitudes des praticiens, et à l'aide d'un traitement de texte courant. Le dossier est alors une suite de documents organisés

selon le mode chronologique, ou tout autre mode en fonction de l'utilisateur.

Dans DOPAMINE, ce dossier doit être susceptible d'être présenté selon un certain nombre de « vues » présentant autant de manière de percevoir et de synthétiser un dossier que de besoins spécifiques des spécialités utilisateur. A titre d'exemple, il est possible à un praticien oncologue de voir un même dossier organisé, alternativement ou en simultanéité, selon les points de vue de l'histoire clinique et des chimiothérapies, tandis que le praticien généraliste pourra accéder à des synthèses organisées selon les points de vue de la prévention primaire ou secondaire, etc.

Du point de vue de l'IA les défis scientifiques majeurs sont les suivants :

Concevoir une ergonomie permettant à chaque utilisateur de retrouver les usages et les modes de consultation du dossier propre à son métier (généraliste, spécialiste, chirurgien, infirmière, aide-soignant, biologiste, pharmacien, patient, ...).

Concevoir des représentations des connaissances et des algorithmes de fouille de données permettant l'extraction (semi) automatique des informations contenues dans le dossier.

Une des originalités de l'approche proposée réside dans la recherche de coopération de techniques, pouvant combiner l'utilisation de connaissances a priori (ontologies médicales) et de reconnaissance de modèles (structure du document), l'analyse statistique de l'information textuelle, et principalement, l'analyse linguistique locale, pour réaliser cette extraction. On cherchera à prendre en compte une notion d'« apprentissage » permettant l'affinement automatique des traitements et l'évolution supervisée des ontologies à partir des rectifications opérées par l'utilisateur en situation.

#### Thèse

- Dalila Bekhouche : « Analyse du texte médical », en cours.
- Dominique Dionisi : « Génération téléologique de systèmes d'expertise décisionnelle et ontologies évolutives », en cours.

#### Équipe SIC (Laboratoire TIMC)

Contact : Catherine Garbay Catherine.  
Garbay@imag.fr  
Michel Dojat Michel.  
Dojat@ujf-grenoble.fr  
Lab. TIMC - IMAG, Institut Bonniot  
Faculté de Médecine - Domaine de la  
Merci  
38706 La Tronche Cedex - France  
Tél : +33 (0)4 76 54 94 85  
Fax : +33 (0)4 76 54 95 49

#### Cadre Général :

Les recherches de l'équipe SIC (Systèmes Intégrés Cognitifs) concernent la conception de systèmes à base d'agents et plus particulièrement le développement de leurs capacités de focalisation, d'adaptation et de coopération. Ces travaux sont appliqués au traitement d'informations spatio-temporelles (signaux, images) avec des applications centrées sur les sciences du vivant : médecine, biologie et neurosciences. L'équipe se compose de 5 chercheurs et de 6 étudiants (thésards ou postdoctorants). Ils s'articulent plus précisément autour de la notion d'agent situé : de tels agents sont à la fois spécialisés (ils sont dotés d'un but précis à atteindre et de compétences dédiées) et ancrés spatialement et temporellement. Ils disposent de capacités autonomes d'adaptation à leur environnement, exprimées comme des compétences de type fouille de données.

#### Projet scientifique :

Nos travaux précédents nous ont permis de mettre en avant l'importance des principes d'interaction et de localité dans le développement de mécanismes de raisonnement complexes. Ils ont révélé les conséquences profondes de ce choix sur la conception même du système. Nos travaux en cours confirment cette analyse et nous conduisent à modéliser un processus complexe comme l'entrelacement de processus mutuellement contraints qui opèrent à différents niveaux d'abstraction, construisent des buts et des résultats

intermédiaires qui introduisent à leur tour de nouvelles contraintes d'analyse. Un décloisonnement des informations est opéré : toute information peut être utilisée à tout moment pour guider l'analyse, indépendamment de son lieu, niveau et moment d'obtention. Une activité d'analyse est dès lors modélisée comme un problème de couplage entre un système et son environnement dont le but est la construction d'objets. Systèmes et environnements sont conçus comme des structures complexes qui évoluent sous l'action d'agents concurrents et situés.

- *L'analyse d'image comme un processus contraint* : Dans le cadre d'un programme Alliance avec l'Université de Manchester, nous avons cherché à intégrer stratégies guidées par les données et stratégies guidées par les modèles pour l'analyse d'images cérébrales anatomiques. Une première phase d'analyse, guidée par un modèle déformable, fournit une pré-structuration du problème par la localisation des principaux éléments de l'image, et facilite ensuite la mise en œuvre des processus guidés par les données. Ce système utilise plusieurs principes de coopération pour l'interprétation d'images. Le principe de coopération intégrative est à la base de la conception, plusieurs phases de traitement étant enchaînées successivement. Selon le principe de coopération augmentative, plusieurs agents d'analyse sont lancés simultanément pour délimiter les supports de la matière grise et de la matière blanche. Enfin, un contour affiné de l'enveloppe corticale est calculé par fusion (coopération confrontative). [Germond00].

- *L'analyse d'image comme un processus constructif* : Dans le cadre d'une collaboration franco-italienne (programme Galilée), nous avons cherché à modéliser l'analyse d'image (ici des biopsies du foie) comme un processus constructif, ouvert et multi-objectifs. Il s'agit d'un entrelacement constant entre analyses mutuellement contraintes qui opèrent à différents niveaux d'abstraction, utilisent toutes les informations disponibles et contribuent à la construction de buts et résultats intermédiaires introduisant de nouvelles contraintes.

Sur le plan opérationnel, le système est fondé sur une approche linguistique de la représentation des plans d'analyse. Le résultat d'analyse est décrit par un graphe orienté, dans lequel les nœuds sont des images ou des descriptions, et les arêtes les liens opératoires. Des règles et méta-règles agissent de manière concurrente ou séquentielle sur le graphe résultat et sur le plan [Bianchi98].

- *L'analyse d'image comme un processus co-évolutif* : Le problème de l'analyse de scènes est ici posé comme un problème de couplage entre un système et un environnement, dont le but est la construction d'objets. L'analyse de séquences d'images de cellules en culture (fibroblastes de souris) a été étudiée dans ce cadre. Chaque agent situé intègre 4 types de comportements asynchrones: perception, interaction, reproduction et différenciation. Le comportement de perception intègre des critères statiques mais aussi dynamiques, afin de tenir compte de l'information de mouvement pour guider la segmentation. Le comportement de reproduction permet la spécification locale des stratégies d'exploration des images. Le comportement de différenciation s'appuie dynamiquement sur les informations déjà collectées dans l'image. C'est le caractère situé de l'analyse effectuée, associé au principe d'adaptation et de coordination, qui confère aux agents la puissance d'analyse nécessaire, et non la présence de compétences complexes de raisonnement et d'analyse [Boucher98].

#### Panorama des Travaux en Cours :

Nos travaux d'extraction d'informations par agents situés pour l'analyse de signaux et des images se poursuivent sur trois types d'applications :

1) le suivi des patients en anesthésie-réanimation. Au sein du projet OSS-CAR (RNTS 2000) nous mettons en œuvre des techniques d'apprentissage pour extraire, à partir des données physiologiques monitorées, des scénarios typiques qu'il s'agira ensuite de représenter, sur la base d'une ontologie du

domaine [Achour01], puis de reconnaître [Ramaux98], [Dojat00];

2) l'extraction de connaissances en électroneuromyographie (projet EMGNet). Le processus d'extraction est formalisé comme un problème distribué et mis en œuvre par une société d'agents [Balter02];

3) la gestion de données en neuroimagerie. Au travers du projet NeuroBase (ACI Technologie pour la santé 2001) une réflexion est en cours pour définir un modèle et une architecture pour fédérer les différents types d'informations en neuroimagerie (cliniques et cognitives) et faciliter des recherches de type « fouille de données ». D'autre part, diverses applications de la neuroimagerie nécessitent une segmentation précise des structures corticales et sous-corticales. Dans la continuation de [Germond00], nous utilisons des agents situés dont les buts et localisations sont déterminés par rapport à des « cartes » associant, de manière qualitative, une zone dans l'espace image à une description en termes d'hypothèses d'interprétation et/ou de propriétés. Les comportements des agents sont spécifiés *a priori*, selon le but à atteindre, en termes de contraintes de traitement que l'agent adaptera de manière autonome au contexte précis de la zone à étudier (e.g. statistiques locales). Nous distinguons pour l'instant 2 étapes dans ce processus: 1) une carte préliminaire est tout d'abord calculée à partir d'éléments de connaissances topologiques définis *a priori* et de statistiques globales calculées sur l'image. Cette première carte permet le placement d'agents qui évoluent localement, par croissance de régions, dans un cadre de coopération augmentative. Les critères de segmentation sont ajustés localement par chaque agent. Les résultats obtenus par ces premiers agents permettent l'obtention d'une information topologique plus précise concernant les régions recherchées, mais aussi une indication de la robustesse de cette information; 2) une seconde carte est calculée qui induit une focalisation sur les zones peu sûres (exemple sillons), ou encore non marquées. Cette seconde carte permet le placement de nouveaux agents, dont le rôle est d'affi-

ner l'étiquetage précédent. Deux approches sont actuellement considérées pour procéder à un tel affinement: i) une approche par négociation (coopération confrontative) selon laquelle des agents voisins vont négocier l'affectation des pixels qu'ils "contrôlent"; ii) une approche par affinement (coopération intégrative), qui implique la mise en œuvre d'agents travaillant à un niveau de grain plus fin (voisinage d'un point), selon une approche génétique proposée par [Liu, Pami 21 (1999) 544-551]. Ces travaux sont appliqués à la segmentation puis au dépliage du ruban cortical.

#### Bibliographie :

1. C. Garbay (eds) Medical Imaging, Special Issue of the Artificial Intelligence in Medicine Journal, 19 (2), 2000.
2. M. Dojat, S. Miksch, and J. Hunter (eds) Knowledge-Based Information Management in Intensive Care and Anaesthesia, Special Issue of the Artificial Intelligence in Medicine Journal 19(3), (2000).
3. B. Dawant, C. Garbay (eds) Data Fusion, Special Issue, IEEE T-BME, 46(10), 1999.
4. J. Balter, A. Labarre-Vila, D. Ziébelin & C. Garbay : A knowledge-driven agent-centered framework for data mining in EMG, Comptes Rendus de l'Académie des Sciences, 2002 (sous presse).
5. S. Achour, M. Dojat, C. Rieux, P. et al.: A UMLS-based knowledge acquisition tool for clinical decision support system development, JAMIA 8:351-360, 2001.
6. L. Germond, M. Dojat, C. Taylor, C. Garbay: A cooperative framework for segmentation of MRI brain scans, Artificial Intelligence in Medicine, 20(1):77-93, 2000.
7. N. Bianchi, P. Bottoni, C. Spinu, C. Garbay, P. Mussio: Situated image understanding in a multi-agent framework, International Journal of Pattern Recognition and Artificial Intelligence, 12 (5):595-624, 1998.
8. A. Boucher, A. Doisy, X. Ronot, C. Garbay: Cell migration analysis after in vitro wounding injury with a multi-agent approach, Artificial Intelligence Review, 12:137-162, 1998.
9. M. Dojat, N. Ramaux, D. Fontaine: Scenario Recognition for Temporal Reasoning in Medical Domains, Artificial Intelligence in Medicine 14:139-155, 1998.

**Équipe SPIM  
Santé Publique et  
Informatique Médicale ( )**

**Adresse :**

Santé Publique et Informatique Médicale (SPIM), UFR Broussais-Hôtel-Dieu, 15 rue de l'école de médecine, 75006 Paris

Secrétariat : 01 42 34 69 83

Mel : [jaulent@spim.jussieu.fr](mailto:jaulent@spim.jussieu.fr)

Le SPIM est issu du regroupement entre le laboratoire de Biostatistiques et Informatique Médicale de l'hôpital Broussais (directeur : P.Degoulet) et le laboratoire de Santé Publique du Professeur Joël Ménard. Sa création a conduit à une synergie entre trois thématiques de recherche : 1) informatisation du dossier médical; 2) méthodologies et outils informatiques d'aide à la décision en médecine ; 3) évaluation des pratiques. L'équipe travaille actuellement sur des projets de recherche visant à utiliser les méthodes et outils d'ingénierie des connaissances dans le contexte de l'e-Santé. L'équipe est actuellement composée de 2 chargées de recherche INSERM (HDR), 3 PU-PH, 3 MCU-PH (dont 1 HDR), 3 assistants hospitalo-universitaires, 2 ingénieurs de recherche, 6 étudiants en thèse d'Informatique Médicale, 4 stagiaires de DEA et une secrétaire.

**Historique**

Les activités de l'équipe depuis dix ans concernent les thèmes de recherche suivants :

***La modélisation de l'information médicale.***

Ces activités de recherche ont concerné le développement de méthodes informatiques destinées à faciliter la gestion automatique des données et connaissances médicales nécessaires à la prise de décision.

***Aide à la décision.***

Les travaux de recherche développés par l'équipe dans ce domaine ont concerné à la fois l'acquisition de la connaissance par des méthodes d'apprentissage ou de consensus, la modéli-

sation et la représentation-structuration de la connaissance, les mécanismes de raisonnement et les modes de communication et transmission des connaissances. La représentation et la gestion de l'imprécis et de l'incertain sont intégrés par le biais de la théorie des probabilités mais également par la théorie des sous-ensembles flous et la théorie des possibilités.

***Evaluation des pratiques de soins.***

Les travaux visent à analyser des déterminants de la qualité des soins et mettre en place et évaluer des interventions destinées à mettre en œuvre les recommandations pour la pratique clinique.

**Contexte**

Les nouvelles technologies s'imposent dans le monde de la santé comme l'un des nouveaux supports de la communication entre le médecin et son patient (e-Santé). Le type de problème auquel est confronté la recherche en ingénierie des connaissances pour l'e-Santé concerne le développement d'environnements de gestion de connaissances sur les réseaux de santé dans des domaines médicaux cibles, permettant de garantir la qualité des connaissances diffusées et la normalisation des échanges dans ces environnements ainsi que d'en mesurer l'impact sur la pratique médicale.

**Travaux actuels**

Les travaux de thèse en cours dans l'équipe sont : 1) Représentation des connaissances pour la construction d'un modèle de prédiction cardio-vasculaire, 2) Outils de médiation pour l'intégration des applications médicales distribuées et hétérogènes, 3) Distances sémantiques entre termes médicaux, application à la pharmacovigilance, 4) opérationnalisation des connaissances, application à l'informatisation des guides de bonnes pratiques cliniques, 4) gestion des connaissances et sevrage tabagique, 5) extraction de connaissances dans une base épidémiologique, application à la mortalité cardio-vasculaire, 6) utilis-

tion des nouvelles technologies de l'information pour l'exploitation médicale d'images anatomo-pathologiques.

Nous présentons plus précisément les deux projets suivants :

***Aide au dépistage et à la prévention cardiovasculaire (projet ESPER)***

Le projet ESPER (ESTimation Personnalisée du Risque ou Examens de Santé Périodiques et Recommandés) vise à démontrer que la démarche scientifique qui consiste à estimer un risque peut devenir une démarche clinique dès lors que le médecin a accès sur le « site de soins » à des estimations plus rationnelles des risques d'un individu, et aux recommandations permettant la gestion de ce risque. Un prototype de serveur Internet pour le projet ESPER a été développé en collaboration avec le CEC1-INSERM (Paris), le CERTIM (Marseille) et le CNGE (Collège National de Généralistes Enseignants). Ce prototype propose une présentation des dix premières causes de mortalité permettant de hiérarchiser les risques et recommandations plus spécifiques (calcul du risque cardio-vasculaire, recommandations de prise en charge des facteurs de risque cardio-vasculaires et recommandations de dépistage du cancer du sein et du cancer colo-rectal).

***Constitution et diffusion de bases de cas multi-média, multiexpertes (projet IDEM)***

L'objectif du projet est l'utilisation des nouvelles technologies de l'information pour l'exploitation médicale de bases de cas anatomo-pathologiques. Il s'agit de répondre au problème bien identifié de variabilité diagnostique en anatomie pathologique dont le retentissement sur la prise en charge diagnostique et thérapeutique des patients a fait l'objet de nombreuses études.

Un premier prototype de l'environnement IDEM est aujourd'hui opérationnel dans le domaine de la pathologie tumorale mammaire. IDEM permet aux pathologistes de visualiser et d'annoter des images, de construire des bases de cas multi-expertes en anatomie pathologique et d'interroger ces bases par le contenu.

**Méthodes****Acquisition des connaissances à partir de textes pour la construction d'ontologies dans le domaine médical**

Utilisation d'outils d'analyse de corpus de comptes rendus médicaux pour construire des ontologies. Une ontologie a été réalisée dans le domaine de la réanimation chirurgicale

Modélisation des recommandations issues des textes de guides de bonnes pratiques cliniques

Définition d'un cadre de comparaison de termes médicaux par réutilisation des nomenclatures et ontologies (distance sémantique, tables de similarité, etc.). Application à la recherche d'effets indésirables proches en pharmacovigilance. Application dans le cadre du projet IDEM

**Acquisition de connaissances à partir de bases de données**

L'aspect particulier qui nous intéresse est l'utilisation de méthodes d'apprentissage symbolique issues de l'Intelligence Artificielle pour construire des modèles de prédiction d'un risque. Nous avons utilisé une méthode basée sur l'induction de prototypes flous pour construire un modèle de prédiction dans le domaine cardio-vasculaire.

**Recherche par le contenu de cas similaires dans des bases de cas**

Dans le domaine de l'imagerie médicale, cette approche est utilisée pour offrir des possibilités de recherche « intelligente » à des bases d'images, notamment par la description du contenu des images, qui soient adaptées à l'utilisation dans le cadre de la démarche diagnostique quotidienne.

**Position de la recherche****Aide à la prédiction du risque cardio-vasculaire**

Dans le domaine de la prévention cardio-vasculaire, plusieurs auteurs ont montré la difficulté des médecins à estimer le risque de façon reproductible. A l'heure actuelle la prédiction quantitative du risque cardio-vasculaire repose sur l'exploitation par des méthodes statistiques des données issues d'enquêtes de cohortes.

Des méthodes d'apprentissage ont été

développées pour établir des modèles de prédiction à partir de bases de données, en particulier des bases de données cliniques. La plus classique est l'approche statistique fournissant des modèles tels qu'une équation de régression logistique. Mais, d'autres existent comme les réseaux de neurones ou les approches symboliques (arbres de décision). Sur la base des différentes approches, de nombreux systèmes et modèles de prédiction ont été développés et validés au moins partiellement dans des domaines précis.

**Recherche de cas similaires dans une base d'images**

L'approche du Raisonnement à Partir de Cas (RàPC) constitue une des techniques permettant non seulement un accès aux images médicales basé sur la description morphologique mais également une utilisation « intelligente » d'une mémoire de cas illustrés pour l'aide au diagnostic. L'utilité réelle d'un système de RàPC est largement conditionnée par la validité de la base de cas. Or, l'étude de la littérature montre que la variabilité d'analyse, de compréhension et de description de l'information contenue dans l'image entre pathologistes est importante. L'utilisation de termes provenant d'un vocabulaire médical contrôlé international comme SNO-MED a été proposé dans de nombreuses études.

Le problème de construire des bases de cas valides est appréhendé ici comme un problème d'acquisition de connaissances à partir d'experts. La mise en place d'une session de consensus autour d'un microscope est la solution la plus naturelle pour que les experts se mettent d'accord sur les caractéristiques pertinentes pour le diagnostic, leur localisation dans les images et les termes utilisés pour les décrire. Par contre, la procédure est souvent rendue difficile par des problèmes d'organisation. Le développement d'outils informatiques peut faciliter la procédure de consensus, améliorer les conditions du consensus et faciliter la construction de ces bases de cas valides.

**Bibliographie**

1. Le Bozec C, Jaulent MC, Zapletal E. IDEM : Remémoration de cas pour l'aide au

diagnostic en Anatomie Pathologique. In : Ingénierie des connaissances. J.Charlet et. Al. (eds). Eyrolles:Paris - Collection technique et scientifique des télécommunications, 2000 ; pp371-386.

2. Bousquet C, Jaulent MC, Chatellier G, Degoulet P. Using semantic distance for the efficient coding of medical concepts. *Proc. AMIA Symp. 2000* ; 96-100

3. Colombet I, Ruelland A, Chatellier G, Gueyffier F, Degoulet P, Jaulent MC. Models to predict cardiovascular risk: comparison of CART, Multilayer perceptron and logistic regression. *Proc. AMIA Symp. 2000* ; 156-60

**Mission de recherche STIM (AP-HP)**

STIM/DPA/DSI/AP-HP

CHU Pitié-Salpêtrière,  
91, boulevard de l'Hôpital,  
75634 PARIS CEDEX 13.

Secrétariat téléphonique :  
01 45 83 67 28

Courriel générique :

[diam@biomath.jussieu.fr](mailto:diam@biomath.jussieu.fr)

URL : <http://www.biomath.jussieu.fr/>

Membres de l'équipe intervenant sur le thème IA et médecine : Jacques Bouaud (01 40 77 96 22), Jean Charlet, Brigitte Séroussi (96 23), Pierre Zweigenbaum, {jb,jc,bs,pz}@biomath.jussieu.fr

**Doctorants** : Américo Carvalho, Yun-Chuang Chiao, Natalia Grabar, Pierre Jacquemart

**Objectifs et cadre général de l'équipe**

Les activités de la mission de recherche en sciences et technologies de l'information médicale (STIM) reposent historiquement sur des travaux de recherche fondamentale et appliquée avec des objectifs de conception, développement, et réalisation de méthodes et d'outils pour (i) l'aide à la décision médicale, (ii) faciliter l'accès aux connaissances médicales dans un objectif de recherche d'informations, que cette recherche mette en œuvre des techniques de traitement du langage naturel ou que ces connaissances soient représentées dans des formats semi-structurés

rés au sein de documents électroniques, et (iii) la prise en compte des informations médicales (dossier patient, guides de bonnes pratiques, etc.) dans un contexte de gestion des connaissances. Issus de la communauté scientifique de l'intelligence artificielle et engagés dans des projets finalisés mettant en œuvre l'évaluation en pratique des outils d'informatique médicale développés pour servir ces trois objectifs, les membres du STIM sont actuellement impliqués dans des réflexions qui placent le texte, le document et plus généralement les approches documentaires au centre du paradigme d'accès à l'information médicale. Ce choix se justifie parce qu'il permet une meilleure prise en compte du contexte, condition indispensable à l'appropriation et à l'usage effectif par les professionnels de santé de systèmes informatiques à base de connaissances.

#### **Modélisation de l'aide à la décision pour la mise en œuvre de guides de bonnes pratiques**

Dans un souci d'optimisation de la qualité des soins, les « guides de bonnes pratiques » (GBPs) visent à harmoniser les pratiques médicales. Toutefois, la simple diffusion de ces GBPs sous une forme textuelle n'ayant pas d'impact sur les pratiques, de nombreux travaux cherchent à intégrer les GBPs au sein de systèmes à base de connaissances qui, utilisés sur les lieux de la décision, sont alors susceptibles de modifier les pratiques. Pourtant, ces systèmes recueillent peu l'adhésion de leurs utilisateurs, ce qui se traduit par une faible observance des recommandations des GBPs. Ces systèmes s'appuient sur les méthodes et outils traditionnels de l'intelligence artificielle où le domaine considéré est classiquement modélisé au travers d'une formalisation.

Nous considérons que la mise en correspondance automatique entre des données patient et des descripteurs codés indépendamment du contexte introduit un « biais de formalisation ». L'approche de l'aide à la décision proposée consiste à déléguer la tâche d'interprétation au médecin, en lui présentant sous une forme documentaire des

connaissances par ailleurs modélisées dans le cadre formel d'un arbre de décision.

#### **OncoDoc, assistance à la décision thérapeutique en oncologie et au recrutement dans les essais cliniques**

Une première application de l'approche a été réalisée sur le traitement du cancer du sein avec le système OncoDoc [2]. La base de connaissances a été constituée en s'appuyant sur l'état de l'art avec l'aide d'un oncologue médical. Le système a été testé lors d'une première expérimentation en routine hospitalière au Service d'Oncologie Médicale Pitié-Salpêtrière. Pour chaque patiente, le clinicien aboutissait à la proposition de recommandations appropriées, à l'issue d'une navigation « centrée-patient » dans la base de connaissances. Le taux d'observance (*i. e.* fréquence du choix d'une recommandation du système pour une patiente) a été mesuré à 80 %, ce qui est largement supérieur aux chiffres généralement publiés.

Dans une expérience de réutilisation, le système a été ensuite évalué à l'Institut Gustave Roussy. Le taux d'observance a été aussi bon que lors de la première expérimentation (85 %) confirmant la qualité des bases de connaissances conformes à l'état de l'art et exemptes des spécificités du site à l'origine de leur développement. L'impact du système a été mesuré dans le cadre d'une étude avant/après. Après utilisation du système, les médecins de l'IGR ont totalement modifié leur décision thérapeutique dans 15 % des cas et le taux de recrutement dans les essais a été amélioré de 50 %.

#### **ASTI, assistance à la stratégie thérapeutique basée sur les guides de bonnes pratiques pour la prise en charge de l'hypertension artérielle**

L'objectif du projet ASTI (Assistance à la Stratégie Thérapeutique Informatisée), partiellement subventionné par le MENRT dans le cadre de l'ACI « Télémedecine et technologie pour la santé », est de concevoir un système qui permette aux médecins généralistes

d'éviter les erreurs de prescription et de suivre les recommandations thérapeutiques des GBPs. Le mode « guidé » s'inspire de l'approche documentaire développée dans OncoDoc. L'objectif est de tester la réutilisation des principes élaborés dans le cas du traitement du cancer du sein, dans la prise en charge d'une maladie chronique, l'hypertension artérielle, nécessitant le suivi des patients sur de nombreuses années et l'exploration selon des niveaux d'intention plus ou moins protocolisés de l'arsenal thérapeutique. La difficulté de ce travail a été d'extraire à partir de GBPs textuels les connaissances à formaliser sous la forme d'un arbre de décision sans l'aide d'un spécialiste du domaine dont le rôle avec OncoDoc avait été de structurer l'état de l'art. Un prototype, associé à un logiciel de dossier médical, doit être évalué dans les cabinets de généralistes [3].

#### **Recherche d'informations médicales**

##### **Traitement automatique du langage médical**

Les travaux de l'équipe en traitement automatique de la langue médicale sont centrés sur la problématique de l'accès aux informations et aux connaissances médicales exprimées sous forme textuelle. Ils s'appuient pour cela sur l'étude de la terminologie médicale, aussi bien à partir de ressources existantes (thesaurus, nomenclatures, classifications, etc.) que de corpus. Ces travaux concernent :

le recensement du vocabulaire (mots, termes) employé en médecine, sous une forme utilisable par des systèmes d'analyse automatique de textes [4]

la mise au point de méthodes pour reconnaître des variantes de termes, en particulier associées à des variantes morphologiques de mots, et aider à l'acquisition et la structuration de terminologie à partir de corpus ;

le recensement des différents types de documents textuels que l'on peut rencontrer dans le domaine médical, et la collecte d'échantillons de documents de ces types, qui doit fournir une base systématique d'étude des points précédents (projet Corpus CLEF <<http://www.bio-math.jussieu.fr/CLEF/>>);

la recherche d'information « trans-langue » : accès à des documents médicaux français et anglais à partir de requêtes monolingues ;

la mise au point et l'évaluation d'outils employant le matériel et les méthodes précédentes pour accéder à des informations et connaissances médicales textuelles ; une collaboration est en cours avec le CHU de Rouen sur l'utilisation de variantes morphologiques pour faciliter l'accès en langue naturelle au catalogue CISMef [5].

### Structuration des connaissances

P2VIE (Prévention et Prise en charge du Vieillessement par l'Information Électronique) est un projet de 3 ans labellisé par le RNTS dans son appel à propositions de 2000 dont l'objectif est de développer un site web à destination des seniors. Au sein de ce projet, l'équipe est responsable de la réalisation d'un serveur de connaissances médicales. Les informations seront structurées pour servir deux types de recherche d'informations: (i) la recherche d'informations générales organisées en documents textuels permettant de synthétiser à un niveau de langage et de détail variables choisis par l'utilisateur des informations médicales en rapport avec la prévention, les premiers symptômes, le diagnostic, le traitement, l'évolution et le pronostic d'un ensemble de pathologies, (ii) la recherche d'informations spécifiques par navigation à travers des bases de connaissances structurées permettant l'instanciation de paramètres centrés patient, la construction d'un contexte, et l'identification de recommandations personnalisées. Un mécanisme d'import-export entre ces contextes et les données d'un carnet de santé électronique doit permettre de capitaliser dans le carnet les informations dérivées des navigations, et de court-circuiter les navigations à partir des données du carnet <<http://www.p2vie.org/>>.

### Ontologies et gestion des connaissances

#### Hospitexte

À la suite du projet DOME, qui avait permis de valider auprès des médecins

une conception hypertextuelle de l'accès à l'information, le projet HOSPITEXTE a développé l'informatisation du dossier médical dans une approche d'« ingénierie documentaire » qui conserve l'information médicale dans son format privilégié d'élaboration, le document textuel. Le prototype réalisé a permis de valider le concept sur 10 dossiers médicaux réels.

L'étape suivante est de tester les capacités du dossier médical à être le pivot d'un système de gestion des connaissances, que ce soit pour naviguer dans le dossier en fonction de critères médicaux ou pour mobiliser des ressources externes comme des guides de bonnes pratiques ou de la bibliographie [1]. La structure conceptuelle permettant d'envisager de tels services est une ontologie qui permette de pointer les concepts informatiques nécessaires à l'informatisation des services proposés. Cette approche va être testée en coopération avec le LARIA, au sein du projet HTSC visant à informatiser le dossier médical d'un service de néo-natologie.

### Thésaurus et ontologies

Les thésaurus médicaux sont des ensembles de termes hiérarchisés, développés au sein de chaque spécialité. Ils permettent aux praticiens de la spécialité de décrire les patients (p. ex. écriture de compte rendu d'hospitalisation) dans un vocabulaire contrôlé fourni par le thésaurus et de « coder » ce patient au sein du codage PMSI. Ce thésaurus, comme la plupart, n'est pas adéquat à un usage informatique, à savoir servir comme référentiel univoque au sein d'un système informatique, en raison d'abord, de sa trop grande proximité avec la langue et, corollaire, de sa structuration mal spécifiée. La solution est de développer une ontologie formant le squelette conceptuel de ce thésaurus, permettant principalement :

- de décrire le thésaurus grâce à des concepts univoques, lui donnant une réelle robustesse dans un contexte d'évolution permanente des connaissances et de la pratique médicale ;
- de compléter le thésaurus ;
- de faire ainsi le lien entre le vocabulaire contrôlé utilisé par le médecin (le

thésaurus) et les services informatiques à rendre (recherche d'informations au sein d'un dossier médical, aide au codage, etc.) ;

Après le développement, par les acteurs du projet MENELAS, d'une ontologie de la chirurgie cardio-vasculaire, une ontologie de la réanimation chirurgicale <<ftp://ftp.biomath.jussieu.fr/pub/papers/LeMoignoDEAIM2001.pdf>> a été développée à partir de corpus en utilisant les acquis de l'Ingénierie des connaissances et avec l'aide de l'outil LEXTER/SYNTEX [].

### Références

- [1] CHARLET J. (1999). Daigne M., Leroux V., Ingénierie des patrimoines informationnels de l'établissement de santé, *Document numérique* vol. 3, 1999. Numéro spécial « Gestion des documents et gestion des connaissances ».
- [2] SEROUSSI B., BOUAUD J. & ANTOINE E.-C. (2001a). OncoDoc, a successful experiment of computer-supported guideline development and implementation in the treatment of breast cancer. *Artificial Intelligence in Medicine*, 22(1), 43-64.
- [3] SEROUSSI B., BOUAUD J., DREAU H., FALCOFF H., RIOU C., JOUBERT M., SIMON C., SIMON G. & VENOT A. (2001b). Asti, a guideline-based drug-ordering system for primary care. In V. L. PATEL, R. ROGERS & R. HAUX, Rédacteurs, *Proceedings of the 10 th World Congress on Medical Informatics*, London, UK.
- [4] ZWEIGENBAUM P. (2001). Ressources pour le domaine médical : terminologies, lexiques et corpus médicaux. *Lettre de l'ELRA. À paraître*.
- [5] ZWEIGENBAUM P., DARMONI S. & GRABAR N. (2001). The contribution of morphological knowledge to French MeSH mapping for information retrieval. *Journal of the American Medical Informatics Association*, 8(suppl), 796-800.

### Intelligence Artificielle et Neurologie à l'INSERM/U455

#### Cadre général

L'unité de recherche 455 « Neuro-imagerie fonctionnelle, plasticité cérébrale et pathologie neurologique » de l'INSERM, dirigée par le Pr. François

Chollet, a pour objectif l'étude des processus d'adaptation du cerveau humain, dans les domaines du langage et de la motricité. Ses travaux concernent en particulier l'adaptation du cerveau à la suite de lésions, dégénératives (maladie de Parkinson, démence de type Alzheimer, Chorée de Huntington) ou focales (accident vasculaire cérébral), ou de pathologies développementales (dyslexie). Groupe pluridisciplinaire d'environ 50 personnes, structuré en 4 équipes, l'U455 utilise les outils de la neuro-imagerie fonctionnelle (Tomographie par Emission de Positrons, IRM fonctionnelle, EEG quantifiée), de la neurologie, de la neuropsychologie, du traitement du signal et de l'image, de la modélisation mathématique et de l'IA.

#### Groupe inter-équipes « IA et Neurologie »

Au sein de l'U455, ce groupe mène des recherches en IA, guidées par la nécessité de résoudre des problèmes posés en Médecine. Les travaux concernent : la modélisation de la dynamique cérébrale à l'échelle des réseaux d'aires, pour une meilleure compréhension des données de neuro-imagerie (projet MITIC<sup>6</sup>), le diagnostic des troubles du raisonnement dans certaines pathologies neurologiques (projet EARTH<sup>7</sup>).

Le groupe comprend 1 ingénieur de recherche, HDR en Informatique (IA), 3 chercheurs spécialistes en neuro-imagerie, dont 2 médecins, 1 chercheur neurologue, 3 étudiants en thèse, dont 1 en IA et 1 en co-direction avec l'ONERA, et 3 stagiaires, dont un neurologue étranger. Le groupe bénéficie de l'appui de la Fédération des Services de Neurologie du CHU Purpan.

Responsable : Josette Pastor  
INSERM U455, Fédération des Services de Neurologie, CHU Purpan, 31059 Toulouse cedex  
Courriel :  
[josette.pastor@toulouse.inserm.fr](mailto:josette.pastor@toulouse.inserm.fr)

#### Projet MITIC

L'utilisation de l'imagerie fonction-

6. Modélisation Intégrée du Traitement de l'Information Cérébrale.

7. Evaluation Artificielle du Raisonnement Temporel Humain.

nelle en neurologie vise à déterminer les correspondances entre structures cérébrales et fonctions cognitives pour a) une meilleure évaluation du handicap possible à la suite d'une pathologie neurologique, b) la mise au point de stratégies de rééducation faisant appel à la plasticité cérébrale. Si la neuro-imagerie apporte des informations sur la localisation de l'activité cérébrale et/ou sur la chronologie des événements cérébraux, l'interprétation ne peut venir que de modèles des mécanismes de traitement et de propagation de l'information dans les réseaux d'aires interconnectées qui sont le support neurologique des fonctions cognitives (Bressler, Brain Research Reviews, 1995).

Le projet MITIC comporte 3 volets :

- Proposer un cadre formel, intégrant une « théorie minimale du cerveau » (connaissance certaine ou consensuelle), dans lequel pourront être exprimés à la fois les réseaux de régions cérébrales, l'information cérébrale et ses mécanismes de traitement de l'information,
- Construire un système, mettant en œuvre le formalisme, suffisamment convivial pour permettre aux neurologues de décrire et de simuler leurs modèles relatifs à certains sous-systèmes cérébraux et interprétatifs des données d'imagerie.
- Proposer des modèles de mécanismes cérébraux généraux qui iront, après validation, enrichir la théorie minimale.

Notre démarche, initiée en 1994, est aux frontières de la neuro-imagerie (Büchel & Friston, Cerebral Cortex, 1997) et des *neurosciences computationnelles* (Arbib et al., MIT Press, 1998), dont elle préfigurait le rapprochement récent (Numéro Spécial Neural Networks 2000). Elle demeure originale sur le choix du formalisme.

Le formalisme doit respecter des contraintes imposées par :

- **la conformité à la "théorie minimale du cerveau" :**
  - *architecture en réseaux orientés* dont les nœuds (les structures cérébrales) peuvent être fonctionnellement différents et, structurellement, à des niveaux de granularité différents,
  - *causalité entre événements céré-*

*braux* (l'activation d'une structure est la conséquence de celles des régions en amont dans le réseau),

- représentation de *l'information cérébrale au niveau intégré* des populations neuronales,
- *dynamique probabiliste et non-linéaire* du cerveau,
- *capacités d'adaptation et d'apprentissage*,
- *coexistence d'automatismes et de processus contrôlés*,

#### • l'objectif :

- *expression explicite des fonctions*,
- modélisation conjointe des phénomènes neuronaux et vasculaires afin de *reproduire les données de neuro-imagerie*,
- prise en compte explicite d'*échelles de temps très différentes* (de la milliseconde pour l'activité cérébrale à plusieurs secondes pour une acquisition en PET),
- traitement de *l'imprécision des données de neuro-imagerie*.

Notre **formalisme, fondé sur les réseaux causaux fonctionnels** (Pearl, Cambridge University Press, 2000), en étend les possibilités afin de se conformer aux contraintes ci-dessus. **MoCCa<sup>8</sup>, le système** implémentant le formalisme, permet une définition graphique des modèles et leur simulation. Bien que toutes les contraintes n'aient pas encore été intégrées, MoCCa a déjà trouvé une application dans le domaine de la dyslexie et une autre a débuté dans celui des accidents vasculaires cérébraux.

Lors d'une collaboration passée avec le LAAS/CNRS, le système BIO-CAEN<sup>9</sup>, fondé sur les réseaux causaux qualitatifs avait été développé. Si les limitations du formalisme de base empêchent son utilisation pour la modélisation cérébrale, il peut avoir des retombées industrielles (processus nécessitant une traçabilité des produits).

#### Publications choisies :

V Labatut, J Pastor, 2001 : Bayesian Modeling of Cerebral Information Processing, Workshop « Bayesian Models in Medicine », AIME'01, pp 41-46.

V Labatut, 2000 : Réseaux causaux : quelle approche pour le cerveau ? DEA RCFR, Université Toulouse III.

8. MODélisation Cérébrale CAusale.

9. BIOlogical Causal ENgine.

J Pastor, M Lafon, L Travé-Massuyès, B Doyon, J-F Démonet, P Celsis, 2000 : Information processing in large-scale cerebral networks: The causal connectivity approach. *Biological Cybernetics*, 82(1) : 49-59.

M Lafon, 1998 : BioCaen : Simulateur d'hypothèses sur la propagation de l'information cérébrale. *RJCAI*, pp 139-142.

### Projet EARTH

Le syndrome frontal peut apparaître dans plusieurs pathologies neurologiques (Brown & Marsden, TINS, 1990) ou psychiatriques. Ce syndrome dégrade le raisonnement (calcul des propositions : Natsopoulos et al., *Cortex*, 1997), en particulier le *raisonnement en situation dynamique*. Cependant, la neuropsychologie ne propose aucun moyen d'évaluation de ce dernier. Or, la prise de décision en environnement naturel, donc très dynamique, y fait appel de façon massive. De plus, la comparaison d'études en neuropsychologie et en ergonomie (Amalberti & Deblon, *International Journal of Man-Machine Studies*, 1992) montre que les patients présentant un syndrome frontal et les sujets sains, fatigués ou confrontés à une situation stressante ou complexe, peuvent commettre les mêmes types d'erreurs.

L'objectif du projet EARTH est donc de définir une méthodologie et un environnement informatique pour l'évaluation du raisonnement en situation dynamique de différents types de sujets (patients, sujets-témoins, opérateurs...). Il comporte 2 volets :

- Construction d'un environnement de test
- Proposition d'une méthodologie d'évaluation du raisonnement

**WAHRPS**<sup>10</sup>, l'environnement expérimental de EARTH, est un générateur de micro-mondes dynamiques. Chaque micro-monde représente un réseau de cuves, dans lequel l'eau coule par gravité et qui peut être contrôlé par des actions sur des vannes. WAHRPS comporte un éditeur graphique de micro-mondes, qui permet de créer la structure la mieux adaptée aux sujets étudiés, et un simulateur, capable d'animer n'im-

porte quel miro-monde. Cet environnement, qui demande peu de pré-requis culturel, peut facilement être proposé à des sujets d'âge et de niveaux culturels très différents. Il est assez réaliste pour créer des situations complexes, imposant de fortes contraintes aux opérateurs industriels.

La méthode d'évaluation est fondée sur 2 principes : l'existence de styles cognitifs différents, fonction de l'éducation ou de la personnalité, et le fait que les seules mesures objectives du raisonnement d'un sujet sont ses actions. A chaque style cognitif, correspond un raisonnement-type, implémenté par un Raisonneur *Artificiel de Référence (RAR)* qui calcule à tout instant la « meilleure » (au sens de critères qui lui sont propres) action à effectuer sur un micro-monde. Le raisonnement d'un sujet est évalué par la comparaison de ses actions à celles du RAR qui implémente le style cognitif le plus proche de celui du sujet. Cette approche a été concrétisée par la définition d'une famille de RARs planificateurs.

Certaines des pathologies-cibles du projet étant des pathologies du vieillissement (maladie de Parkinson), une première étape a été l'étude de l'impact du vieillissement normal. Les premières expériences sur les patients sont en cours, ainsi qu'une étude visant à étudier l'impact sur le raisonnement de certains médicaments. Les applications industrielles se font, en collaboration avec l'ONERA de Toulouse, dans le domaine de l'Aéronautique.

### Publications choisies

J Pastor, J-F Démonet, A Agniel, P Celsis, 2001 : *From Neuropsychology to Human Factors: Dynamic Micro-worlds and AI-based Models for the Assessment of Reasoning* (soumis)

J Pastor, 2000 : *Cognitive Performance Modeling – Assessing Reasoning with the Earth methodology* -. COOP'2000, Workshop « Modelling Human Activity », pp 5-17.

L Bayssié, M Boyer, L Chaudron, J Pastor, 2000 : *Programme d'études des corrélations entre le bruit et la gêne en termes de performances cognitives*. ALCAAH, pp 30-42.

J Pastor, A Agniel, P Celsis, 1998 : *Artificial Reasoners for the Cognitive Assessment of Patients with Parkinson's Disease*. ECAI98, pp 119-123.

**Unité INSERM 494 - imagerie médicale quantitative**  
**Détection, caractérisation et classification de motifs par des méthodes adaptatives dans des images biomédicales et des séquences génomiques.**

Unité INSERM 494 (imagerie médicale quantitative)

91 bd de l'hôpital

75634 PARIS Cedex 13

Bernard FERTIL CNRS

[fertil@imed.jussieu.fr](mailto:fertil@imed.jussieu.fr)

### Contexte

Une part sans cesse croissante de l'activité de Recherche est consacrée de nos jours à ce qu'il est convenu d'appeler l'analyse statistique de motifs (Statistical Pattern Recognition). Elle concerne des applications aussi diverses que la recherche de documents sur Internet, l'analyse de séquences génomiques et protéiques, la reconnaissance de la parole, l'analyse d'images....

Notre projet de Recherche s'inscrit dans l'étude d'objets complexes (images, séquences génomiques) caractérisés par de nombreuses variables (pixels dans les images, mots dans les séquences génomiques par exemple). Ces objets sont en grand nombre et, en général, une ou plusieurs expertises de haut niveau permettent de structurer les bases de données qui les regroupent. Notre but est de mettre en place des systèmes qui permettent d'avoir une vision synthétique de ces ensembles et d'en faciliter l'exploration.

### Approche méthodologique

Pour réaliser notre projet, nous avons choisi de développer une méthodologie basée sur la mise en œuvre de techniques d'apprentissage par l'exemple. Ce choix résulte de l'analyse que nous avons faite de problèmes concrets qui nous ont été soumis par des cliniciens et des biologistes. Notre approche est contextuelle, dans la mesure où les connaissances nécessaires pour traiter chaque problème sont acquises de

10. Worlds for Assessing Human Reasoning in Process Supervision

manière spécifique pour chaque situation et adaptative car les systèmes que nous construisons s'ajustent par apprentissage, éventuellement de manière dynamique, aux données à analyser. Les méthodes dites neuromimétiques, en conjonction avec des méthodes statistiques classiques et quelques outils de traitement de signal plus particuliers, constituent nos moyens d'exploration et d'analyse privilégiés.

### Applications

Quantification de la progression virale dans les rétinoopathies chez les patients atteints du SIDA

*Objet d'étude* : zones texturées dans des angiographies numériques du fond de l'œil.

*Problématique* : segmenter les images et étudier les caractéristiques de la frontière qui sépare zones saines et infectées.

*Participants, Soutien* : D. Brahmi, N. Cassoux, C. Serruys, A. Giron, B. Fertil (CRI CNRS), - Pr P. LeHoang (Service d'Ophthalmologie de la Pitié Salpêtrière), ANRS, Sidaction

L'enjeu du projet est de mettre au point une méthode permettant, par la comparaison automatique d'angiographies, de quantifier l'avancée du front caractéristique de la nécrose active. Il s'agit d'évaluer l'efficacité du traitement et de dépister précocement la survenue de récurrence pour préserver autant que possible la rétine saine. Sur les angiographies numérisées, la progression de la rétinite est évaluée par l'extension, sous la forme d'un front de nécrose, de lésions préexistantes ou par l'apparition de nouveaux foyers. À l'aide d'images réalisées à intervalles réguliers, on espère pouvoir évaluer finement l'avancée de la rétinite. L'approche proposée s'appuie sur une méthode qui apprend à reconnaître les caractéristiques des zones de rétine saines et infectées à partir de segmentations réalisées par les ophtalmologistes. Le système est en phase d'évaluation au service d'Ophthalmologie de la Pitié Salpêtrière où un grand nombre de patients atteints de CMV sont traités en routine.

Brahmi D, Serruys C, Cassoux N,

Giron A, Triller R, LeHoang P, Fertil B. « Extraction of features from medical images using a two-stage approach which combines supervised and unsupervised learning » SPIE medical imaging, 2000; 3979:265-272.

### Classification automatique des tumeurs noires de la peau (détection des mélanomes)

*Objet d'étude* : Formes, textures, coloration spécifique dans des images scannées de tumeurs noires.

*Problématique* : Reconnaître les signes de malignité

*Participants, Soutien* : C. Serruys, D. Brahmi, A. Giron, B. Fertil - Dr. R. Triller (Centre International de Dermatologie, Levallois-Perret), Dr G. Aitken, Dr. JP Vignali (Hôpital Louis Mourier, Colombes) - Ligne Nationale contre le Cancer, Société Française de Dermatologie

Il s'agit de développer et d'évaluer une méthodologie adaptative de traitement d'images pouvant être intégrée à une stratégie diagnostique ou de dépistage des mélanomes. Le système doit déterminer la typologie de la tumeur qui lui sera présentée et proposer un diagnostic motivé. Nous mettons en place une approche qui doit permettre d'éviter certaines phases critiques de l'analyse classique. Le système de classification des tumeurs noires est fondé sur la reconnaissance automatique des signes spécifiques de malignité tels qu'ils sont décrits par les dermatologues. Toutefois, il n'est pas nécessaire de proposer une description paramétrée des différentes formes, textures et couleurs que peuvent prendre ces signes. Le système élabore, de façon autonome, sa propre "vision" à partir des exemples fournis par les dermatologues au cours d'une expertise préalable de la base d'apprentissage. Une base d'images de 1800 tumeurs noires, dont une partie a été expertisée par 5 dermatologues, a été constituée. Un module de segmentation des tumeurs noires et un prototype capable de reconnaître des textures caractéristiques ont été construits.

Fertil B, Serruys C, Brahmi D, Cassoux N, Triller R, LeHoang P and Giron A. « A versatile learning by

sample strategy to perform Statistical Pattern classification in biomedical images ». SPIE electronic Imaging, 2001 (in press).

### Quantification et diagnostic en imagerie TEMP cérébrale :

*Objet d'étude* : Images isotopiques multidimensionnelles.

*Problématique méthodologique* : Amélioration de la qualité des images, rehaussement spécifique de structures et de phénomènes d'intérêt.

*Participants, Soutien* : P. Maksud, MO. Habert, G. El Fakhri, B. Fertil - Pr. SC Moore (Brigham and Women's Hospital et Harvard Medical School, Boston)

Les images obtenues à la suite d'examen en tomographie par émission monophotonique (TEMP) sont affectées par divers phénomènes physiques qui en limitent l'exploitation. Il existe des méthodes de correction des images isotopiques, mais aucune ne peut être considérée, à ce jour, comme vraiment satisfaisante dans une perspective de quantification des distributions d'activités. Nous nous proposons de mettre en place une stratégie optimale de traitement de l'image isotopique appliquée à des examens réalisés chez des patients, d'une part en imagerie de perfusion dans une pathologie débutante, d'autre part en imagerie de neurotransmission, dans la recherche d'une réponse thérapeutique. Pour cette étude, il est souhaitable de pouvoir observer de faibles variations d'activité en scintigraphie cérébrale de perfusion et de neurotransmission. Nous avons validé l'approche mise en place pour la correction de la diffusion (absorption photoélectrique, diffusion Compton et Rayleigh) en imagerie planaire et tomographique ainsi que pour la séparation des acquisitions multi-isotopiques, par des simulations de Monte Carlo de modèles réalistes de cœur et de cerveau. La méthode adaptative a ainsi permis de réaliser une quantification précise de l'activité de chaque isotope.

El Fakhri G, Moore SC, Maksud P, Aurengo A, Kijewski MF: « Absolute activity quantitation in simultaneous 123I/99Tc brain SPECT ». J Nucl Med, 2001; (in press).

**Etude de l'organisation des génomes**

*Objet d'étude* : fragments d'ADN ou génomes entiers.

*Problématique* : trouver des invariants structurant les séquences pour mettre en évidence la spécificité d'espèce.

*Participants, Soutien* : P. Deschavanne, C. Dufraigne, C. Chapus, A. Giron, B. Fertil - Pr. Edwards (Burke Museum, University of Washington, Seattle)

La mise à disposition récente de génomes complets permet maintenant de travailler sur l'organisation globale des séquences d'ADN. Les régularités tout comme les particularités de ces séquences sont susceptibles d'être utilisées pour comparer les génomes entre eux et obtenir des informations interprétables en termes d'évolution, de systématique moléculaire. Ainsi, l'analyse de la fréquence, le long du génome, de courtes chaînes de nucléotides (des « mots »), permet d'observer des similitudes et des singularités propres à chaque espèce. L'ensemble des fréquences des mots de différentes longueurs peut être considéré comme une signature génomique susceptible d'être visualisée sous forme d'image paramétrique.

L'objectif du projet est d'étudier l'utilisation des oligonucléotides dans les séquences génomiques de différents organismes dans un but de caractérisation et classification. Trois aspects complémentaires seront particulièrement développés :

- Identification de l'origine de fragments d'ADN. On trouvera des applications directes dans la reconnaissance d'organismes rares ou pour lesquels des problèmes techniques empêchent une caractérisation rapide (micro-organismes à croissance lente par exemple) ainsi que dans le domaine contrôle qualité pour vérifier en temps réel la production des séquenceurs.
- Analyse de l'homogénéité du génome. On s'intéressera en particulier aux zones du génome dont la signature s'écarte significativement de la signature globale. Ces zones peuvent provenir en effet de transferts horizontaux pour lesquelles l'espèce

d'origine pourra être recherchée, ou correspondre à des régions de fonctionnalités intéressantes.

- Phylogénie moléculaire. Nous avons observé que « distance » entre signatures génomiques et distance phylogénétique étaient partiellement liées. Nous chercherons à déterminer plus précisément cette relation en définissant une métrique plus appropriée à la mesure des relations phylogéniques sur la base de simulations de l'évolution de molécules d'ADN au cours du temps.

Deschavanne P, Giron A, Vilain J, Fagot G, Fertil B: « Genomic signature : characterization and classification of species assessed by chaos game representation of sequences. » *Mol Biol Evol*, 1999, 16(10): 1391-1399

**ARTICLES COURTS**

**Contrôle « centré-médecin »  
de la contextualisation  
des données du patient :  
une approche documentaire  
de l'aide à la décision  
médicale**

Brigitte Séroussi et Jacques Bouaud  
MRSTIM/DPA/DSI/AP-HP  
SIM. 91, bd de l'Hôpital  
75634 Paris cedex 13  
{bs,jb}@biomath.jussieu.fr

Les premiers systèmes d'aide à la décision médicale (SADMs) reposaient sur des approches statistiques. Ils ont fait la preuve de leur expertise comme l'atteste l'étude de De Dombal menée à Leeds il y a près de 30 ans sur les douleurs aiguës de l'abdomen et qui conduisait à plus de 90 % de bons diagnostics. Néanmoins, ces systèmes n'ont jamais réussi à s'imposer dans les cabinets médicaux ou les consultations hospitalières, le facteur limitant leur diffusion étant le côté abrupt d'une décision quantitativement justifiée qui rendait difficile une appropriation du système par les utilisateurs.

La seconde vague de SADMs a été constituée par les systèmes à base de connaissances initialement qualifiés de systèmes experts. Forts des enseignements de leurs prédécesseurs dits « numériques », ils proposent une approche « symbolique », propice à la justification qualitative des raisonnements mis en oeuvre au niveau de l'implémentation (inférences). Pourtant, s'ils ont tout d'abord séduit les médecins qui rêvaient de systèmes « magiques » que la sémantique du terme « Intelligence Artificielle » leur laissait supposer, ils ont finalement déçu en pratique par leurs performances moyennes dès lors que l'on sortait de cas jouets.

Plus récemment, la crédibilité de l'avis d'expert étant largement controversée, la médecine factuelle, ou evidence-based medicine, est venue proposer une méthodologie scientifiquement fondée (méta-analyses) basée sur l'étude de la littérature du domaine et la prise en compte des résultats des essais cliniques publiés pour l'élaboration de guides de bonnes pratiques (GBPs), ou guidelines. L'expertise d'un domaine n'émane plus alors d'un individu qui bien qu'expert peut se tromper, mais elle est représentée par un ensemble de recommandations dont la valeur objective est quantifiée en termes de niveaux de preuve établis selon la qualité des études (randomisation des échantillons) dont ces recommandations sont dérivées.

Qu'ils exploitent les connaissances d'un expert, ou maintenant les connaissances d'un GBP, les SADMs se présentent comme des systèmes à base de connaissances qui visent à fournir une assistance décisionnelle en pratique quotidienne. Les méthodes et outils sous-jacents sont ceux de la théorie de la décision et de l'ingénierie des connaissances : règles situation-action, modules logiques réflexes, tables de décision, plans d'action, etc. Aujourd'hui, des langages génériques de représentation de GBPs comme par exemple PROforma ou GLIF, ont été élaborés pour représenter les différents concepts que la formalisation des soins implique. Pourtant, malgré la multiplication des systèmes de GBPs implémentés, il existe peu d'éléments objectivant le respect en pratique par les médecins des recom-

mandations proposées par ces systèmes.

Les raisons psychologiques, classiquement évoquées pour expliquer cette faible observance, sont basées sur la crainte des praticiens de perdre leur liberté de choix diagnostique et thérapeutique, et en pratique leur refus de se faire usurper la partie noble de la décision médicale, c'est-à-dire, le raisonnement. Par ailleurs, si les médecins ne suivent pas, pour un patient donné, les recommandations d'un système qui implémente les GBPs, qui sont par construction l'expression objective de l'état de l'art du domaine, c'est aussi parce ces recommandations ne correspondent pas aux conclusions auxquelles ils aboutissent à l'issue de leur propre analyse de la situation. Nous pensons que ces désaccords naissent des modalités d'implémentation de ces systèmes et de l'utilisation des modèles formels qui sont développés dans les bases de connaissances.

La médecine n'est pas une science exacte, certains disent même que c'est un art. Les connaissances médicales se transmettent à travers un ensemble de notions exprimées en langage naturel sous la forme de termes qui véhiculent une sémantique dont les contours sont délibérément vagues. Mais ce mode de communication ne pose pas de problème de compréhension aux médecins qui interprètent les messages qu'ils reçoivent en fonction d'un contexte, contexte qui dépend à la fois du patient et du médecin. Car ce qui fait la richesse du langage naturel, c'est qu'il permet une variabilité d'interprétation et une flexibilité dans la contextualisation d'une information qu'aucun modèle formel, par définition, ne peut autoriser. En effet, la formalisation des connaissances passe par une étape ontologique où l'inventaire des concepts d'un domaine est réalisé. Ces concepts ont un sens pour un contexte donné. Mais comme tous les contextes possibles ne peuvent être raisonnablement envisagés, les formalisations résultantes sont des réductions nécessaires de la connaissance d'origine. Nous pensons que la démarche diagnostique et thérapeutique du médecin s'alimente d'un ensemble de facteurs contextuels non formalisables d'envergure bien plus large que l'ensemble des concepts formalisés qui existent dans

une base de connaissances. Or, le modèle de la démarche diagnostique et thérapeutique reproduit par un système informatique s'alimente, dans les approches traditionnelles de l'aide à la décision, d'un ensemble de données décontextualisées décrivant de façon réductrice, donc simpliste, l'état du patient sous la forme d'un équivalent « dégradé » du patient réel. Ce sont ensuite les données de cette version dégradée qui subissent un traitement automatique de classification permettant d'aboutir à un équivalent formel de l'équivalent dégradé, et donc du patient réel. Ce filtre constitue un biais formel de représentation de l'état du patient puisque l'intermédiaire dégradé s'appuie sur une formalisation ne reflétant pas tout le sens des notions médicales qui doivent être interprétées en contexte pour générer des inférences pertinentes.

C'est cette problématique que nous avons abordée dans l'élaboration du système OncoDoc. Cette approche part d'un constat difficile à accepter par les acteurs d'une communauté scientifique puisqu'il établit les limites structurelles de la formalisation du monde réel, en général, et de la médecine, en particulier. La richesse de la réalité socio-psycho-physiologique d'un patient, perçue par le médecin à l'issue de sa consultation, ne pourra jamais être correctement représentée par l'ensemble, par construction réduit, des paramètres pris en compte par le modèle sous-jacent à tout système informatique. La réduction du patient réel en son dual dégradé qui fixe arbitrairement le contexte des paramètres qui le construisent ne doit pas être automatisée comme le font les SADMs classiques : le patient réel se caractérise par un état dont l'interprétation doit être contextualisée par un environnement où il n'existe pas d'équation a priori entre le tout et la somme même pondérée de ses parties. C'est au médecin que revient la tâche de contextualisation et le contrôle de la catégorisation du patient réel par le « meilleur » équivalent formel afin d'en dériver les recommandations.

Dans une approche documentaire de l'aide à la décision médicale, qui privilégie le contexte d'une situation clinique plutôt qu'un ensemble de paramètres caractérisant de façon approximative

l'état d'un patient, OncoDoc s'appuie sur la rigueur et la systématique du cadre formel d'un modèle décisionnel pour la représentation des GBPs et autorise la flexibilité d'une interprétation contextuelle des connaissances par la lecture. Sur le plan formel, la base de connaissances est encodée sous la forme d'un arbre de décision dont les branches exclusives et exhaustives garantissent l'existence d'un chemin unique pour la représentation du meilleur équivalent formel de tout patient réel. Sur le plan informel, le système est développé de manière à proposer une documentation structurée de l'état de l'art au médecin qui par sa navigation hypertextuelle à travers la base de connaissances va instancier les paramètres patient et contrôler ainsi l'opérationnalisation des connaissances et la construction du meilleur équivalent formel duquel les recommandations thérapeutiques seront dérivées. Cette approche respecte l'expertise clinique du médecin puisque c'est lui qui choisit les modalités décrivant au mieux l'état de son patient plutôt qu'un traitement automatique des informations à travers le filtre de la formalisation et la production d'un patient réel dégradé approximatif. Cette approche, en évitant le biais formel, propose également d'aboutir au meilleur équivalent formel, et donc aux meilleures recommandations thérapeutiques ce qui permet d'améliorer la qualité des suggestions émises par le système, qui mieux ciblées sont mieux acceptées par les médecins. Cette approche, actuellement appliquée au traitement du cancer du sein avec le système OncoDoc, a été testée par des expérimentations sur sites pilotes. Les résultats en termes d'observance (de l'ordre de 80 %) et de recrutement dans les essais thérapeutiques (+ 50 % après utilisation du système) ont été largement meilleurs que les chiffres publiés dans la littérature.

#### Bibliographie

1. OHNO-MACHADO L., GENNARI J. H., MURPHY S. H., JAIN N. L., TU S. W., OLIVER D. E., PATTISON-GORDON E., GREENES R. A., SHORTLIFFE E. H. & BARNETT G. O. (1998). The Guideline Interchange Format: A model for representing guidelines. *J Am Med Inform Assoc*, 5(4), 357-372.

2. SEROUSSI B., BOUAUD J. & ANTOINE E.-C. (2001). OncoDoc, a successful experiment of computer-supported guideline development and implementation in the treatment of breast cancer. *Artif Intell Med*, 22(1), 43–64.
3. TIERNEY W. M., OVERHAGE J. M. & MCDONALD C. J. (1996). Computerizing guidelines: Factors for success. *J Am Med Inform Assoc*, 3(suppl), 562–566.

## DES AGENTS POUR SIMULER ET DETECTER DES EPIDEMIES

Alain-Jérôme Fougères  
SeT \_ UTBM

Université de Technologie de Belfort  
Montbéliard, Belfort Technopôle, 90010  
BELFORT – France

Mél : [Alain-Jerome.Fougères@utbm.fr](mailto:Alain-Jerome.Fougères@utbm.fr)

### 1. Présentation

Pour modéliser, concevoir ou simuler des systèmes complexes dont les composants sont en fortes interactions, le paradigme multi-agents semble dorénavant admis. Ce court article présente une approche générique de la modélisation du comportement d'un système d'aide à la détection d'épidémies où la capacité de prise de décision des composants (agents coopérants [2]) peut être exploitée. Les épidémiologistes disposent d'un certain nombre d'indices qui leur permettent de différencier, sur la base d'accumulation de symptômes sur des individus, les cas isolés d'une maladie ou la conséquence d'une contagion (épidémie). Le système conçu sur la base de l'organisation de la santé publique, permet de simuler des cas de maladies (locaux ou épars) et d'établir régulièrement des rapports de diagnostics d'éventuelles épidémies.

Notre présentation est structurée comme suit : dans la section 2 nous décrivons le système de détection d'épidémies, ainsi que le système de simulation de ces épidémies. La section 3 est consacrée à la modélisation agent du système précédent, à savoir une modélisation systémique de l'organisation, facilitant la cohésion entre la structure décisionnelle et la structure physique, et

la description des éléments de l'architecture des agents cognitifs en utilisant les réseaux de Petri. Finalement, dans la section 4, nous donnons quelques conclusions sur notre approche de modélisation, puis nous évoquons les prolongations de notre travail.

### 2. Le système de détection d'épidémies

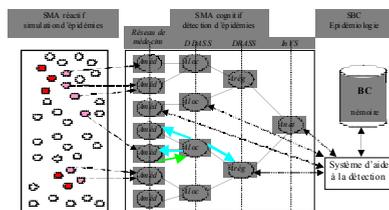


Figure 1.

L'architecture globale du système

Le système global est décrit par la figure précédente. Il est composé de 3 sous-systèmes : un SMA pour la simulation des épidémies, un SMA pour la détection d'éventuelles épidémies et un système d'aide à la décision exploitant des connaissances médicales permettant de diagnostiquer à la fois des maladies et des épidémies. Chaque catégorie d'agents du système de détection d'épidémies possède ses propres connaissances pour établir les diagnostics en relation avec les rôles qu'ils jouent au sein de l'organisation et de leur environnement. La prise de décision est ainsi répartie. Le système d'aide a donc pour réelle vocation d'assister les agents sur leur initiative, ainsi que de participer au processus de mémorisation.

#### 2.1. Le SMA de simulation d'épidémies

Ce SMA réactif, apparenté aux systèmes mis en œuvre dans les projets de vie artificielle [1,7], est composé de deux types d'agents réactifs :

- Les agents *contaminateurs* (*Acont*). Ce sont des agents d'infections, porteurs des maladies admises dans le système, notamment les maladies à déclaration obligatoire [8], qui vont être les vecteurs de la contamination. Ils disparaissent lorsque la transmission à un agent individu est effectuée.

- Les agents *individu* (*Aind*) qui peuvent être soit sains, soit contaminés. Dans ce dernier cas ils consultent leur médecin.

Un agent *contaminateur* est introduit localement par un agent de simulation qui active les maladies (grippe, méningite, coqueluche, listériose,...). Il peut, en raison de sa proximité, infecter un agent *individu* sain. Ce dernier devient porteur de la maladie et peut la véhiculer dans l'espace considéré en contaminant d'autres agents *individus*.

#### 2.2. Le SMA de détection des épidémies

Ce SMA de détection et d'alerte qui doit permettre une réaction rapide des autorités sanitaires, reprend l'organisation du système de santé public. Il comprend 4 niveaux d'agents cognitifs (acteurs), dont les rôles sont distincts :

- Les agents *médecins* (*Améd*), constitués en réseaux (notamment le réseau national de médecins sentinelles), représentent le premier niveau du système de santé en relation directe avec les malades. Ils ont pour rôle de recenser des symptômes, de diagnostiquer des maladies et de recueillir les informations à communiquer aux agents DDASS.
- Les agents *DDASS* (*Aloc et Ains*, Direction Départementale d'Action Sanitaire et Sociale) ont des fonctions de consultation, de conseil, de contrôle et de communication avec les médecins et les laboratoires. Leur mission est de fournir une aide au diagnostic pour les médecins, une remontée d'information aux instances régionales ou nationales et une mise en application des directives nationales pour la lutte ou la prévention de maladies.
- Les agents *DRASS* (*Areg*, Direction Régionale d'Action Sanitaire et Sociale) sont des agents centralisateurs (recueil d'informations régionales). Ils transmettent des informations à l'agent InVS ou des décisions sanitaires aux agents DDASS. Ce niveau a cependant un rôle limité dans l'organisation, car il est bien souvent court-circuité.
- Les agents *nationaux* dont l'agent *InVS* (*Anat*, Institut de Veille Sanitaire, l'organisme national de surveillance de l'état de santé de la population) collectent des informations

pour assurer un rôle de contrôle/décision sur le territoire national ou établir des conclusions/rapports statistiques ou qualitatifs sur des phénomènes épidémiologiques.

Chaque agent est susceptible de prendre une décision localement et peut se faire assister par le système d'aide à la détection.

Nous réfléchissons en ce moment à l'extension de cette organisation aux autres acteurs du système sanitaire impliqués dans la chaîne de lutte contre les maladies (laboratoires et notamment le Centre National de Référence, Instituts Pasteur, services de surveillance de l'eau,...).

### 3. La modélisation du système

#### 3.1. Une modélisation systémique

Nous proposons dans cette section de définir une stratégie d'organisation de systèmes complexes centrée sur la notion de réactivité. Cette nécessité de réactivité provient d'une nouvelle vision des organisations dans lesquelles les acteurs se voient augmenter leur degré d'autonomie et de flexibilité (distribution de la prise de décision et rapprochement entre la structure décisionnelle et la structure physique). L'autonomie des acteurs d'une organisation permet alors de développer localement des processus d'adaptation efficaces. La modélisation systémique proposée par Jean-Louis Hermine dans [3] a inspiré notre réflexion. Elle reprend le schéma classique OID (Opérations, Information-mémorisation et Décision) auquel elle intègre un quatrième système permettant la circulation des connaissances.

#### 3.2. La modélisation des agents cognitifs

##### 3.2.1. Eléments de modélisation

Notions de SMA réactif :

Def *SMAreactif* ::= <Agents, Environnement> [1]

Def *AgentRéactif* ::= <Perception, Interprétation, Réaction>

Notions de SMA cognitif : Def *SMAcognitif* ::= <Agents, Environnement, Interactions, Organisation> [6]

Def *AgentCognitif* ::= <Perception, Communication, Raisonnement, Actions, Connaissances>

Le principal intérêt des SMA est qu'ils permettent de distribuer des agents, entités communicantes, autonomes, réactives et dotées de compétences. Pour réaliser un SMA selon ces critères, il faut doter chaque agent, des trois propriétés suivantes :

- Autonomie : un agent doit posséder son propre ensemble de ressources et de connaissances.
- Communication : des agents qui collaborent doivent pouvoir communiquer pour s'échanger des informations.
- Intelligence : les agents poursuivent des buts en fonction de leurs compétences, ils doivent donc posséder des connaissances et des mécanismes pour raisonner avec ces connaissances.

Nous avons proposé une architecture générale d'un agent cognitif respectant les trois propriétés d'indépendance, de communication et d'intelligence. Celle-ci (figure 2) est composée de cinq modules gérant la mémoire, la perception, la communication, le contrôle et le raisonnement de l'agent.

- Le module de perception (*Mper*) : C'est le processus permettant l'acquisition des connaissances sur l'environnement (données et variables) dans lequel évolue l'agent.
- Le module de communication (*Mcom*) : c'est le mécanisme d'interactions de l'agent cognitif avec la communauté d'agents. Il reçoit des messages, les interprète et peut en transmettre sur décision du module de contrôle. Pour cela il utilise un protocole de communication basé sur le langage KQML (Knowledge Query and Manipulation Language).
- Le module de contrôle (*Mctl*) : il gère l'ensemble des modules de l'agent cognitif à l'aide de connaissances sur la synchronisation et le contrôle des tâches internes (par exemple l'analyse de messages), et des tâches externes concernant la collaboration avec d'autres agents. Le traitement d'informations ou de messages reçus se décompose en quatre phases : une formalisation de l'information, la détermination des tâches à réaliser, la transmission des actions possibles pour résoudre ces tâches au module de raisonnement et la transmission de la réponse

(décision) au module de communication.

- Le module de décision (*Mdec*) : c'est le processus qui permet à l'agent de raisonner pour prendre des décisions de manière autonome. En fonction des connaissances contenues dans le module cognitif (compétences, intentions, rationalités, croyances, accointances) et du type de décisions à prendre, le module de décision infère les séquences d'actions qu'il doit transmettre au module de contrôle.
- Le module de mémoire (*Mmém*) : ce module est caractérisé par des accointances (connaissances sur les autres agents du système), des compétences (ensemble de connaissances sur les règles de fonctionnement et sur l'état du système), ainsi que des intentions/rationalités qui correspondent aux motivations personnelles de l'agent (intentions) et aux modes d'évaluation (rationalités).

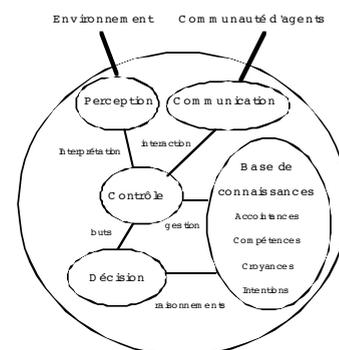


Figure 2. Architecture modulaire d'un agent cognitif

#### 3.2.2. Un modèle réseau de Petri pour l'agent cognitif

Pour faciliter la modélisation et l'analyse de systèmes complexes, nous avons développé sous forme de réseau de Petri un ensemble de 5 modules génériques, permettant d'établir le modèle des agents cognitifs présenté ci-dessus. La figure 3 présente le réseau de Petri complémentaire aux cinq précédents, à savoir le cycle de fonctionnement d'un agent cognitif. Un cycle alterne deux modes sur sollicitation périodique du module de contrôle : la poursuite d'activités selon le but que s'est fixé l'agent, ou l'acquisition d'informations par l'observation de l'environnement (module de perception) et par la consultation de

sa boîte aux lettres (module de communication).

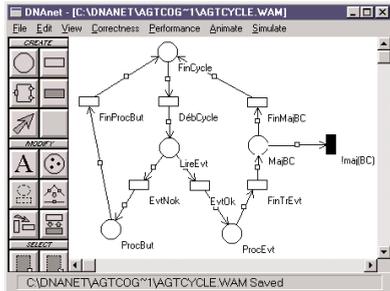


Figure 3.  
Cycle de base d'un agent cognitif.

#### 4. Conclusion

Nous venons de présenter un cadre générique de modélisation d'agents cognitifs définis dans le but de répartir la prise de décisions dans des systèmes complexes, pour en améliorer les performances. L'approche formelle corrélée consiste à définir une architecture modulaire pour les agents cognitifs et à utiliser des réseaux de Petri pour spécifier leurs comportements [4]. Les avantages majeurs qui ressortent de l'utilisation des réseaux de Petri sont, d'une part, la possibilité de donner des spécifications à la fois formelles et graphiques de nos agents, et d'autre part, la possibilité de simuler le système avant son exécution, puis d'évaluer une partie de ces propriétés. Nous avons illustré notre approche sur la définition d'un système de simulation et de détection d'épidémies qui allie complexité (3 sous-systèmes de conception distincte à intégrer) et clarté de présentation. L'élaboration d'une méthodologie de définition formelle de tels systèmes intelligents, dédiés à la modélisation et à la simulation de systèmes complexes, constitue nos perspectives de recherche.

#### Bibliographie

- [1] Drogoul A. and Hutzler G., Le jardin des hasards : peinture abstraite et I.A.D. réactive, 1996, Rapport de recherche LAFORIA, 96/04.  
[2] Ferber J., 1997. Les systèmes multi-agents : un aperçu général, *Technique et Science Informatiques*, 16(8), 979-1012.  
[3] Hermine J.-L., 2000. *Les systèmes de*

*connaissances*, 2<sup>e</sup> éditions, Hermès Science Publications, Paris.

- [4] Holvoet T., 1995. Agents and Petri Nets, in O. Herzog, W. Reisig, and R. Valk, editors, *Petri Net Newsletters*, 3, number 49.  
[5] Jennings N.R. and Wooldridge M.J., 1998. *Applications of intelligent agents*, in Springer-Verlag (Ed.), *Agent Technology: Foundations, Applications and Markets*.  
[6] Ocello M., Koning J.-L. and Baeijs C., 2001. Conception de systèmes multi-agents : quelques éléments de réflexion méthodologiques, *Technique et Science Informatiques*, 20(2), 233-263.  
[7] Quinqueton J. and Hamadi Y., 1999. Communication et Emergence : une épidémie chez les termites, *JFIADSMSA'99*.  
[8] Note de la direction générale de la santé sur la transmission de données individuelles à l'autorité sanitaire, BEH, *Bulletin épidémiologique hebdomadaire*, n° 47, 23 novembre 19

#### Introduction

Pour un utilisateur, toute application informatique est l'intégration de plusieurs éléments cognitifs : les connaissances des experts du domaine, les connaissances informatiques qui président au choix des représentations des connaissances sous contraintes de l'état de l'art de la technologie, les savoir-faire présumés des utilisateurs pour mettre en œuvre ces représentations par les outils de la technologie.

Le projet I.C.R.E.M. (Informatisation des Compte Rendus d'Examens Médicaux) (chef du projet Bernard Huet) de l'équipe *Métafor* (animateur Gilles Blain) du groupe *Objets et Agents pour les Systèmes d'information et la Simulation (O.A.S.I.S.)* du lip6 étudie l'intégration des éléments cognitifs énoncés ci-dessus dans l'informatisation d'une composante du « dossier médical » (DM): le recueil, l'archivage, l'interprétation et la transmission des résultats d'examen médico-techniques (gastroscopie, échocardiographie, imagerie par résonance magnétique nucléaire,...). Ce projet nous a permis de faire émerger divers problèmes cognitifs.

Au sens le plus extensif, le DM voudrait recouvrir tous les événements de santé d'un individu de sa naissance à sa mort. Dans la pratique le DM est un état partiel résultat d'un recueil limité dans le temps et dans l'espace ; même en se limitant au seul aspect hospitalier le dossier médical est considérablement polymorphe car il est le résultat de la contribution de plusieurs médecins différenciant par leur point de vue : chirurgie, médecine, biologie, leur formation, leur spécificité, leur degré d'expertise,...

La représentation des examens médico-techniques est généralement exprimée dans le cadre d'un consensus entre experts très spécialisés constituant une « micro-communauté » à l'intérieur d'un domaine, (par exemple, en gastro-entérologie le pôle d'endoscopie avec la gastroscopie, la fibroscopie, la recto-sigmoïdoscopie,...). Ce consensus est caractérisé par l'existence d'une intersection cognitive très forte entre les différents experts.

Polymorphisme des experts versus monomorphisme des outils informatiques actuels

#### Projet ICREM Risques cognitifs liés à la représentation actuelle des connaissances dans les applications médicales

<sup>1,2</sup>Huet B., <sup>1</sup>Blain G.,

<sup>1</sup>Laboratoire d'Informatique de Paris 6 (lip6) CNRS UMR 7606, Université P. et M. Curie, Paris, France

<sup>2</sup>Université Paris XIII, Hôpital Avicenne (Assistance Publique – Hôpitaux de Paris), Santé Publique, Biostatistiques, Informatique Médicale, Bobigny, France  
*bernard.huet@lip6.fr*  
*gilles.blain@lip6.fr*

#### Résumé

Cet article vise à montrer que l'écart entre les outils informatiques monomorphiques disponibles et les outils polymorphiques dont les experts ont besoin engendre un risque cognitif. L'enjeu de la réduction de ce risque est une maximisation de la transmission de l'expertise ; la Méta-Modélisation est l'une des solutions possibles pour réduire cet écart.

La médecine est exercée sur des patients, c'est-à-dire des systèmes complexes en interactions dont on ne connaît que peu d'informations sur leur état au moment de prendre une décision (la décision sous incertitude de résultats caractérise la décision thérapeutique). Il en résulte que l'exercice de la médecine est dominé par le mode de l'expertise.

La diversité des experts (chacun a ses connaissances propres, ses expériences, ses cas rencontrés) et des spécialités aboutit à un polymorphisme des approches diagnostiques et thérapeutiques. Il en résulte que les outils employés dans l'élaboration des diagnostics et des décisions thérapeutiques doivent être eux-mêmes polymorphes pour être vraiment bien adaptés à chaque expert ; ce polymorphisme est encore accru par l'évolution des connaissances du domaine, des approches et des outils au fil du temps. Le « dossier médical », outil de base du médecin est de ce fait un *concept intrinsèquement générique* qui s'instancie dans une multitude de formes plus ou moins complexes.

En face de ce besoin d'outils polymorphes, force est de constater que les outils informatiques actuels de gestion de l'information médicale sont très monomorphes.

Ce monomorphisme se traduit par les aspects suivants : la majorité des applications sont bâties autour d'une représentation figée d'une partie du vecteur d'état censé renseigner sur l'état du patient, de plus les modalités d'usage des outils actuels de la mise à jour d'une telle structure renforcent ce monomorphisme : ainsi, la dénomination des rubriques d'un tel vecteur (par exemple le document de compte-rendu d'examen d'échocardiographie, de gastroscopie...) est fixée au niveau du document type d'un service clinique et non au niveau de l'expert utilisateur; il en est de même pour la définition des limites des rubriques par rapport au vecteur d'état du patient, pour l'affectation du corpus des termes médicaux, pour la gestion des ontologies, pour la structuration des unités significatives pour l'expert. Enfin, l'unicité et la rigidité du déroulement opératoire proposé dans les applicatifs contribuent au monomorphisme.

### Le risque consécutif à la non-adaptation des logiciels actuels vis-à-vis des experts

Chaque expert a ses spécificités du fait que chacun a une expérience unique et en partie non-transmissible : les cas rencontrés, les connaissances académiques, les expériences phénoménologiques tout cela aboutit à un *contexte cognitif propre à chaque expert*, au sein d'un pôle d'expertise spécialisé, avec des ramifications et des extensions cognitives profondément personnelles.

Ce *contexte cognitif* propre à chaque expert peut se traduire par un supplément interprétatif spécifique de l'expert, ce supplément interprétatif est fondé sur le fait que dans une micro-communauté d'experts, il y a une capacité interprétative commune à tous les experts (fournissant le noyau commun interprétatif d'un compte-rendu) ou *généralité cognitive* de la micro-communauté et une capacité interprétative complémentaire spécifique propre à chaque expert (celle-ci est plus ou moins développée selon que l'on ait affaire à un expert *débutant, confirmé, senior*), c'est la spécificité cognitive de chaque expert.

Le monomorphisme des logiciels actuels conduit à une faible captation des arborescences les plus fines de la *spécificité cognitive* de chaque expert.

Le *risque* est donc d'appauvrir les modèles des experts (le risque est d'autant plus élevé que l'expert est un *senior*) en incitant ceux-ci à abandonner tout ou partie de leur spécificité cognitive (notamment les arborescences les plus fines) pour les inciter à s'aligner plus ou moins sur la *généralité cognitive*.

Il s'agit là d'une démarche *dynamique d'appauvrissement cognitif*.

Il est clair que la capacité d'expertise d'un expert est indépendante, en soi, de la technologie, mais la disponibilité d'un outil capable de prendre seulement en compte une faible partie des arborescences les plus fines d'expert *senior* est un facteur qui ne facilite pas pour ces derniers la gestion et la navigation dans leurs connaissances; tout ceci contribue à diminuer et non à développer la *spécificité cognitive* d'un expert : pourquoi se lancer dans l'interprétation d'arborescences fines si l'on n'est pas possible de les

exprimer dans le cadre des outils de la pratique courante ?

Par ailleurs, le processus d'enrichissement du corpus de la *généralité cognitive* par intégration des spécificités cognitives s'effectue le plus souvent par intervention de « Sociétés Savantes ». Chaque *spécificité cognitive* parvenant à la micro-communauté est intégrée à la *généralité cognitive* après évaluation positive par des experts de la micro-communauté. La diffusion de la *généralité cognitive complétée* est ensuite assurée par la « micro-communauté ». La « remontée » des *spécificités cognitives* vers les « Sociétés Savantes » est aujourd'hui très partielle, ce qui constitue un facteur limitant entraînant une *minimisation de la dynamique du processus d'enrichissement de la généralité cognitive*.

*L'enjeu de la réduction du risque cognitif est donc au niveau individuel une maximisation de la transmission de l'expertise et au niveau d'une communauté d'experts l'accélération de la dynamique du processus d'enrichissement de la généralité cognitive.*

### Comment les nouvelles techniques de développement des applicatifs peuvent-elles faire face à ce défi ?

Le problème réel est de fournir à chaque micro-communauté d'utilisateur-experts les outils leur permettant de maximiser l'exploitation de la *généralité cognitive* de la micro-communauté et de la *spécificité cognitive* des experts. Pour cela il faudrait pouvoir intégrer économiquement dans les applicatifs la très grande variété possible des *contextes cognitifs* liés à leur usage ; or les techniques classiques ne le permettent pas, par ailleurs il n'est pas possible de faire un inventaire exhaustif de ces *contextes cognitifs* permettant leur typage.

Les techniques de Méta-Modélisation doivent pouvoir desserrer l'étau du monomorphisme des applicatifs actuels en travaillant sur la conception de modèles (structure, données, ontologies,...) *génériques* et non sur la conception de modèles (structure, données, ontologies,...) *spécifiques*; telle est l'*idée clé qui est à la base de la recherche en Méta-Modélisation*.

Alors qu'un modèle de structure de données spécifique répond au modèle d'un expert ou au modèle consensuel d'un groupe d'experts ; *un modèle de structure de données générique* est potentiellement capable (sous réserve de disposer des outils logiciels ad hoc) de permettre la génération d'un grand nombre de modèles de structure spécifiques, ce qui permet de répondre à la variété des contextes cognitifs des experts ; car les modèles des experts d'une même communauté médicale (par exemple en échocardiographie, gastroscopie,...) sont généralement très voisins, en sorte qu'il y a souvent une grande variété de modèles très voisins.

Le projet *I.C.R.E.M.* du *lip6* est basé sur la méta-modélisation des compte rendus médicaux techniques (échocardiographie, gastroscopie, fibroscopie,...). Une première version du méta-modèle (mmv0) est en cours d'évaluation dans un service hospitalier de cardiologie, l'élément logiciel servant de base à ce Méta-modèle est l'environnement de recherche et développement logiciel *Méta-Gen* conçu et développé par l'équipe *Métafor / O.A.S.I.S. / lip6*. Les outils logiciels actuels orientés vers la méta-modélisation, outre *Méta-Gen*, sont *ProtégéII* (Musen, Stanford), et *Méta-Edit* (Lyytinen, Finlande). Il a été noté récemment des spécifications pour l'industrie du logiciel dans le cadre du document « Meta Object Facility ».

### Deux avantages attendus de la méta-modélisation en médecine

D'une part, un avantage concret pour le patient, puisque la possibilité de maximiser la *spécificité cognitive* d'un expert *senior* peut conduire ce dernier à poser un diagnostic rare qui ne l'aura pas été par d'autres experts dont la *spécificité cognitive* est plus pauvre.

D'autre part un avantage cognitif pour la micro-communauté d'experts, car la captation *massive* des arborescences les plus fines aura pour conséquence d'accélérer la *dynamique du processus d'enrichissement de la généralité cognitive* de chaque micro-communauté d'experts. De proche en proche la *généralité cognitive* s'enrichira des spécificités cognitives pour constituer un corpus de connaissances transmissibles permettant de mettre à la disposition de tous les experts l'état de l'art des connaissances d'une micro-communauté d'experts dans les meilleurs délais.

*La dynamique d'appauvrissement cognitif des experts à travers le monomorphisme des applications figées agit comme un facteur limitant la dynamique du processus d'enrichissement de la généralité cognitive d'une micro-communauté d'experts.*

La Méta-modélisation en redonnant toute leur responsabilité aux utilisateurs-experts dans la réalisation finale de l'application devrait permettre au contraire une dynamique d'enrichissement cognitif individuel agissant comme un accélérateur de la dynamique

du processus d'enrichissement de la généralité cognitive .

Bernard Huet et Gilles Blain remercient le professeur J.L. Pourriat (Professeur des Universités, U.P. VI – Praticien Hospitalier, Hôtel-Dieu, Assistance Publique – Hôpitaux de Paris) pour sa contribution lors de la relecture de cet article.

L'équipe *Métafor* (animateur Gilles Blain) du groupe *Objets et Agents pour les Systèmes d'information et la Simulation (O.A.S.I.S.)* du *lip6* a construit un environnement de recherche et développement « MétaGen » basé sur les techniques de « Méta-Modélisation » et de transformation de programmes. Cette plate-forme a déjà servi de générateur d'applications dans divers domaines (banque, assurance, œnologie,...) et dans un projet Européen (FIBOF), celle-ci est l'un des éléments du projet *IMPACT* (RNNTL 2002) ; dans l'article ci-dessous, le domaine médical est illustré à travers le projet *I.C.R.E.M.* (Informatisation des Comptes Rendus d'Examens Médicaux) (chef du projet Bernard Huet).

### Références

1. Huet B., Artigou J.Y., Poirier J., Blain G.,
2. Meta-Modelling; the solution for a family of applications, *Proc. Xth World Conf. Med. Inform. MEDINFO 2001*, London, G.B., Sep. 2 - 5, 2001 (Ed. Patel V.L., Rogers R., Haux R) IOS press Inc., Amsterdam, NL, 623 - 629.

## LIVRES



### Vous avez *aimé* un *livre*

Vous voulez faire partager votre intérêt, deux options sont possibles.

Vous pouvez en faire une critique d'une ou deux pages et l'envoyer à Amadeo Napoli.

Mais vous pouvez aussi lui 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.

## COMPTE RENDU DE CONFÉRENCES

### Compte rendu de la conférence EA'01 Evolution Artificielle

*Le Creusot les 29-30-31 Octobre 2001  
Centre Condorcet  
Université de Bourgogne,*

*<http://www.cmap.polytechnique.fr/~ea01/home.html>  
Marc Schoenauer et Michèle Sebag (membres du bureau de l'AFIA)*

La conférence Évolution Artificielle 01 (EA'01) s'est tenue du 29 au 31 octobre 2001 au Creusot, au Centre Condorcet, antenne délocalisée de l'Université de Bourgogne.

La qualité du programme scientifique a été assurée par une sélection sévère des communications (28 papiers présentés sur 67 soumissions) : le Comité d'Organisation tenait d'une part à ne pas sombrer dans le mercantilisme à tout prix, d'autre part à prolonger la « tradition » (forte de maintenant 5 éditions) de séances plénières (pas de sessions parallèles). Tous les papiers acceptés ont ainsi pu être présentés en 30mn, laissant largement le temps

pour des questions. De nombreux exposés ont ainsi été l'origine de véritables débats, qui se sont poursuivis lors des pauses-café et autres déjeuners.

Le Comité de programme a décerné le prix AFIA, récompensant le meilleur papier, à Rodolphe Leriche et Frédéric Guyon (CNRS et Université de Clermont-Ferrand) pour leur article « Dual Evolutionary Optimisation ». Rodolphe Leriche a ainsi reçu une inscription gratuite à ECAI 2002 (juillet 2002 à Lyon), offerte par l'AFIA, ainsi qu'un magnum de Givry, offert par l'association Evolution Artificielle. Un prix du meilleur papier d'application a également été attribué à

Pierre-yves Oudeye pour « Origins and Learnability of Syllable Systems : a Cultural Evolutionary Model ». Enfin, un jury « populaire », composé de cinq participants volontaires choisis lors de la séance inaugurale, a décerné le prix de la meilleure présentation à Ingo LaTendresse pour sa présentation du papier « The Effects of Partial Restarts in Evolutionary Search ». Ces prix ont été solennellement remis lors de la séance de clôture de la conférence.

Côté participation, 69 personnes s'étaient inscrites, dont 47 français, 11 européens non français, et 11 non européens. Mais 9 des non-européens n'ont pas osé affronter les transports aériens en ces temps troublés, même alors que leur papier avait été accepté.

Mais la série de conférences Évolution Artificielle (1994 à Toulouse, 1995 à Brest, 1997 à Nîmes et 1999 à Dunkerque) est également réputée pour la qualité des prestations extra-scientifiques ! Et cette spécificité n'est sans doute pas étrangère à la participation relativement importante de la communauté internationale, même en ces temps où le nombre de conférences du domaine évolutionnaire explose - sans parler des tristes événements internationaux récents.

EA'01 n'a pas dérogé à la règle, avec un programme social évidemment tourné vers le vignoble et la bonne chère : visite d'une cave à Givry, gracieusement offerte et

commentée par le viticulteur en personne, excellent dîner où nos amis anglo-saxons ont dégusté (ou refusé de déguster), pour la plupart pour la première fois, de véritables escargots de Bourgogne, mais également à signaler les trois déjeuners de haute tenue dont le moindre aurait fait pâlir d'envie bien des banquets de conférences internationales que nous ne nommerons pas... Signalons enfin que tout ceci était offert à un prix défiant toute concurrence (130 Euros incluant les déjeuners, l'après-midi de visite et le banquet) grâce aux efforts sans faille de Pierre et Valérie Collet, tant au niveau financier, pour récupérer des subventions locales et nationales, qu'au niveau pratique, Valérie et lui assurant avec efficacité toutes ces petites tâches d'organisation peu spectaculaires, certes, mais indispensables au déroulement sans heurt d'un tel événement.

Signalons enfin, parmi les atouts locaux, la situation géographique, réunissant les avantages de la province et un accès très facile grâce au TGV (1h30 depuis Paris, 2h30 depuis Lille, 3h depuis Marseille).

Le programme scientifique, ainsi qu'un échantillon de photos prises lors de la conférence, sont disponibles à l'URL

<http://www.cmap.polytechnique.fr/~ea01/home.html>.

## Compte rendu de la conférence JF IAD SMA '01 Journées Francophones Intelligence Artificielle Distribuée et Systèmes Multi-Agents

*Montréal, les 12-13-14 Novembre 2001*  
**CRIM (Centre de Recherche Informatique de Montréal)**  
[https://www.crim.ca/jfiadsma\\_2001](https://www.crim.ca/jfiadsma_2001)  
**Amal Elfallah Seghrouchni, Sylvie Pesty**

La recherche dans le domaine des systèmes multi-agents (SMA) prend de plus en plus d'ampleur, en particulier grâce à l'émergence dans le domaine de nouvelles technologies de systèmes à base d'agents logiciels. Ces agents offrent des modèles de coopération et de coordi-

nation qui font leur succès. Ils constituent également un nouveau paradigme de développement logiciel d'applications distribuées ouvertes et adaptables. Depuis 9 ans, les JF IAD SMA, c'est-à-dire les Journées Francophones pour l'Intelligence Artificielle Distribuée (IAD) et les

Systèmes Multi-Agents (SMA) sont une occasion de rencontres et d'échanges entre chercheurs francophones de ce domaine. Pour cette neuvième édition, 87 participants ont été accueillis chaleureusement par Laurent Magnin et l'équipe du Centre de Recherche Informatique de Montréal (CRIM).

Le thème principal de la conférence était cette année « Fondements des Systèmes Multi-Agents : Modèles, Spécifications formelles et Vérification ». Après les traditionnels mots de bienvenue des organisateurs, une première conférence invitée, donnée par Pierre-Yves Schobbens (Université de Namur – Belgique) sur le thème de la vérification automatique de logiques orientées agent, a ouvert la conférence. La première journée a été consacrée au thème « Modèles formels pour la coopération dans les SMA et pour leur conception ». Ont été présentés durant cette journée des mécanismes de coopération tels que la négociation multi-agents, la formation de coalition, la révision des croyances et l'extraction de comportement à partir d'interactions, ou encore des méthodes formelles pour la spécification des interactions multi-agents, une sémantique formelle pour la composition et la substitution de composants dans les SMA ainsi que la conception de SMA orientée patterns.

La seconde journée a débuté par la deuxième conférence invitée de Sylvain Giroux (Université de Sherbrooke, Canada), intitulée « mobilité, personnalisation et information géo-référencée », qui a ainsi amorcé le thème « Ouverture, distribution et mobilité dans les SMA ». Ce thème a adressé le problème de la gestion distribuée et de la mobilité des agents, de l'ouverture dans

les SMA ainsi que le problème de l'allocation et de l'exécution distribuée de tâches dans les SMA.

La troisième et dernière journée a été quant à elle dédiée aux présentations relevant d'approches pluridisciplinaires ou décrivant des applications réalisées selon le paradigme SMA. Enfin, d'autres aspects techniques de conception et/ou de mise en œuvre ont été présentés également à travers la session posters et quelques démonstrations.

Cette année, le collège SMA de l'AFIA et les organisateurs ont alloué huit bourses à de jeunes chercheurs afin de couvrir leurs frais d'inscription. Un « prix AFIA » du meilleur article a également été décerné. Pour cela, trois articles ont été tout d'abord nominés, à savoir :

- « Extraction des comportements à partir des interactions locales. » - Nicolas sabouret et Jean-Paul Sansonnet du LIMSI - Orsay
- « Modélisation des satisfactions personnelles et interactives d'agents situés coopératifs. » - Olivier Simonin et Jacques Ferber du LIRMM - Montpellier
- « La spécification de protocoles d'interaction par réseaux de Petri » - Christophe Sibertin-Blanc, Janette Cardoso et Chihab Hanachi de l'IRIT et du CERISS – Toulouse

C'est finalement la contribution d'Olivier Simonin et Jacques Ferber qui a été choisie et c'est Olivier qui, ayant fait la présentation, a reçu le prix sous forme d'une inscription gratuite à ECAI 2002. Bravo !

## Compte rendu de la conférence VALGO 2001 Journées Jeunes Chercheurs en Connexionisme

**Larnas (Ardèche), les 17-18-19 Octobre 2001**

<http://www.supelec-rennes.fr/acth/JourneesValgo/journeesvalgo.html>

**David Mercier**

**Président de l'ACTH (Association des Connexionnistes en Thèse)**

<http://www.supelec-rennes.fr/acth>

Les journées « Valgo » ont pour objectif de réunir les jeunes chercheurs (doctorants) francophones travaillant dans le domaine du connexionisme afin qu'ils partagent leurs expériences mais aussi qu'ils aient une idée plus précise des différentes approches existantes et des divers débouchés. Elles sont une occasion pour chacun de s'exercer à synthétiser et présenter ses travaux devant une assemblée critique mais amicale.

Ces journées entrent dans le cadre de "l'Opération Doctorants" du GDR-PRC ISIS (Information, Signal, Images, viSion) ce qui permet également aux doctorants appartenant à des laboratoires membres du GDR-PRC ISIS, d'avoir leur frais de déplacement pris en charge. La revue en ligne Valgo (ISSN: 1625-9661) sert d'actes de ces journées et permet ainsi une bonne diffusion des travaux présentés ou envisagés.

Du mercredi 17 au vendredi 19 octobre 2001 s'est donc déroulée sur la commune de Larnas, au domaine d'Imbours, l'édition 2001 des Journées Valgo.

Arrivés dans le courant du mercredi après-midi, nous avons commencé par profiter du soleil et du splendide paysage de l'Ardèche sur les kilomètres séparant Montélimar et le domaine d'Imbours !

Le lendemain, la journée a débuté par les présentations de Vincent Martin sur « Recherche de dynamiques neuronales pour la vision en robotique » et d'Emilie Poisson sur « Réseaux de neurones à convolution appliqués à la reconnaissance de l'écriture manuscrite non contrainte ». Les questions et remarques qui suivirent chaque exposé furent nombreuses et nous menèrent à midi, heure à laquelle Sylvie Pesty, notre correspondante AFIA arriva. Après une discussion sur l'AFIA et les Journées Valgo, Sylvie Pesty a remis le prix pour le meilleur article des Journées à Emilie Poisson, sous la forme d'une inscription gratuite à ECAI 2002 (15th European Conference on Artificial Intelligence).

Après le déjeuner et une courte détente, la deuxième séance de travail a commencé par la présentation d'Anne-Sophie Dujardin sur « Pertinence d'une approche hybride multi-neuronale dans la résolution de problèmes liés au diagnostic industriel ou médical » et a continué avec celle de David Mercier sur « Utilisation des STANN en audio et son illustration en reconnaissance de chiffres ». Après les questions et remarques sur ces exposés et un court en-cas, nous avons poursuivi par une séance de travail dans laquelle chacun a pu présenter les problèmes ou difficultés qu'il rencontre dans son travail de thèse.

La séance du vendredi matin s'est ouverte par une présentation de Michaël Aupetit, en tant qu'orateur libre, sur l'ensemble de sa thèse « Auto-organisation, coopération et apprentissage dans les réseaux connexionnistes » suivie de la présentation de Bruno Scherrer sur « L'auto-organisation modulaire d'une architecture intelligente ».

Voici donc le bref résumé de ces journées qui se sont passées dans un très bon état d'esprit et auquel il ne manque finalement que les remerciements : tout d'abord à tous les membres du conseil d'administration de l'ACTH qui ont co-organisé ces journées avec moi, c'est-à-dire Michaël Aupetit, Christophe Foucher, Christelle Gaudin, Sébastien Marcel et Jean-Denis Muller; ensuite à nos « partenaires » et leurs correspondantes, c'est-à-dire Elisabeth Depré pour le GDR-PRC ISIS et Sylvie Pesty pour l'AFIA; également à tous les relecteurs qui nous ont permis de sélectionner les articles, et à leurs auteurs de les améliorer, c'est-à-dire les docteurs et professeurs Frédéric Alexandre (LORIA), Daniel Collobert (FranceTelecom), Patrick Gallinari (LIP6), Pierre-Yves Glorennec (IRISA), Yves Granvalet (UTC), Kurosh Madani (I2S), Sébastien Marcel (IDIAP), Daniel Memmi (IMAG) et Jean-Denis Muller (CEA); enfin à tous les participants de ces Jour Bravo !

## Compte rendu de la conférence Reconnaissance des Formes et Intelligence Artificielle RFIA 2002

*Angers, 7-10 janvier 2002*

*Catherine garbay et Roger Mohr,  
Co-présidents du congrès RFIA 2002*

RFIA 2002, treizième édition du congrès francophone AFRIF-AFIA de Reconnaissance des Formes et d'Intelligence Artificielle, s'est du 7 au 10 janvier 2002 au centre de congrès d'Angers.

Ce congrès est l'occasion de rencontres particulièrement fructueuses entre les communautés de la Reconnaissance des Formes et de l'Intelligence Artificielle, dont la tradition est maintenant bien établie. Il constitue pourtant un point singulier, dans le paysage scientifique européen et international, dont il faut rappeler l'importance.

En effet, face à l'essor grandissant des sciences du traitement de l'information, il nous paraît plus que jamais nécessaire de mettre en avant les spécificités de ces deux disciplines et leurs complémentarités. L'évolution des supports informationnels et des moyens de calcul et de communication offre des possibilités accrues de communiquer par la parole et par l'image, et suscite le développement d'outils pour les analyser, les combiner, les interpréter et les visualiser ; dans le même temps, la diversification des usages pousse le développement de nouveaux modèles pour l'extraction et la formalisation des connaissances et du raisonnement, et pour la résolution de problèmes.

De nouveaux enjeux s'ouvrent ainsi, qui impliquent notre capacité à créer des synergies entre le monde du numérique et celui du symbolique ; le congrès RFIA 2002 est précisément l'occasion d'affronter ces enjeux . Vous avez été nombreux à répondre à l'appel à communication, et nous tenons à vous remercier tous pour l'intérêt que vous avez ainsi manifesté pour ce congrès.

La sélection s'est avérée difficile, étant donné le grand nombre d'articles de qualité que nous avons reçus

: sur les 210 soumissions, seulement la moitié a pu être retenue. Ce travail de sélection est un travail délicat, et nous avons été assisté dans cette tâche par deux comités : le comité éditorial et le comité de programme. Le rôle du comité de programme, qui comporte 44 membres répartis de manière égales entre nos deux communautés, est de procéder à l'évaluation des papiers soumis, en faisant éventuellement appel à des relecteurs extérieurs. Chaque papier a ainsi été évalué par 2 relecteurs indépendants, selon les critères habituels d'originalité, d'importance de la contribution, de qualité scientifique et technique, et de qualité de la rédaction. Nous avons souhaité qu'à cette exigence de qualité soit associée une volonté d'ouverture à la diversité des écoles de pensée (travaux à la frontière entre plusieurs domaines, ou de natures diverses : théoriques, méthodologiques, expérimentaux...). Le congrès RFIA est en effet l'occasion d'affronter cette diversité. Le comité éditorial, qui comporte 16 membres répartis de manière égales entre nos deux communautés, est de superviser l'ensemble du processus de sélection. A ce titre, il participe à la constitution du comité de programme, lui attribue les articles soumis pour évaluation en les affectant de manière individuelle, enfin participe aux décisions finales concernant la sélection des articles et à la mise en place du programme. Il joue un rôle de première importance dans le processus d'évaluation, en fournissant un avis supplémentaire en cas de conflit entre les relecteurs, et en rédigeant une fiche de synthèse utilisée lors de la décision finale. Il veille ainsi à la rigueur du processus d'évaluation. Un travail remarquable a été effectué, tant par les membres du comité de programme que par le comité éditorial. Qu'ils soient tous chaleureusement remerciés pour ce travail difficile.

Ce travail d'évaluation nous a permis de remarquer l'excellence de plusieurs contributions : deux d'entre

elles ont été sélectionnées, l'une issue de la communauté RF, l'autre issue de la communauté IA, et un prix leur a été décerné par l'AFIA et par l'AFRIF.

Un programme de qualité a ainsi été élaboré, qui comprenait plus de trente sessions. Une formule originale a été mise en place pour la présentation des communications, qui offrait la possibilité d'associer présentation orale, en session, et présentation sous forme de poster et/ou de démonstration. L'intérêt d'une telle formule était de permettre à la fois une présentation formelle, classique, des éléments principaux du travail, et une présentation plus détaillée et aussi plus interactive de certains aspects. Une session particulière a été dédiée à ces présentations supplémentaires, qui a été un succès, malgré le faible nombre de présentations démonstrations/posters proposées. Nous espérons qu'une telle formule pourra être renouvelée dans les prochaines éditions RFIA, et nous vous incitons à y contribuer plus nombreux.

Des tutoriaux avaient également été organisés, la veille du congrès, afin de permettre l'approfondissement de grands thèmes d'actualité, et nous remercions vivement tous les collègues qui se sont mobilisés pour organiser ces tutoriels. Un seul d'entre eux a pu avoir lieu, du fait du faible nombre d'inscrits. Là aussi, cet état de fait est regrettable, et nous vous invitons à

répondre plus nombreux à l'avenir à ce type de proposition. Enfin, 4 conférenciers invités nous font le plaisir de participer à cet événement et nous en les remercions très vivement.

Un tel travail n'aurait pas été possible sans le soutien du comité d'organisation, grâce à qui nous nous sommes réunis à Angers : par la mise en place d'un site web associant informations pratiques et scientifiques et gestion des soumissions, par le suivi constant du processus d'évaluation dans toutes ces phases, ils ont soutenu l'ensemble du travail effectué par le comité de programme, et nous tenons à les remercier très chaleureusement pour cet investissement sans faille. Le congrès a été grâce à eux un grand succès, et nous avons particulièrement apprécié la qualité de l'accueil, la convivialité des locaux et la qualité de l'environnement technique, ainsi que l'excellence des repas et du dîner de gala qui s'est tenu dans un lieu prestigieux (le Musée de Lurçat) dans une ambiance musicale très appréciée.

Nous espérons que ce congrès a été pour vous aussi enrichissant qu'utile, par la richesse de son programme scientifique et par les opportunités de rencontres et d'échanges qu'il est dans sa tradition d'offrir, et nous espérons que vous participerez nombreux à la prochaine édition de ce congrès.

# CALENDRIER

## Les Conférences AFIA en 2002

8 –10 Janvier	<b>RFIA</b>	Angers	Reconnaissance des Formes et Intelligence Artificielle
21 – 23 Janvier	<b>EGC</b>	Montpellier	Extraction et Gestion des Connaissances
28 – 30 Mai	<b>IC</b>	Rouen	Ingénierie des Connaissances
21 – 26 Juillet	<b>ECAI</b>	Lyon	15th European Conference on Artificial Intelligence
24 – 26 Juillet	<b>PAIS</b>	Lyon	Prestigious Applications of Intelligent Systems
22 – 23 Juillet	<b>STAIRS</b>	Lyon	First European Starting AI Researcher Symposium
1 –5 Juillet	<b>IPMU</b>	Annecy	9th Information Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems

### Calendrier des conférences en IA en 2002

2-4 Janvier	<b>AMAI'2002</b>	Fort Lauderdale (USA)	Artificial Intelligence and Mathematics
8-10 Janvier	<b>RFIA'2002 (AFIA)</b>	Angers	Reconnaissance des Formes et IA
15-18 janvier	<b>NF'2002</b>	La Capitolio de la Habana (Cuba)	Neuro Fuzzy
19-20 janvier	<b>PADL'2002</b>	Portland (USA)	Practical Aspects of Declarative Languages
21-23 janvier	<b>EGC 2002 (AFIA)</b>	Montpellier	Extraction et Gestion des Connaissances
10-14 Mars	<b>SAC'2002</b>	Madrid (Spain)	Coordination Models, Languages and Applications
25-27 Mars	<b>IAS'2002</b>	Marina del Rey (USA)	Intelligent Autonomous Systems
2-5 Avril	<b>AISB 2002</b>	London (UK)	AI and the Simulation of Behaviour
7-10 Avril	<b>AIS 2002</b>	Lisbon (Portugal)	AI, Simulation and Planning
8-11 Avril	<b>COORDINATION 2002</b>	York (United Kingdom)	Coordination Models and Languages
16-18 Mai	<b>FLAIRS 2002</b>	Pensacola (USA)	Artificial Intelligence
27-29 Mai	<b>AI 2002</b>	Calgary (Canada)	Artificial Intelligence
28-30 Mai	<b>IC 2002 (AFIA)</b>	Rouen	Ingénierie des connaissances
4-7 Juin	<b>INFORSID 2002</b>	Nantes	Systèmes d'informations
10-12 Juin	<b>ISWC 2002</b>	Sardinia (Italy)	Semantic Web
10-12 Juin	<b>DSV-IS 2002</b>	Rostock (Germany)	Design, Specification and Verification of Interactive Systems
17-19 Juin	<b>CAp 2002</b>	Orléans	Apprentissage
27-29 Juin	<b>ISMIS 2002</b>	Lyon	Methodologies for Intelligent Systems
1-5 Juillet	<b>IPMU'2002 (AFIA)</b>	Annecy	Knowledge-Based Systems
1-5 Juillet	<b>AISC'2002</b>	Marseille	Artificial Intelligence and Symbolic Computation
3-6 Juillet	<b>CIAA'2002</b>	Tours	Implementation and Application of Automata
8-12 Juillet	<b>ICML'2002</b>	Sydney (Australia)	Machine Learning
9-13 Juillet	<b>GECCO 2002</b>	New York (USA)	Genetic and Evolutionary Computation
14-18 Juillet	<b>SCT'2002</b>	Orlando (USA)	Systemics, Cybernetics and Informatics
15-19 Juillet	<b>AAMAS'2002</b>	Bologna (Italy)	Autonomous Agents and MultiAgent Systems
21-26 Juillet	<b>ECAI'2002 (AFIA)</b>	Lyon	Intelligence Artificielle
24-26 Juillet	<b>PAIS'2002 (AFIA)</b>	Lyon	Prestigious Applications of Intelligent Systems
22-23 Juillet	<b>STAIRS'2002 (AFIA)</b>	Lyon	STarting Artificial Intelligence Researchers Symposium
28 Juillet-1 Août	<b>AAAI'2002</b>	Edmonton (Canada)	Intelligence Artificielle
24 Août-1 Septembre	<b>Coling 2002</b>	Taipei (China)	Computational Linguistics
25-30 Août	<b>IFIP'2002</b>	Montreal (Canada)	Information Processing
25-30 Août	<b>IIP'2002</b>	Montreal (Canada)	Intelligent Information Processing
16-18 Septembre	<b>KES 2002</b>	Podere d'Ombriano (Italy)	Knowledge-Based Intelligent Information Engineering Systems & Allied Technologies
1-4 Octobre	<b>EKAW 2002</b>	Siguneza (Spain)	Knowledge Engineering and Knowledge Management
18-22 Novembre	<b>ICONIP 2002</b>	Singapore (Singapore)	Neural Information Processing
18-22 Novembre	<b>SEAL 2002</b>	Singapore (Singapore)	Simulated Evolution And Learning
18-22 Novembre	<b>FSKD 2002</b>	Singapore (Singapore)	Fuzzy Systems and Knowledge Discovery

# CALENDRIER

## Calendrier des Workshops ECAI

Workshops ECAI : <a href="http://ecai2002.univ-lyon1.fr/show_fr.pl">http://ecai2002.univ-lyon1.fr/show_fr.pl</a>	Date de soumission
<b>Lundi 22 juillet</b>	
W1 Ontologies and Semantic Interoperability	15-mars
W4 Configuration	8-mars
W5 Knowledge Management and Organizational Memories	15-mars
W6 Artificial Intelligence in Mobile Systems	22-mars
W10 Computational Complexity in Artificial Intelligence	25-février
W12 Knowledge Discovery from Temporal and Spatial Data	25-mars
W13 Agents Applied in Health Care	31-mars
W15 Innovative Software Engineering Techniques	25-mars
W18 Computational Models of Natural Argument	1-avril
W19 E-alliances : Process Support Tools for Wide Area Inter-organisational Alliances	22-avril
<b>Mardi 23 juillet</b>	
W2 Advances in Formal Concept Analysis for Knowledge Discovery in Database	28-avril
W7 Knowledge Transformation for the Semantic Web	15-avril
W8 Spatial and Temporal Reasoning	15-mars
W9 Modelling and Solving Problems with Constraints	6-mai
W11 Causal Networks from Inference to Data Mining	24-mars
W14 Binding Environmental Sciences and Artificial Intelligence	15-mars
W20 Agent Technologies in Logistics	31-mars
W21 Intelligent Data Analysis in Medicine and Pharmacology	22-mars
W22 Semantic Authoring, Annotation and Knowledge Markup	28-avril
<b>Lundi 22 juillet et Mardi 23 juillet</b>	
W3 Model Checking and Artificial Intelligence	1-avril
W16 Machine Learning and Natural Language Processing for Ontology Engineering	15-mars
W17 Creative Systems: Approaches to Creativity in AI and Cognitive Science	5-avril

## Appel à communication

### FIRST EUROPEAN STARTING AI RESEARCHER SYMPOSIUM STAIRS-2002

*Affiliated to ECAI'2002, Lyon, July 22-23, 2002*

#### AIMS AND SCOPE

STAIRS'2002 is the first European Starting AI Researcher Symposium, an international meeting aimed at AI researchers, from all countries, at the beginning of their career: PhD students or people holding a PhD for less than one year. It offers them :

1. a first experience on submitting and presenting a paper in an international forum with a broad scope and thorough selection process;
2. the opportunity to gather knowledge and exchange ideas related to their research problems and approaches, and to attend ECAI at a reduced cost;
3. information on European research careers and mobility.

STAIRS'2002 will be held in Lyon, in the period immediately preceding ECAI'2002.

#### IMPORTANT DATES

Title and abstract due to the PC co-chairs:

March 1, 2002

Paper submission deadline: March 6, 2002

Notification of acceptance sent to authors:

April 19, 2002

Camera-ready copies due to the PC co-chairs:

May 6, 2002

Symposium venue: July 22-23, 2002

## TOPICS

Submissions are invited on substantial, original, and previously unpublished research in all fields of Artificial Intelligence, including, but not limited to:

- Adaptive Systems
- Automated Reasoning
- Autonomous Agents
- Case-Based Reasoning
- Causal and Probabilistic Reasoning
- Cognitive Science in AI
- Constraints
- Data Mining and Information Retrieval
- Description Logics and Conceptual Graphs
- Diagnosis and Abduction
- Distributed AI
- Genetic Algorithms and Soft Computing
- Intelligent User Interfaces
- Intelligent Web Applications
- Knowledge Representation and Reasoning
- Logic Programming
- Machine Learning
- Model-based and Qualitative reasoning
- Multi-Agent Systems
- Natural Language Processing and Speech Recognition
- Neural Networks
- Non monotonic Reasoning
- Ontology
- Planning and Scheduling
- Real-time AI and Control
- Reasoning about Actions and Change
- Robotics
- Search
- Spatial and Temporal Reasoning
- Uncertainty in AI
- Vision

Additionally, we welcome contributions addressing practical results and lessons learned from the integration of artificial intelligence techniques into industrial applications, either completed or in progress. Such papers should be clearly identified as « Application paper » and will be treated separately by the Program Committee.

## SUBMISSION GUIDELINES

Submission is a two-stage process. Authors are first asked to submit a summary of their research paper by March 1st, 2002, by sending it to [stairs@dis.uniro-ma1.it](mailto:stairs@dis.uniro-ma1.it), including the title, the names of all authors with full affiliations and email addresses, the contact author, and an abstract of the paper (maximum 200 words), plus keywords drawn from the above list of topics (plus other keywords if appropriate). Please do not forget to add « Application paper » when relevant.

The full paper should be sent electronically to the same address by March 6th, 2002, in either PostScript or PDF format (the latter being highly preferred). Free services to convert a number of widely used file formats to PDF are available at <http://createpdf.adobe.com/> and at <http://wheel.compose.cs.cmu.edu:8001/cgi-bin/browse>. If electronic submission is problematic, please contact both PC co-chairs by February 15 (addresses are below). It is highly recommended to submit papers using the final

camera-ready formatting style (including authors and affiliations on the first page). Submissions must not exceed six pages in camera-ready format, or 10 pages in plain article format (such as the Springer LNCS format). Over length submissions will be rejected without review. Guidelines on the format of submissions and style files will be available on the STAIRS'02 webpage <http://stairs2002.univ-lyon1.fr/>. The final version of accepted papers will be required to conform strictly to these guidelines, and will be allocated six pages in the proceedings.

Papers received after March 6th 2002 (23:59 CET) will not be reviewed. Notification of receipt of full papers will be emailed to the corresponding author soon after receipt.

## REVIEW PROCESS

Papers should address significant research works undertaken by the authors. At least one of the authors must be a PhD student, or a person holding a PhD for less than a year. Preference will be given to papers whose authors are all in such condition.

All submissions will be subject to academic peer review by the STAIRS'2002 Program Committee. Each paper will be assigned at least two reviewers. Review criteria include originality of ideas, technical soundness, significance of results, and quality of presentation.

Notification of acceptance or rejection of submitted papers will be emailed to the corresponding author by April 19, 2002.

## CONFERENCE PROCEEDINGS

The conference proceedings will be published and distributed by IOS Press. The authors will be responsible for producing camera-ready copies of papers, conforming to the STAIRS'2002 formatting guidelines, for inclusion in the proceedings. The deadline for receipt of the camera-ready copy is May 6, 2002.

For each accepted paper, at least one author, and most preferably one of the beginner researchers involved in the work, is required to attend the conference to present the paper.

## AWARDS

At least one best paper of the conference will be selected by the STAIRS'2002 Program Committee, and announced during the Gala Dinner on the evening of July the 22nd. The authors of this paper, as well as those of other papers selected by the Program Committee, will be invited to submit a longer version for a fast review process for publication in the AI Com journal.

## PROGRAM CO-CHAIRS

Thierry Vidal  
ENIT Tarbes,  
47, avenue d'Azereix - BP 1629  
F-65016 Tarbes cedex, France  
Tel: +33-5-62442762  
Email: thierry@enit.fr

Paolo Liberatore  
Dipartimento di Informatica e Sistemistica,  
University of Roma La Sapienza,  
Via Salaria 113, I-00198, Roma, Italy  
Tel: +39-0649918488  
Email: liberato@dis.uniroma1.it

## PROGRAM COMMITTEE

The PC consists of 39 junior established researchers, most of them having completed their PhD in the past decade, and either European or working in a European research team.

Liliana Ardissono, University of Turin, Italy  
Carlos Areces, University of Amsterdam, The Netherlands  
Alessandro Artale, UMIST University of Manchester, United Kingdom  
Bernhard Beckert, University of Karlsruhe, Germany  
Gianluca Bontempi, Interuniversitair Micro-Elektronica Centrum, Belgium  
Jean-Cédric Chappelier, Swiss Federal Institute of Technology in Lausanne, Switzerland  
George Coghill, University of Wales, UK  
Silvia Coradeschi, Örebro University, Sweden  
Carlos Damásio, Universidade Nova de Lisboa, Portugal  
Khalid Daoudi, INRIA Lorraine, France  
Nadja De Carolis, University of Bari, Italy  
Alvaro del Val, Universidad Autónoma de Madrid, Spain  
Luc de Raedt, Albert-Ludwigs University, Freiburg, Germany

Yannis Dimopoulos, University of Cyprus, Cyprus  
Alexis Drogoul, University Paris 6, France  
Florence Dupin de Saint-Cyr, Université d'Angers, France  
Dieter Fox, University of Washington, USA  
Jose Manuel Gutierrez, University of Cantabria, Spain  
Ivana Kruijff-Korbayová, Universität des Saarlandes, Germany  
Patrick Lambrix, Linköpings universitet, Sweden  
Vincenzo Lombardo, Università del Piemonte Orientale « A. Avogadro », Italy  
Vittorio Maniezzo, Università di Bologna, Italy  
Dunja Mladenic, Jozef Stefan Institute Ljubljana, Slovenia  
Abdel-Ilah Mouaddib, University of Caen, France  
Bart Netten, Delft University of Technology, The Netherlands  
Barry O'Sullivan, University College Cork, Ireland  
Chris Reed, University of Dundee, UK  
Jochen Renz, Vienna University of Technology, Austria  
Jussi Rintanen, Albert-Ludwigs University, Freiburg, Germany  
Florence Sellini, Airbus, UK  
Barry Smyth, University College Dublin, Ireland  
Steffen Staab, University of Karlsruhe, Germany  
Csaba Szepesvári, Mindmaker Ltd, Hungary  
Hans Tompits, Vienna University of Technology, Austria  
Ioannis Tsamardinos, University of Pittsburgh, USA  
Sofie Verbaeten, Katholieke Universiteit Leuven, Belgium  
Rob Vingerhoeds, Siemens, Toulouse, France  
Lluís Vila, Technical University of Catalonia, Spain  
Pinar Öztürk, Norwegian Univ. of Science and Technology Thronheim, Norway

## ORGANIZATION COMMITTEE

Nathalie Guin-Duclosson  
Laboratoire d'Ingénierie des Systèmes d'Information (LISI)  
Université Claude Bernard - Lyon 1  
Nautibus (Bâtiment 710), 8 Boulevard Niels Bohr  
Domaine scientifique de la Doua,  
69622 Villeurbanne Cedex, France  
Email : [stairs2002@bat710.univ-lyon1.fr](mailto:stairs2002@bat710.univ-lyon1.fr)

with Olivier Aubert, Pierre-Antoine Champin, Stéphanie Jean-Daubias, Serge Fenet, Sandra Nogry, Simone Pimont, Arnaud Stuber, Olivier Teytaud

## Appel à communication

# 13èmes Journées d'Ingénierie des Connaissances IC-2002

**Rouen, 28-30 mai 2002**

<http://ic2002.insa-rouen.fr>

### Historique

Placée sous l'égide du GRACQ (Groupe de Recherche en Acquisition des Connaissances, <http://www.irit.fr/GRACQ>), la conférence IC est un lieu d'échange et de réflexion de la communauté française sur les problématiques de l'Ingénierie des connaissances. Elle fait suite aux conférences IC 2001 (plate forme AFIA, Grenoble), IC 2000 (IRIT, Toulouse), IC'99 (École Polytechnique, Palaiseau), IC'98 (Nancy), IC'97 (Roscoff) et aux Journées d'Acquisition des Connaissances (JAC) qui se sont tenues tous les ans de 1990 à 1996 (Lannion, Sète, Dourdan, Saint-Raphaël, Strasbourg, Grenoble, Sète). Elle se déroulera les 28-29-30 mai 2002 à Rouen.

### L'Ingénierie des connaissances

De manière générale, l'Ingénierie des connaissances porte sur l'instrumentation technique des contenus pour les exploiter dans un cadre où ils sont mobilisés pour leur signification. Par conséquent, l'Ingénierie des connaissances porte sur la représentation informatique du sens pour que les systèmes construits puissent avoir leur comportement paramétré par cette représentation. Le sens peut être le sens d'une activité (comment le système s'adapte à ce qui fait sens pour l'utilisateur à un moment donné), d'un document (comment le système traite le document en fonction de ce qu'il signifie), etc. Plus précisément, l'Ingénierie des connaissances propose des concepts, méthodes et techniques permettant de d'acquérir, de modéliser et de formaliser des connaissances pour les mobiliser dans l'activité d'une organisation. Structurées, formalisées et opérationnalisées pour être intégrées dans le fonctionnement d'un système à base de connaissances (SBC), ces connaissances deviennent des informations destinées à être, in fine, interprétées par un humain dans son interaction avec l'artefact construit. La réussite d'une démarche d'ingénierie des connaissances peut alors se mesurer par rapport à 3 objectifs :

- Construire un SBC ayant un comportement appréhendable et compréhensible par l'utilisateur. Cela passe par une modélisation à un niveau d'abstraction pertinent qui fait sens pour les différents acteurs impliqués

(ingénieur cogniticiens, experts métiers, utilisateurs, etc.) et leur permet de s'appropriier le comportement du système et d'interagir avec lui. C'est le niveau des connaissances ».

- Appréhender le fonctionnement et l'utilisation d'un SBC par rapport à son intégration dans une organisation. L'ingénierie des connaissances doit alors permettre de modéliser et d'explicitier le fonctionnement cognitif de l'utilisateur ou du collectif dans sa relation au SBC.
- Développer et mettre en place une réelle démarche d'ingénierie, reproductible et proposant des instrumentations techniques, i.e. des méthodes et des outils fondées sur des approches rigoureuses de modélisation des connaissances et de formalisation logique, mathématique et informatique.

L'IC s'associe par conséquent à de nombreuses disciplines :

- d'une part, dans sa démarche d'ingénierie, l'IC mobilise les concepts et techniques de la représentation des connaissances, les méthodes d'analyse et de conception à objets, le raisonnement à base de cas, l'ingénierie documentaire ou l'ingénierie éducative, la conception de systèmes d'information, etc. ;
- d'autre part, dans sa démarche de modélisation des connaissances, l'IC doit se rapprocher de disciplines permettant de caractériser et décrire les connaissances d'un domaine et d'évaluer leur mise en oeuvre dans les SBC. La sociologie, la gestion ou l'ergonomie peuvent ainsi concourir à une démarche d'explicitation de ce que sont les connaissances dans un contexte humain et organisationnel.

### Thématiques de l'Ingénierie des connaissances

Ce rôle central et pluridisciplinaire de l'Ingénierie des connaissances se retrouve dans les thèmes privilégiés (mais non exhaustifs) dans l'édition 2002 de sa conférence annuelle :

#### *Méthodes de résolution de problèmes, ontologies et langages de spécification de connaissances*

Différents types de langages permettent de représenter les connaissances issues des modèles de raisonnement (Méthodes de résolution de problèmes) ou des

modèles du domaine (Ontologies). Ces langages sont généralement dédiés à la modélisation, la représentation ou l'opérationnalisation des connaissances dénotant l'expertise visée (connaissances ontologiques, connaissances de contrôle, etc.). Les contributions peuvent en particulier aborder l'analyse,

la définition ou l'exploitation de langages de représentation des connaissances, et la définition, exploitation et réutilisation de modèles de raisonnement ou d'ontologies.

### *Approches cognitives et activité des opérateurs*

Ce thème s'intéresse aux modélisations s'appuyant sur une analyse des systèmes cognitifs naturels en situation. Il part de l'hypothèse qu'il est nécessaire d'analyser de façon approfondie les pratiques de l'utilisateur qu'on se propose d'assister et de cerner son fonctionnement cognitif pour concevoir des systèmes d'aide adaptés. On cherche à susciter l'exposé de travaux de modélisation des connaissances basés sur une analyse de l'activité réelle des opérateurs, que celle-ci soit vue à l'échelle de l'activité individuelle ou à l'échelle de l'activité collective au sein d'une organisation. Les modélisations qui seront soumises peuvent avoir été produites dans le cadre de la conception d'outils informatiques ou dans un objectif d'étude plus global. Les travaux concernant les activités de conception, notamment collective, sont particulièrement recherchés.

### *Textes et ontologies*

Les outils et méthodes d'analyse de corpus sont utiles pour la construction ou l'enrichissement de ressources lexicales et conceptuelles à partir de corpus spécialisés. Relèvent du thème les méthodes mobilisant la terminologie et l'analyse linguistique des textes pour la construction d'ontologies, les outils (outils de TALN, outils statistiques, outils d'apprentissage et de « text mining », etc.), et les ressources comme les ontologies, thesaurus, lexiques sémantiques, etc.

### *Hypertextes, hypermédia et Ingénierie des connaissances documentaires*

L'accès à l'information passe de plus en plus par l'intermédiaire de système documentaire, dont le Web sémantique est en passe de proposer un paradigme unificateur. Les contributions peuvent porter sur les modèles de représentation des connaissances (structurales et sémantiques) liés aux documents hypermédiés et aux techniques de constitution et d'exploitation de ces systèmes (annotation, navigation, hypertextes dynamiques et adaptatifs, etc.).

### *Systèmes d'Information et Ingénierie des connaissances*

Ce thème privilégie la présentation d'articles por-

tant sur les convergences entre les travaux de recherche menés en système d'information et en ingénierie des connaissances. Il porte entre autres sur les systèmes de recherche d'informations au sein de sources hétérogènes et distribuées, les systèmes favorisant la réutilisation à différents niveaux d'abstraction et de granularité (du PGI au composant logiciel), sur les aspects représentation et recherche de connaissances réutilisables.

### *Intranet/Internet et recherche d'informations*

L'Internet ou l'Intranet au sein d'une entreprise peuvent être des moyens privilégiés pour l'acquisition, la modélisation, la diffusion, le partage, l'échange et la gestion des connaissances de manière distribuée et coopérative. Les contributions pourront porter sur les bases documentaires ou de connaissances sur le Web, la recherche d'information, les technologies de type XML/RDF, ainsi que les problèmes liés à la mémoire d'entreprise et la fouille de données.

### *Ingénierie des connaissances et théorie des organisations*

L'objectif de ce thème est de susciter l'exposé de travaux qui éclairent d'une part les apports des concepts, outils et méthodes de l'IC aux chercheurs en management d'entreprise, d'autre part les apports des modélisations des gestionnaires, plus globales et permettant de « situer » les connaissances pour l'action dans leur contexte organisationnel et dans les finalités opérationnelles de l'organisation.

### *Ingénierie des connaissances pour les systèmes de travail coopératif Médiatisé*

Le but de la thématique « Ingénierie des Connaissances pour les systèmes de travail coopératif médiatisé » est de contribuer à ouvrir un espace de dialogue sur la nature des relations entre analyse des pratiques professionnelles situées, modélisation des connaissances et conception de supports technologiques aux activités coopératives. En termes de conception de systèmes, l'hypothèse de travail sous-jacente est que pour modéliser ces nouvelles applications, et donc les informations qu'elles mettront à disposition des collectifs de travail, il faut expliciter les connaissances mises en jeu dans la coopération : coordination, résolution collective de problème, dialogues...

### *Gestion des connaissances et mémoire d'entreprise*

La gestion des connaissances est une nouvelle approche de l'organisation basée sur la valorisation (au sens très large) du capital de connaissances composé des savoirs, savoir-faire, comportements, compétences etc., accumulé par une entreprise. Elle concerne les pro-

cessus de capitalisation, partage, enrichissement externe par l'intelligence économique et la veille, etc.

L'objectif du thème est de montrer en quoi l'ingénierie des connaissances est un outil puissant et un support de la gestion des connaissances.

### *Ingénierie et modélisation de connaissances à l'aide d'objets*

Les objets sont utilisés aujourd'hui pour informatiser la plupart des domaines applicatifs (les objets « métier »). Des convergences apparaissent entre le génie logiciel d'une part et l'ingénierie et la représentation de connaissances d'autre part, qui rendent nécessaire une réflexion commune. L'ingénierie des objets rapproche ces deux communautés à la fois sur le plan de la recherche et sur le plan industriel, malgré des différences de vocabulaire. On pourra aborder la construction de modèles de connaissances et d'ontologies, les modèles de savoir-faire (p. ex. « patterns »), les modèles d'architecture et leur utilisation (p. ex. « frameworks »), les techniques de transfert ou de transformation de modèles (XML, ...), etc.

### *Ingénierie éducative*

La construction d'un environnement informatique dédié à la formation est un problème complexe dont de nombreux aspects relèvent de l'IC. Les contributions pourront porter sur l'acquisition, modélisation et représentation des connaissances du domaine enseigné ou des connaissances pédagogiques et didactiques ; l'utilisation des TIC pour la médiatisation ; les méthodes d'ingénierie de construction des environnements de formation (rapport avec celles de l'IC) ; et les problèmes liées au « training industriel », la simulation et la mise en place de systèmes d'apprentissage coopératifs ou collaboratifs.

### *Raisonnement à partir de cas*

Un système de raisonnement à partir de cas (RàPC) résout un nouveau problème en s'inspirant d'un cas résolu. Face à une requête spécifiant un nouveau problème à résoudre (cas cible), le système se remémore un cas d'une mémoire de cas (cas source), réutilise la solution du cas source en l'adaptant pour

le problème cible, révisé la solution obtenue si celle-ci n'est pas satisfaisante et éventuellement mémorise le nouveau cas. A chaque étape du cycle de raisonnement une modélisation du savoir-faire des experts est nécessaire pour laquelle les techniques d'IC sont pertinentes. On s'intéressera en particulier à la conception d'un système de RèPC, aux systèmes RèPC interactifs et textuels et aux systèmes RèPC coopératifs.

### *Expériences pratiques*

Ce thème privilégie la présentation d'articles tirant des leçons d'expériences pratiques de recueil et de formalisation de connaissances allant ou non jusqu'à la réalisation de systèmes. Des descriptions et bilans d'applications seront également intéressants si on souligne les interactions entre ces applications et leurs utilisateurs.

Les auteurs choisiront de soumettre une démonstration et/ou une communication de forme courte ou longue selon leur capacité à (volonté de) se situer de manière critique par rapport aux travaux du domaine, mettre en valeur des évaluations de l'état de l'art et souligner les apports ou les manques des approches actuelles.

## Informations pratiques

### Calendrier

1 février : date limite des soumissions  
22 mars : résultat de la sélection des papiers  
26 avril : réception des papiers définitifs  
28-30 mai : IC'2002

### Soumissions attendues

Les auteurs sont invités à soumettre des propositions d'article dans les 2 catégories ci-après :

1. Communications de recherche (format A4 long : 12 p. max., interligne simple, times 12) Sont concernés des travaux de recherche originaux, théoriques ou expérimentaux. Une rigueur devra être apportée à la rédaction des articles, tant sur le plan de la forme que sur celui du contenu scientifique.
2. Communications appliquées (format A4 court : 6 p. max.) Sont concernés des applications industrielles et opérationnelles significatives, des études de cas, des prototypes avancés. Le contexte de l'application devra être clairement décrit et les auteurs mettront l'accent sur les problèmes rencontrés, les enseignements tirés et les retours d'expériences.

Les auteurs sont invités à indiquer, sur la première page de l'article, le ou les thèmes dans lesquels s'inscrivent leurs propositions. Les articles ne relevant pas de ces thèmes s'attacheront à mettre clairement en évidence leur lien avec l'ingénierie des connaissances.

Chaque proposition est relue par 3 membres du comité de programme, a priori par 2 spécialistes et un non spécialiste.

Toute communication écrite (de recherche ou appliquée) donnera lieu à un exposé oral de 30 mn pour les communications de recherche et de 15 mn pour les communications appliquées.

## Formats des soumissions

Soumissions des communications

Les auteurs sont invités à envoyer leur communication:

- sous forme électronique (format Word, PDF ou Postscript ) à l'adresse suivante : [ic2002@insa-rouen.fr](mailto:ic2002@insa-rouen.fr) - et sous forme papier à la présidente du comité d'organisation :

Catherine Barry-Gréboval

PSI

INSA de Rouen

Place Emile Blondel

BP 08

76 131 Mont Saint Aignan Cedex

Pour tout problème ou renseignement portant sur le contenu du papier contacter le président du comité de programme :

Bruno Bachimont

Directeur Scientifique

Institut National de l'Audiovisuel

4, Ave de l'europe,

94366 Bry sur Marne cedex

Courriel : [bbachimont@ina.fr](mailto:bbachimont@ina.fr)

tél. : 01 49 83 29 08

fax : 01 49 83 25 82

Pour tout problème ou renseignement portant sur l'envoi du papier, envoyer un message à [ic2002@insa-rouen.fr](mailto:ic2002@insa-rouen.fr)

## Comité de programme

B. BACHIMONT, INA/UTC, Bry sur Marne / Compiègne (président)

P. ALBERT, ILOG, Gentilly

N. AUSSENAC-GILLES, IRIT, Toulouse

C. BARRY-GREBOVAL, PSI, INSA de Rouen

J.-F. BOUJUT, INP, Grenoble

J. Bézivin, LRSG, NANTES

B. BIEBOW, LIPN, Villetaneuse

D. BOURIGAULT, ERSS, Toulouse

J. CHARLET, AP-HP, Paris

M. CHEIN, LIRMM, Montpellier

F. DARSES, CNAM, Paris

R. DIENG, INRIA Sophia-Antipolis

J.-L. ERMINE, Tech-CICO, Troyes

J.-G. GANASCIA, LIP6, Paris

N. GIRARD, INRA, Toulouse

G. GROSZ, CRI, Paris 1

M.-C. JAULENT, SIM Broussais, Paris

G. KASSEL, LaRIA, Amiens

J.-M. LABAT, CRIP5, Paris 5

Ph. LAUBLET, CAMS, Paris 4

J. LE MAITRE, SIS, Toulon

A. MILLE, LISI, Lyon 1

S. MOISAN, INRIA Sophia-Antipolis

A. NAPOLI, INRIA Lorraine, Nancy

Y. PRIE, LISI, Lyon 1

M. REVENU, GREYC, Caen

Ch. REYNAUD, LRI, Orsay

P. SALEMBIER, GRIC-IRIT, Toulouse

P. TCHOUNIKINE, LIUM, Le Mans

R. TEULIER, GRID, ENS Cachan

F. TRICHET, IRIN, Nantes

M. ZACKLAD, Tech-CICO, Troyes

## Comité d'organisation

Catherine Barry-Gréboval, PSI, INSA de Rouen (présidente)

Marie-Hélène Abel, HEUDIASYC, Compiègne

Nathalie Chaignaud, PSI, INSA de Rouen

Nicolas Delestre, PSI, INSA de Rouen

Pierre Héroux, PSI, INSA de Rouen

Jacques Labiche, PSI, INSA de Rouen

Eric Trupin, PSI, INSA de Rouen

## Appel à Communications

### International Conference on Electronic Commerce COLLECTeR (Europe) 2002

Toulouse, France April 20, 2002

<http://ebusiness.newcastle.edu.au/COLLECTeREurope>

COLLECTeR (Collaborative Electronic Commerce Technology and Research) is a joint venture by more than 25 universities around the world which, through a loose federation of research centres - one located at each institution - provides both depth and breadth of expertise in Electronic Commerce.

The COLLECTeR (Europe) Conference on Electronic Commerce 2002 will be held on April 20 2002 immediately preceding KR2002 (the Eighth International Conference on Principles of Knowledge Representation and Reasoning <http://www.kr.org/kr02>

This follows the successful colocation of COLLECTeR (USA) 2000 with KR2000 in Breckenridge, Colorado.

COLLECTeR (Europe) 2002 is intended to attract researchers interested in all aspects of Electronic Commerce and has been timed to precede KR2002 to draw interest from researchers who are particularly interested in the Semantic Web, eBusiness Agents, Knowledge Representation, Knowledge Management, and Knowledge Technologies.

Suggested research topics for COLLECTeR(Europe) 2002 include (but are not limited to):

- Knowledge Management
- ontologies for eBusiness
- eBusiness Software Agents
- eBusiness Applications for the Semantic Web
- eCommerce/eBusiness strategies and models
- Mobile Commerce
- Interactive Marketing
- online marketing strategies and issues
- Intra organisational communications
- Electronic Payment Systems
- Internetbased Electronic Commerce
- Virtual Communities/Community Networks
- Logistics Issues for Electronic Commerce
- Business Reengineering Issues for Electronic Commerce
- Government Electronic Procurement and Service Delivery
- Legal, Auditing or Security Issues for Electronic Commerce
- Privacy Issues for Electronic Commerce

- Requirements Engineering Approaches for Electronic Commerce
- eHealth
- eEducation

Researchers are encouraged to submit either « work in progress » papers of around 2,000 words, or full papers (of around 5,000 words) on completed research. Papers should be original and not previously published. All papers will be refereed by experienced reviewers with expertise in the area. Accepted papers will be published in the Conference Proceedings,

which will be available at the Conference itself. Additional copies of these Proceedings will also be available for purchase. We also expect to produce an online version of the proceedings.

The conference offers an opportunity for all those interested in researching Electronic Commerce to meet in a specialised venue to discuss research activities, findings and experiences. Attendees at the conference are encouraged to participate in discussion of all papers presented.

#### Submission Process

Send a copy of your paper in pdf format to one of the Conference

Co-Chairs:

Paul Swatman ([swatman@uni-simt.de](mailto:swatman@uni-simt.de)) or

Mary-Anne Williams ([maryanne@ebusiness.newcastle.edu.au](mailto:maryanne@ebusiness.newcastle.edu.au))

by March 8, 2002.

#### Important Dates

Deadline for Submission: March 8, 2002

Notification of Acceptance: April 1, 2002

Camera Ready Copy Due: April 10, 2002

#### Conference Organisers

Paul Swatman, Stuttgart Institute of Management and Technology, Germany

(Co-Chair)

Mary-Anne Williams, University of Newcastle, Australia (Co-Chair)

Danni Fowler, University of Baltimore, USA

Donald McCubbrey, University of Denver, USA  
Paula Swatman, University of Koblenz, Germany

## Programme Committee

Joan Cooper, University of Wollongong, Australia  
Tim Finin, University of Maryland Baltimore  
County, USA

Danni Fowler, University of Baltimore, USA  
Felix Hampe, University of Koblenz, Germany  
Daniel Le Berre, Centre de Recherche en  
Informatique de Lens, France  
Steve Elliot, University of Newcastle, Australia  
Don McCubbrey University of Denver, USA  
Siva Multhay, University of Newcastle,

AustraliaLeora Morgenstern, IBM T.J. Watson  
Research Center, USADan O'Leary, University of  
Southern CaliforniaPavlos Peppas, Patras University,  
GreeceRobert Rugimbana, University of Newcastle,  
AustraliaSuku Sinnappan, University of Newcastle,  
AustraliaPaula Swatman, University of Koblenz,  
GermanyVladimir Zwass Fairleigh Dickinson  
University, USA

Mary-Anne WilliamsDirector, Center of Electronic  
Commerce for Global BusinessFaculty of Business and  
LawThe University of NewcastleNSW 2308,  
AustraliaeFax & eVoiceMail: +1-650-745-3301  
<http://ebusiness.newcastle.edu.au/maryannemaryanne@ebusiness.newcastle.edu.au>

## Secondes Journées Francophones d'Extraction et de Gestion des Connaissances

### EGC 2002

*21 - 23 janvier 2002 – Montpellier*

Aujourd'hui, de grandes masses de données structurées ou semi-structurées sont accessibles dans les bases de données d'entreprises ainsi que sur la Toile. Aussi les entreprises ont-elles besoin de méthodes et d'outils capables d'extraire les connaissances pertinentes pour les décideurs, de les stocker, et de les diffuser aux différents acteurs de l'organisation. Pour répondre à cette attente du tissu industriel, de nombreux projets de recherche se développent autour de l'extraction de connaissances à partir de données (Knowledge Discovery in Data), ainsi que sur la gestion de connaissances (Knowledge Management).

L'objectif de ces journées consiste à rassembler, d'une part les chercheurs des disciplines connexes

(apprentissage, statistiques et analyse et données, systèmes d'information et bases de données, ingénierie des connaissances, ...), et d'autre part les spécialistes d'entreprises qui déploient les méthodes et les outils adaptés à leurs besoins, afin de contribuer à la formation d'une communauté scientifique dans le monde francophone autour de cette double thématique de l'extraction et de la gestion de connaissances.

Cette manifestation sera ainsi le lieu privilégié d'échanges entre chercheurs et spécialistes industriels afin de présenter l'état de l'art de cette thématique centrale, d'en faire émerger les besoins et les enjeux, ainsi que d'en stimuler le développement.

## The 9th Information Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems

### IPMU 2002

*1 - 5 juillet 2002, - Annecy*

The IPMU Conference is organized every two years with the focus of bringing together scientists working on methods for the management of uncertainty and aggregation of information in intelligent systems. This

conference provides a medium for the exchange of ideas between theoreticians and practitioners in these and related areas.

## Séminaire d'Intelligence Artificielle de l'Ile de France

### Séminaire IAdIF

Paris

Pour plus d'informations, consulter le site du séminaire (IAdIF).

### Appel à communications

## CAp 2002 - Conférence d'Apprentissage

Orléans 17-19 juin 2002

<http://www.univ-orleans.fr/CONFERENCES/CAp2002/>

Date limite de soumission : 21 février 2002

#### Objectifs :

L'apprentissage automatique est une composante majeure dans la conception de systèmes informatiques, plus particulièrement en Intelligence Artificielle. Il permet d'améliorer les performances des systèmes ou d'acquérir des connaissances sur un domaine. L'intérêt pour cette discipline s'est accru ces dernières années avec le développement de grandes bases de données, la nécessité de constituer des mémoires d'entreprises, et le besoin d'outils intelligents pour exploiter les flux d'informations sur le web. Ceci a conduit, notamment, à l'émergence de l'ECD (extraction de connaissances à partir de données) et de la fouille de textes.

La Conférence d'Apprentissage (CAp) est un rendez-vous annuel permettant aux chercheurs universitaires et industriels des diverses communautés francophones de se rassembler pour échanger des résultats théoriques et appliqués, des modèles, des méthodes, des outils et des expériences. Elle s'est construite sur le rapprochement des Journées Francophones d'Apprentissage (1985-1998) et du Groupe de Travail Apprentissage (groupe AFCET-AFIA, intitulé GTRA, 1995-1999). Ce rapprochement a finalement conduit à la création récente, dans le cadre de l'AFIA, du Collège Apprentissage, Fouille et Extraction (CAFE). Ainsi, grâce au travail accompli depuis maintenant une quinzaine d'années, une culture commune aux divers tenants de l'apprentissage automatique s'est constituée, alliant les aspects symboliques et formels issus de l'IA aux aspects statistiques, sans oublier l'analyse de données, et certains concepts de la physique.

L'édition 2002 sera organisée à Orléans, du 17 au 19 juin 2002, par le LIFO (Laboratoire d'Informatique Fondamentale d'Orléans). Nous vous invitons vivement à contribuer à cet événement qui favorise les échanges entre les différents courants de la communauté apprentissage.

#### Calendrier :

Date limite de soumission des articles :

21 février 2002

Notification aux auteurs de l'acceptation :

29 mars 2002

Réception des articles : 19 avril 2002

Réception des propositions de posters et démos :

7 mai 2002

Conférence : 17-19 juin 2002

#### Thèmes :

Voici une liste, non exhaustive, des sujets intéressant la conférence d'apprentissage CAp 2002 :

- Apprentissage symbolique et algébrique : arbres de décision, programmation logique inductive, apprentissage multi-instances, treillis de Galois, recherche de motifs fréquents, ...
- Apprentissage numérique ou statistique : réseaux neuronaux, réseaux bayésiens, machines à support de vecteurs, apprentissage semi-supervisé, co-training, ...
- Approches hybrides : algorithmes d'évolution, raisonnement à partir de cas, analogie, approches cognitives, ...
- Reformulation : sélection de descripteurs, sélection d'instances, propositionnalisation, abstraction, ...
- Etude formelle de l'apprentissage : apprenabilité, évaluation des systèmes d'apprentissage, liens avec l'expert et aspects interface homme-machine (biais, visualisation, modélisation de l'expert, etc.), ...
- Apprentissage à partir de données semi-structurées, de textes (fouille de textes, web mining, ...), de génomes, ... Applications : fouille de données, fouille de textes, bio-informatique, médecine, chimie, ...

#### Modalités de soumission des articles :

La longueur des papiers est limitée à 12 pages en

interlignes simples, police Times 12 points (les papiers trop longs ne seront pas pris en compte). La première page comportera classiquement le titre de la communication, le nom des auteurs et leur affiliation, les coordonnées et l'email de l'auteur à contacter, un résumé de la communication de 200 à 300 mots ainsi qu'une liste de mots clés décrivant les thèmes abordés dans l'article. Les articles pourront être soumis en français ou en anglais, cependant en cas d'acceptation, ils devront obligatoirement être publiés en français. Les présentations à la conférence seront faites en français.

La soumission devra être effectuée exclusivement par le biais du site web de la conférence (<http://www.univ-orleans.fr/CONFERENCES/CAP2002/>). En cas de problème, contactez-nous à l'adresse électronique de la conférence [CAp2002@lifo.univ-orleans.fr](mailto:CAp2002@lifo.univ-orleans.fr)

Les formats électroniques acceptés, classés par ordre de préférence décroissant sont les suivants : PDF, PS (compressé en .zip) et RTF.

Les articles soumis seront évalués par au moins deux relecteurs indépendants. Les critères d'acceptation sont l'originalité, la portée scientifique dans le cadre d'un travail théorique, le soin apporté à la validation dans le cadre d'un travail expérimental, le souci de réaliser une jonction entre l'apprentissage et des domaines connexes (ingénierie des connaissances, raisonnement à partir de cas, évolution artificielle, etc.) et enfin, la reproductibilité des algorithmes et des expériences.

Proposition de démonstrations ou de posters :

Par ailleurs, afin d'encourager les discussions entre les participants, une session Démonstrations & Posters se déroulera dans le cadre de la conférence. Les participants pourront y présenter librement les résultats de leurs recherches, c'est-à-dire, de manière indépendante du processus classique de soumission au comité de programme. Bien évidemment, seuls seront acceptés les posters et les démonstrations qui rentrent dans la thématique de CAP.

Les contributions pourront prendre deux formes :

- Présentation sous la forme d'un « poster » de travaux originaux.
- Démonstration d'un système d'apprentissage ou d'une application.

Les personnes qui présenteront une démonstration ou un poster devront décrire en 2 pages les travaux exposés. L'ensemble de ces contributions sera édité dans un fascicule, indépendant des actes de la conférence, qui sera distribué à tous les congressistes. Les deux pages devront être rédigées en Times 12 pts, interlignes simples, la première comportant un titre, le nom des

auteurs, leur adresse et une liste de mots clés. La soumission devra être effectuée par le biais du site web de la conférence (<http://www.univ-orleans.fr/CONFERENCES/CAP2002/>) avant le 7 mai 2002.

Il faut noter que pour toute soumission acceptée (article, démonstration ou poster), au moins un des auteurs devra s'inscrire à la conférence et s'engager à venir présenter les travaux.

## COMITE DE PROGRAMME :

Présidente du comité de programme :  
Christel Vrain (LIFO, Orléans)

Florence d'Alché-Buc (LIP6, Paris)  
Marc Bernard (EURISE, St Etienne)  
Gilles Bisson (LEIBNIZ, Grenoble)  
Jean-François Boulicaut (LISI, INSA Lyon)  
Laurent Brehelin (LIRMM, Montpellier)  
Antoine Cornuejols (LRI, Orsay)  
François Coste (IRISA, Rennes)  
François Denis (LIM, Marseille)  
Pierre Dupont (UCL, Louvain-la-Neuve, Belgique)  
Frédéric Garcia (INRA, Toulouse)  
Jean-Yves Jaffray (LIP6, Paris)  
Yves Kodratoff (LRI, Orsay)  
Stan Matwin (SITE, Ottawa, Canada & LRI, Orsay)  
Lionel Martin (LIFO, Orléans)  
Laurent Miclet (ENSSAT, Lannion)  
Claire Nédellec (INRA, Paris)  
Nicolas Pasquier (I3S, Nice)  
Hélène Paugam-Moisy (ISC, Lyon)  
Céline Rouveirol (LRI, Orsay)  
Lorenza Saitta (M&EC, Torino, Italie)  
Michèle Sebag (LRI / Polytechnique-LMS, Orsay)  
Henri Soldano (LIPN, Villetaneuse)  
Fabien Torre (Grappa-LIFL, Lille)  
Gilles Venturini (LI, Tours)  
Djamel Zighed (ERIC, Lyon2)  
Jean-Daniel Zucker (LIP6, Paris)

## COMITE D'ORGANISATION

Président du comité d'organisation :  
Matthieu Exbrayat (LIFO, Orléans)

Sylvie Billot, Agnès Braud, Guillaume Cleuziou, Isabelle Debourges, Sylvie Haouy-Maure, Lionel Martin, Frédéric Moal, Ansaf Salleb, Teddy Turmeau

## Informations pour les organisateurs de conférences en IA

### Le parrainage :

Tout organisateur de conférence, colloque, at-

lier,... dans le domaine de l'IA peut solliciter l'AFIA pour obtenir un parrainage de l'événement qu'il prépare. Le parrainage de l'AFIA est en quelque sorte une garantie, une caution scientifique qui est obtenue en se fondant sur différents critères tels que les objectifs scientifiques de la conférence, la composition du comité de programme, la procédure de sélection des articles... Le parrainage de l'AFIA se veut également être un indicateur pour tous ceux qui s'interrogent sur l'intérêt de participer à telle ou telle manifestation.

L'octroi du parrainage de l'AFIA s'accompagne d'un engagement réciproque entre l'organisateur et l'AFIA. Cet engagement permet à l'organisateur de bénéficier de la plupart des moyens de diffusion propres à l'AFIA. En contrepartie le parrainage permet à l'AFIA de se faire mieux connaître, de donner aux membres de l'AFIA des informations plus complètes sur l'activité de la communauté IA et d'obtenir des réductions pour ses membres pour les inciter à participer aux manifestations parrainées.

### **Que faire ?**

#### *Cas général (pour tous)*

Pour toutes les conférences, que celles-ci soient organisées par l'AFIA, par les groupes de travail ou collègues, comme pour les conférences qui ont obtenues le parrainage, l'organisateur est invité à faire parvenir :

l'URL du serveur de la manifestation (à adresser à Sylvie PESTY) ainsi qu'au responsable des « Groupes de travail et Collèges » (Jean-Paul BAQUIAST - voir adresse page 2 du Bulletin).

Les différentes annonces de la conférence, une page environ pour annoncer la conférence et une à deux pages pour faire connaître le bilan de la conférence (à adresser au responsable des conférences dans le bulletin (Sylvie PESTY).

Cas particulier des conférences organisées par l'AFIA, ses groupes de travail ou ses collègues

L'organisation est déléguée à un organisateur local. Un contrat doit être entre l'AFIA et l'organisateur local pour définir les conditions de la délégation, essentiellement au niveau du partage des risques financiers et du résultat de la manifestation (bénéfice ou déficit).

Trois modèles de contrat sont proposés aux organisateurs locaux ; ces contrats sont aménageables (nous faire une proposition : Sylvie PESTY)

un modèle de contrat avec délégation des risques et résultats à l'organisateur local.

un modèle de contrat avec délégation des risques et résultats à l'AFIA.

un modèle de contrat pour un partage des risques et résultats entre l'organisateur local et l'AFIA.

Pour toute information supplémentaire, n'hésitez pas à consulter le site de l'AFIA ou à contacter :

Sylvie PESTY [Sylvie.Pesty@imag.fr](mailto:Sylvie.Pesty@imag.fr)

Laboratoire Leibniz-IMAG

46, Avenue Felix Viallet

F-38031 Grenoble Cedex

Tél : 04 76 57 48 07

Fax : 04 76 57 50 81

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 texte (mél + fichier attaché) à Sylvie PESTY (voir « l'ours » page 2 de ce bulletin).

# SOMMAIRES DES REVUES

*N'hésitez pas à envoyer un message à Brigitte Grau (grau@limsi.fr) pour lui indiquer toute suggestion permettant d'améliorer cette rubrique. Les revues figurant régulièrement au sommaire mais n'ayant pas de nouveau numéro apparaissent seulement avec leur nom et leur adresse WEB.*

## REVUE D'INTELLIGENCE ARTIFICIELLE

[http://www.editions-hermes.fr/rev\\_archiv.asp?code=ria](http://www.editions-hermes.fr/rev_archiv.asp?code=ria)

### REVUE D'INTELLIGENCE ARTIFICIELLE V15 N°1

- Une approche orientée agent pour la conception du travail coopératif, *S. AKNINE, S. PINSON*
- Problèmes d'évaluation dans la communication orale homme-machine, *W. MINKER*
- Apprentissage supervisé et raisonnement approximatif, l'hypothèse des imperfections, *A. BORGI, H. AKDAG*
- Conceptual graphs with constrained reasoning, *D. CORBETT*

## ELECTRONIC TRANSACTIONS ON ARTIFICIAL INTELLIGENCE (ETAI)

Free publication and interactive reviewing on the internet  
<http://www.ida.liu.se/ext/etai/>

## ARTIFICIAL INTELLIGENCE

<http://www.elsevier.nl:80/inca/publications/store/5/0/5/6/0/1/>

### ARTIFICIAL INTELLIGENCE V133 N°1-2 DECEMBER 2001

- On decision-theoretic foundations for defaults, *R. I. BRAFMAN, N. FRIEDMAN*
- An argument-based approach to reasoning with specificity, *P. M. DUNG, T. C. SON*
- The problem of survival from an algorithmic point of view, *J.-D. FOUKS, L. SIGNAC*
- Recognition of gestures in Arabic sign language using neuro-fuzzy systems, *O. AL-JARRAH, A. HALAWANI*
- Reasoning about nonlinear system identification, *E. BRADLEY, M. EASLEY, R. STOLLE*
- A general formulation of conceptual spaces as a meso level representation, *J. AISBETT, G. GIBBON*
- Defeasible reasoning with variable degrees of justification, *J. L. POLLOCK*

## AI MAGAZINE

<http://aaai.org/Magazine/magazine.html>

### AI MAGAZINE V22 N°3 2001

- Creativity at the metalevel AAI-2000 presidential address,

### B.G. BUCHANAN

- A gamut of games, *J. SCHAEFFER*
- AIPS'00 planning competition: The fifth international conference on Artificial Intelligence Planning and Scheduling systems, *F. BACCHUS*
- FF: The fast-forward planning system, *J. HOFFMANN*
- The GRT planner, *I. REFANIDIS, I. VLAHAVAS*
- MIPS: The model-checking integrated planning system, *S. EDELKAMP, M. HELMERT*
- A planner called R, *F. LIN*
- Heuristic search planner 2.0, *B. BONET, H. GEFFNER*
- STAN4: A hybrid planning strategy based on subproblem abstraction, *M. FOX, D. LONG*
- TOKENPLAN: A planner for both satisfaction and optimization problems, *Y. MEILLER, P. FABIANI*
- ALTALT: Combining graphplan and heuristic state search, *B. SRIVASTAVA ET AL.*
- The shop planning system, *D. NAU, Y. CAO, A. LOTEM, H. MUNOZ-AVILA*
- TALPLANNER: A temporal logic-based planner, *P. DOHERTY, J. KVARNSTROM*
- Planning in the fluent calculus using binary decision diagrams, *H.-P. STORR*

## COGNITIVE SCIENCE

[http://www.elsevier.nl/gejng/29/18/show/Products/CITE/frames/frame\\_right.htm?area=Journals&factor=1](http://www.elsevier.nl/gejng/29/18/show/Products/CITE/frames/frame_right.htm?area=Journals&factor=1)

## APPLIED ARTIFICIAL INTELLIGENCE

<http://www.tandf.co.uk/journals/tf/08839514.html>

### APPLIED ARTIFICIAL INTELLIGENCE V15 N°9 OCTOBER 2001

- A neuro-fuzzy hybrid scheme for design and simulation of human machine systems, *X. F. ZHA*
- Combined method of genetic programming and association rule algorithm, *A. NIIMI, E. TAZAKI*
- Rough set-aided keyword reduction for text categorization, *A. CHOUCOULAS, Q. SHEN*

### APPLIED ARTIFICIAL INTELLIGENCE V15 N°10 NOVEMBER 2001

- Automatic generation of communication and teamwork within multi-agent teams, *K. E. BIGGERS, T. R. IOERGER*
- Agent interaction for bioinformatics data management, *K. BRYSON, M. LUCK, M. JOY, D. T. JONES*
- Learning-based scheduling of flexible manufacturing systems using case-based reasoning, *P. PRIORE, D. DE LA FUENTE, R. PINO*

## COMPUTATIONAL INTELLIGENCE

<http://www.blackwellpublishers.co.uk/asp/journal.asp?ref=08247935&src=cts>

# SOMMAIRES DES REVUES

## COMPUTATIONAL INTELLIGENCE V17 N°4 NOVEMBER 2001

- A Complexity-Level Analysis of the Sensor Planning Task for Object Search, *Y. YE, J. K. TSOTSOS*
- Control States and Complete Agent Architectures, *D. N. DAVIS*
- Semantic Forcing in Disjunctive Logic Programs, *M. DE VOS, D. VERMEIR*
- Discovering Maximal Generalized Decision Rules Through Horizontal and Vertical Data Reduction, *X. HU, N. CERCO-NE*
- Granularity in Relational Formalisms - With Application to Time and Space Representation, *J. EUZENAT*
- Social Mental Shaping: Modelling the Impact of Sociality on the Mental States of Autonomous Agents, *P. PANZARASA, N. R. JENNINGS, T. J. NORMAN*
- A Pattern-Weight Formulation of Search Knowledge, *R. LEVINSON, G. FUCHS*
- Letter to the Editor of Computational Intelligence: Issues Concerning Kohanim and Johnson's "Using Images as a Foundation for Natural Language Processing" (Nov. 2000), *T. C. MANSFIELD*

## MINDS AND MACHINES

### JOURNAL FOR ARTIFICIAL INTELLIGENCE PHILOSOPHY AND COGNITIVE SCIENCE

<http://www.wkap.nl/prod/j/0924-6495>

#### MINDS AND MACHINES V11 N°3 AUGUST 2001

- Machine Mentality and the Nature of the Ground Relation, *D. WHOBREY*
- Does Intentional Psychology Need Vindicating by Cognitive Science?, *J. KNOWLES*
- Tractable Competence, *M. FRIXIONE*
- vLearning Concepts by Arranging Appropriate Training Order, *Y.-T. HSU, T.-P. HONG, S.-S. TSENG*
- Attractive and In-discrete, *C. ELIASMITH*
- Equivalence and Format: On Strategies for Recognition. Remarks on Kosslyn's Reply, *M. ROLLINS*

#### Book review

- Rodney Cotterill, *Enchanted Looms: Conscious Networks in Brains and Computers.*, *R. WILSON*
- Susan Blackmore, *The Meme Machine.*, *M. ELTON*
- R. B. Horn, J. Yoshimi, M. Deering, R. McBride, D. Fleischman, T. DiDonna, J. Wedel (Editor), *Mapping Great Debates. Can Computers Think?: 7 Maps and a Handbook.*, *A. P. SAYGIN*
- M. Spitzer, *The Mind Within the Net. Models of Learning, Thinking, and Acting.*, *K. AIZAWA*
- F. Cowie, *What's Within? Nativism Reconsidered.* *Philosophy of Mind Series.*, *P. ALWARD*

#### MINDS AND MACHINES V11 N°4 NOVEMBER 2001

- Explanation and Complexity, *J. SYMONS*
- Coexisting Spatio-Temporal Scales In Neuroscience, *A. PEREIRA JÚNIOR*

- Understanding Neural Complexity: A Role For Reduction, *J. BICKLE*
- The Compatibility of Complex Systems and Reduction: A Case Analysis of Memory Research, *W. BECHTEL*
- Explanations at Multiple Levels1, *A. RUEGER*
- Explanation, Representation and the Dynamical Hypothesis, *J. SYMONS*
- Computational versus Causal Complexity, *M. SCHEUTZ*
- Intelligence is not Enough: On the Socialization of Talking Machines, *E. M.A. RONALD, M. SIPPER*
- Real Logic is Nonmonotonic, *H. E. KYBURG JR.*

#### Book review

- N. Griffith and P. M. Todd, eds., *Musical Networks: Parallel Distributed Perception and Performance*, *W. TABOR*
- E. Lepore and Z. Pylyshyn, *What is Cognitive Science?*, *D. KILFOYLE*

## ARTIFICIAL INTELLIGENCE REVIEW

[http://www.elsevier.nl/gej-ng/29/18/show/Products/CITE/frames/frame\\_right.htm?area=Journals&factor=1](http://www.elsevier.nl/gej-ng/29/18/show/Products/CITE/frames/frame_right.htm?area=Journals&factor=1)

#### AI REVIEW V16 N°1 SEPTEMBER 2001

- A survey on case-based planning, *L. SPALAZZI*
- Repeating and remembering foreign language words: Implications for language teaching systems, *C. DAVIS, J. KIM*
- Meta-patterns and higher order meta-patterns in cellular systems, *P. VAN LOOCKE*
- Incorporating software visualization in the design of intelligent diagnosis systems for user programming, *H.A. RAMADHAN, F. DEEKAND, K. SHIHAB*

#### AI REVIEW V16 N°2 OCTOBER 2001

- Efficient case-based structure generation for design support, *K. BORNER*
- Current directions in computational humour, *G. RITCHIE*
- Conceptual spaces for computer vision representations, *A. CHELLA, M. FRIXIONE, S. GAGLIO*
- Scalable parallel genetic algorithms, *W. RIVERA*

#### AI REVIEW V16 N°3 NOVEMBER 2001

- Understanding the crucial role of attribute interaction in data mining, *A.A. FREITAS*
- Incremental evolution in ANNs: Neural nets which grow, *C. MACLEOD, G.M. MAXWELL*
- Goal-seeking behaviour in a connectionist mode, *T.E. PORTEGYS*

## INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES

<http://www.academicpress.com/ijhcs>

#### IJHCS V55 N°5 NOVEMBER 2001

- Effects of training and representational characteristics in

# SOMMAIRES DES REVUES

icon design, *R. S. GOONETILLEKE, H. MARTINS SHIH, H. K. ON, J. FRITSCH*

- Graphic and numerical methods to assess navigation in hypertext, *J. E. MCENEANEY*
- The effects of contextualized access to knowledge on judgement, *JI-YE MAO, I. BENBASAT*
- Implications of buyer decision theory for design of e-commerce websites, *B. G. SILVERMAN, M. BACHANN, K. ALAKHARAS*
- Task structure and the apparent duration of hierarchical search, *N. TRACTINSKY, J. MEYER*

## COMPUTATIONAL LINGUISTICS

<http://mitpress.mit.edu/catalog/item/default.asp?sid=8563C099-9701-4DD2-85C8-8F3502E9C8AE&tttype=4&tid=10>

### COMPUTATIONAL LINGUISTICS V27 N° 3 SEPTEMBER 2001

- The Interaction of Knowledge Sources in Word Sense Disambiguation, *M. STEVENSON, Y. A. WILKS*
- A Statistical Model for Word Discovery in Transcribed Speech, *A. VENKATARAMAN*
- Automatic Verb Classification Based on Statistical Distributions of Argument Structure, *P. MERLO, S. STEVENSON*
- Towards Constructive Text, Diagram, and Layout Generation for Information Presentation, *J. BATEMAN, J. KLEINZ, T. KAMPS, K. REICHENBERGER*

#### Book review

- Edited by Merle Horne, Prosody: Theory and Experiment. Studies Presented to Gösta Bruce, *C. SHIH, R. SPROAT*
- Edited by Frank Van Eynde and Dafydd Gibbon, Lexicon Development for Speech and Language Processing, *C. BREW*
- Stephen C. Levinson, Presumptive Meanings: The Theory of Generalized Conversational Implicature, *N. GREEN*
- Richard Sproat, A Computational Theory of Writing Systems, *K. R. BEESLEY*

## NATURAL LANGUAGE ENGINEERING

<http://www.journals.cambridge.org/bin/bladerunner?REQUEST=976614197&REQSESS=5004848&116000REQEVENT=&REQINT1=5&REQSTR1=NLE&REQAUTH=0>

## USER MODELING AND USER-ADAPTED INTERACTION

<http://www.wkap.nl/jrnltoct.htm/0924-1868>

### USER MODELING AND USER-ADAPTED INTERACTION V12 N°1 FEBRUARY 2002

- Assessment of User Affective and Belief States for Interface Adaptation: Application to an Air Force Pilot Task, *E. HUDLICKA, M. D. MCNEESE*

- Modeling Multimodal Expression of User's Affective Subjective Experience, *N. BIANCHI-BERTHOUBE, C. L. LISETTI*
- Affective Computing, *R. W. PICARD*
- Affective Interactions: Towards a New Generation of Computer Interfaces, *A. PAIVA*

## COMPUTER SPEECH AND LANGUAGE

<http://www.idealibrary.com/links/toc/csla/latest>

### COMPUTER SPEECH & LANGUAGE V15 N°4 OCTOBER 2001

- Augmenting standard speech recognition features with energy gravity centres, *R. DE MORI, L. MOISA, R. GEMELLO, F. MANA, D. ALBESANO*
- Synthesis of Arabic from short sound clusters, *Y. A. EL-IMAM*
- Spectral stability based event localizing temporal decomposition, *A. C. R. NANDASENA, P. C. NGUYEN, M. AKAGI*
- A bit of progress in language modelling, *J. T. GOODMAN*

## MACHINE LEARNING

<http://www.wkap.nl/jrnltoct.htm/0885-6125>

### MACHINE LEARNING V46 N°1-3 JANUARY 2001

- Editorial: Kernel Methods: Current Research and Future Directions, *N. CRISTIANINI, C. CAMPBELL, C. BURGESS*
- On a Connection between Kernel PCA and Metric Multidimensional Scaling, *C. K.I. WILLIAMS*
- Bayesian Methods for Support Vector Machines: Evidence and Predictive Class Probabilities, *P. SOLLICH*
- Hierarchical Learning in Polynomial Support Vector Machines, *S. RISAU-GUSMAN, M. B. GORDON*
- A Probabilistic Framework for SVM Regression and Error Bar Estimation, *J.B. GAO, S.R. GUNN, C.J. HARRIS, M. BROWN*
- On the Dual Formulation of Regularized Linear Systems with Convex Risks, *T. ZHANG*
- Choosing Multiple Parameters for Support Vector Machines, *O. CHAPELLE, V. VAPNIK, O. BOUSQUET, S. MUKHERJEE*
- Training Invariant Support Vector Machines, *D. DECOSTE, B. SCHÖLKOPF*
- Support Vector Machines for Classification in Nonstandard Situations, *Y. LIN, Y. LEE, G. WAHBA*
- An Analytic Center Machine, *T. B. TRAFALIS, A. M. MALYSCHEFF*
- Linear Programming Boosting via Column Generation, *A. DEMIRIZ, K. P. BENNETT, J. SHAWE-TAYLOR*
- Large Scale Kernel Regression via Linear Programming, *O.L. MANGASARIAN, D. R. MUSICANT*
- Efficient SVM Regression Training with SMO, *G. W. FLAKE, S. LAWRENCE*
- A Simple Decomposition Method for Support Vector

## SOMMAIRES DES REVUES

Machines, C.-W. HSU, C.-J. LIN

- Feasible Direction Decomposition Algorithms for Training Support Vector Machines, P. LASKOV
- Convergence of a Generalized SMO Algorithm for SVM Classifier Design, S.S. KEERTHI, E.G. GILBERT
- The Relaxed Online Maximum Margin Algorithm, Y. LI, P. M. LONG
- Gene Selection for Cancer Classification using Support Vector Machines, I. GUYON, J. WESTON, S. BARNHILL, V. VAPNIK
- Text Categorization with Support Vector Machines. How to Represent Texts in Input Space?, E. LEOPOLD, J. KINDER-MANN

### NEURAL NETWORKS

[http://www.elsevier.nl/gejng/29/18/show/Products/CITE/frames/frame\\_right.htm?area=Journals&factor=1](http://www.elsevier.nl/gejng/29/18/show/Products/CITE/frames/frame_right.htm?area=Journals&factor=1)

#### NEURAL NETWORKS V14 N°10 DECEMBER 2001

- Connectionist inference models, A. BROWNE, R. SUN
- Analysis of the frequency response of the saccadic circuit: numerical simulations, M. E. JACKSON, O. LITVAK, J. W. GNADT
- Unified stabilization approach to principal and minor components extraction algorithms, T. CHEN, S.-I. AMARI
- Improvement of generalization ability for identifying dynamical systems by using universal learning networks, K. HIRASAWA ET AL.
- Self-organization in the one-dimensional SOM with a decreasing neighborhood, J. A. FLANAGAN
- Upper bound of the expected training error of neural network regression for a Gaussian noise sequence, K. HAGIWARA, T. HAYASAKA, N. TODA, S. USUI, K. KUNO
- The POP learning algorithms: reducing work in identifying fuzzy rules, C. QUEK, R.W. ZHOU
- Oriented principal component analysis for large margin classifiers, S. BERMEJO, J. CABESTANY
- A remark on 'On stability of nonlinear continuous-time neural networks with delays', Y. CHEN
- On using the Poincaré polynomial for calculating the VC dimension of neural networks, M. SCHMITT

- Response: on using the Poincaré polynomial for calculating the V-C dimension of neural networks, M. A. CARTER, M. E. OXLEY
- Evaluating the Vapnik-Chervonenkis dimension of artificial neural networks using the Poincaré polynomial, Neural Networks, 12, 403-408, (1999), M. A. CARTER, M. E. OXLEY

### ARTIFICIAL INTELLIGENCE IN MEDECINE

[http://www.elsevier.nl/gejng/29/18/show/Products/CITE/frames/frame\\_right.htm?area=Journals&factor=1](http://www.elsevier.nl/gejng/29/18/show/Products/CITE/frames/frame_right.htm?area=Journals&factor=1)

#### AI IN MEDICINE V23 N°3 NOVEMBER 2001

- A method for automated temporal knowledge acquisition applied to sleep-related breathing disorders, G. GUIMARÃES, J.-H. PETER, T. PENZEL, A. ULTSCH
- A neural network approach to the outcome definition on first treatment with sertraline in a psychiatric population, L. FRANCHINI ET AL.
- A modular approach for representing and executing clinical guidelines, P. TEREZIANI, G. MOLINO, M. TORCHIO
- A sequential neural network model for diabetes prediction, J. PARK, D. W. EDINGTON

### INTERNATIONAL JOURNAL OF APPROXIMATE REASONING

[http://www.elsevier.nl/gejng/29/18/show/Products/CITE/frames/frame\\_right.htm?area=Journals&factor=1](http://www.elsevier.nl/gejng/29/18/show/Products/CITE/frames/frame_right.htm?area=Journals&factor=1)

#### IJAR V29 N°1 JANUARY 2002

- Using probability trees to compute marginals with imprecise probabilities, A. CANO, S. MORAL
- Belief function independence: I. The marginal case, B. BEN YAGHLANE, P. SMETS, K. MELLOULI
- Soft evidential update for probabilistic multiagent systems, M. VALTORTA, Y.-G. KIM, J. VOMLEL

## Démarche méthodologique de conception de systèmes d'information coopératifs inter agents pour la gestion des connaissances Par Imed Boughzala,

Thèse de l'Université Paris 6 Jussieu, soutenue le 20 décembre 2001 à Paris 6.

### Résumé.

Dans cette thèse, nous nous intéressons aux Systèmes d'Information Coopératifs des entreprises étendues, en particulier quand ceux-ci traitent des problématiques de gestion des connaissances. Ces systèmes d'information possèdent des spécificités intrinsèques liées au caractère potentiellement hétérogène des sources d'information, des processus et des savoirs mis en relation et à l'indépendance que les agents souhaitent conserver.

Ces spécificités nécessitent la mise en place d'une démarche méthodologique originale qui intègre des dimensions organisationnelles pour proposer des méthodes d'ingénierie des connaissances basées sur des modèles conceptuels génériques de coopération inter agents adaptés au contexte de l'entreprise étendue. C'est l'objectif que nous poursuivons pour la construction de la démarche MeDICIS (Methodology for Designing Interenterprise Cooperative Information System).

Nous proposons d'abord une approche théorique de la notion de coopération et de la manière dont celle-ci se déroule entre les entreprises dans le contexte de « l'entreprise étendue » que nous différencions de l'entreprise réseau, du réseau d'entreprises et de l'entreprise virtuelle. Nous introduisons ensuite les notions essentielles permettant d'aborder la conception du système d'information et des systèmes de connaissances avant d'introduire les concepts issus des systèmes multi-agents et du CSCW qui nous permettront de présenter notre modèle d'entreprise étendue. Celui-ci est caractérisé par l'introduction d'un « agent coopérateur », un « acteur » humain, qui s'appuie sur le système d'information coopératif interentreprises pour assurer la communication, la coordination, et gestion des connaissances et gérer les conflits « sociaux » pouvant surgir dans l'entreprise étendue.

Nous définissons ensuite les différents modèles de MeDICIS qui permettent de spécifier le système d'information de l'entreprise étendue, le modèle d'affaire, de coopération, d'agent, de communication, de coordination et de résolution collective de problème qui s'appuient sur des langages de modélisation comme UML et UEML et des méthodes de gestion de connaissances comme CommonKADS et MASK. Nous appliquons ces modèles dans le cadre d'une expérimentation au sein de la filière textile et présentons un prototype en deux versions implémentant les propriétés du modèle de communication.

**Mots clés :** Système d'information, gestion des connaissances, coopération interentreprises, entreprise étendue, CSCW.

**Adresse :** Imed Boughzala

Département GSID – UTT, 12, rue de Marie Curie - BP 2060, 10010 Troyes cedex

**Email :** [Imed.boughzala@utt.fr](mailto:Imed.boughzala@utt.fr)

Téléphone : 03 25 71 56 88

Télécopie : 03 25 71 56 99

## Svetlan', un système de structuration du lexique guidé par la détermination automatique du contexte thématique, par Gaël de Chalendar

Thèse de l'Université Paris XI, soutenue le 21 décembre 2001 au LIMSI à Orsay.

### Résumé.

Des connaissances sémantiques sont obligatoires pour le Traitement Automatique des Langues. Les bases de connaissances générales posent des problèmes. Jusqu'à présent, on savait extraire automatiquement des connaissances sémantiques depuis les textes de spécialité mais pas de façon adéquate depuis des textes de la langue dite "générale". Cette thèse a pour but de montrer que ce type d'extraction est possible.

Nous montrons qu'une analyse thématique de textes non spécialisés (journalistiques en particulier) permet la plupart du temps de se ramener dans le cadre d'un problème classique de traitement de corpus spécialisé, tout en nécessitant des interventions humaines très réduites.

Dans notre approche, le thème des segments de textes est détecté par l'analyse statistique des distributions des mots. Après avoir défini des notions de similarité et d'agrégation, les mots des segments similaires sont agrégés pour former des domaines thématiques dans lesquels les mots de poids élevés décrivent un thème. On regroupe les noms qui apparaissent comme argument d'un même verbe dans les divers segments de texte appartenant à un certain thème, ce qui forme des classes. Notre argumentation est implémentée dans un système informatique, Svetlan', qui a été testé sur des corpus de plusieurs millions de mots en français et en anglais. L'analyse empirique des résultats montre que, comme prévu, les mots sont très souvent en relation sémantique forte les uns avec les autres dans les classes obtenues, et ce dans le contexte déterminé par le thème.

Le jugement humain de classes n'étant pas suffisamment sûr, nous validons de manière indirecte les connaissances acquises par Svetlan' en les utilisant dans une tâche d'extension de requêtes en vue de l'amélioration des performances d'un système de réponse à des questions en langue naturelle.

**Mots clés :** traitement automatique des langues, sémantique lexicale, apprentissage de classes de termes, méthode hybride symbolique et statistique, acquisition en contexte.

**Contact :** Gaël de Chalendar

LIMSI/CNRS, BP 133, 91403 Orsay Cedex.

Email : [Gael.de.Chalendar@limsi.fr](mailto:Gael.de.Chalendar@limsi.fr)

**Adresse Web :** <http://www.limsi.fr/Individu/gael/>

Téléphone : 33 1 69 85 80 04

Télécopie : 33 1 69 85 80 88

## Codage et traitement phylogénétique de caractères structuraux de génomes entiers par Cyril Gallut

Thèse de l'Université Pierre et Marie Curie (Paris 6), soutenue le 18 décembre 2001 à Paris.

### Résumé.

Une approche originale de reconstruction phylogénétique

tique à partir de l'organisation de génomes entiers dans un contexte cladistique est proposée. Cette approche se fonde sur la comparaison globale de l'organisation des génomes étudiés, sans hypothèses a priori sur les remaniements. Deux codages sont proposés : « Position relative » et « Jonctions », avec deux options différentes. Ces possibilités de codage sont analysées et comparées avec le codage « Jonctions signées » de Cosner et al. ; ils permettent de représenter l'ordre d'unités fonctionnelles le long des chromosomes, les unités fonctionnelles pouvant être des gènes ou des segments chromosomiques homologues par exemple. Le codage « Position relative » regroupe trois types de caractères : des caractères de position, d'orientation et de présence/absence des unités. Le codage jonction est basé sur des caractères binaires : polarité d'unité, présence/absence d'unités et de jonctions d'unités. Le codage « Jonctions signées » est un codage entièrement fondé sur la présence/absence de jonctions signées. Ce

dernier présente l'avantage de prendre en compte l'orientation et la position simultanément. Par contre, le codage « Position relative » permet de reconstituer les génomes ancestraux a posteriori, ce qui est d'un grand intérêt pour l'interprétation de l'évolution du génome. Le génome mitochondrial des métazoaires ainsi que les chromosomes du genre *Mastomys* ont été analysés avec succès au moyen des trois codages.

**Mots clés :** cladistique, phylogénie, codage, génomique, chromosome, ordre de gènes.

**Contact :** Cyril Gallut, Laboratoire d'Informatique et Systématique 12, rue Cuvier 75005 Paris.

**Email :** [gallut@ccr.jussieu.fr](mailto:gallut@ccr.jussieu.fr)

**Adresse Web :** <http://lis.snv.jussieu.fr/~gallut>

Téléphone : 01 44 27 65 21

Télécopie : 01 44 27 65 60

## Modélisation et traitement de formes spatiales et dynamiques Reconstruction 3D, mouvement et Geste, Architecture du Génome.

par Rachid Gherbi

Habilitation à diriger des recherches de l'Université de Paris-Sud XI, soutenue le 5 décembre 2001, LIMSI à Orsay.

### Résumé.

Mes travaux de recherche en Informatique concernent les domaines de la Vision par Machine, de la Reconnaissance de Formes et de la

Modélisation d'Objets 3D pour des problématiques pluridisciplinaires et applications en Communication Homme-Machine, en Réalité Virtuelle et Augmentée et en Bio-Informatique. Le point commun à ces travaux est l'image, dans sa signification la plus large, qui est un objet de nature spatiale et dynamique. De nos jours, on parle d'images réelles ou virtuelles, voire d'images mixtes alliant le réel au virtuel. L'image est utilisée pour ses capacités à séduire, à explorer, à révéler, à changer de point de vue, à prouver et à expliquer. Les domaines du traitement de l'information et de la communication ne sont pas restés insensibles à cette information multidimensionnelle, allant d'une part du traitement d'images à la représentation des objets, et d'autre part de la conception d'ob-

jets à la synthèse d'images.

Dans ce contexte de problématiques de recherche fondamentale menés depuis une dizaine d'années, plusieurs systèmes ont été développés d'une part pour le suivi, la modélisation et l'interprétation du regard et des gestes humains ; et d'autre part en reconstruction d'objets 3D par des approches de coopération analyse/synthèse d'images. Par ailleurs, un environnement logiciel, appelé ADN-Viewer est en développement qui offre la représentation et la visualisation stéréoscopique de séquences d'ADN à l'échelle génomique.

Dans les années à venir, je compte mettre l'accent sur l'interface entre informatique et biologie pour consolider et développer davantage la synergie entre ces disciplines. De façon spécifique, je me focaliserai sur l'analyse, la représentation, la reconnaissance et l'interprétation de structures spatiales et dynamiques des séquences génomiques et des images issues du Transcriptome/Protéome.

En parallèle, les travaux sur ces données factuelles seront mis en lien avec les travaux sur les données textuelles en collaboration avec des spécialistes en extraction d'information et data-mining. De façon globale, les modèles informatiques doivent être souvent (voire toujours) validés biologiquement et affinés en retour.

**Mots clés :** Image, reconstruction 3D, reconnaissance de formes, geste, bioinformatique.

**Contact :** Rachid Gherbi,

LIMSI-CNRS - Université Paris-Sud, Orsay

Téléphone : +33/01 69 85 81 64

Télécopie : +33/01 69 85 80 88

**Email :** [Rachid.Gherbi@limsi.fr](mailto:Rachid.Gherbi@limsi.fr)

**Adresse Web :** <http://www.limsi.fr/Individu/gherbi>

## Expérience de couplage entre bases de données factuelles et bases de données bibliographiques : Identification dans Medline des gènes décrits dans Flybase et application à l'extraction d'informations sur les interactions génétiques ou moléculaires à partir de publications par Ambroise Ingold

Thèse de l'Université d'Aix-Marseille 3, soutenue le 4 Janvier 2002.

### Résumé.

La thèse propose des solutions pour mettre automatiquement en relation des informations bibliographiques avec des informations factuelles. Les bases de données bibliographiques donne accès à l'information la plus exhaustive et la plus précise tandis que les bases de données factuelles organisent le savoir accumulé dans des encyclopédies électronique ou dans des banques de résultats d'expériences. Coupler ces deux types de bases de données est nécessaire. Il s'agit soit de documenter des bases de données factuelles avec des références bibliographiques, soit d'extraire de l'information directement à partir de la littérature scientifique.

Nous avons pris l'exemple des gènes et de leurs interactions chez la Drosophile. La Drosophile est un organisme modèle en génétique et l'analyse des interactions génétiques

ou moléculaires permet de comprendre comment plusieurs gènes collaborent à une même fonction.

Dans un premier temps, nous avons construit un système qui permet de créer des liens entre Flybase et Medline. Flybase est une encyclopédie électronique sur la Drosophile. Medline est la plus grande base de données bibliographiques dans le domaine des sciences de la vie. Ces liens consistent à identifier dans Medline des gènes décrits dans Flybase. Cette tâche est difficile à automatiser en raison de la complexité de la nomenclature (existence d'alias, d'abréviations et de termes vagues, composition de termes utilisant des noms de gènes, etc.) et de la confusion possible entre certains noms de gènes et des mots du vocabulaire courant.

Dans un second temps, notre travail a consisté à établir une liste d'interactions probables à partir d'un ensemble de résumés issus de Medline. Cela a été fait par l'analyse statistique du vocabulaire utilisé. La méthode a été testée avec succès et le détail de la mise en œuvre est donné dans le document.

**Mots clés :** Extraction d'informations, informatique documentaire, statistique textuelle, terminologie, interaction génétique.

**Contact :** Ambroise Ingold,  
136, rue Amelot 75011 Paris.

**Email :** [ingold@free.fr](mailto:ingold@free.fr)

**Adresse Web :** [ingold.free.fr](http://ingold.free.fr)  
Téléphone : 06 61 66 70 90.

---

## Des modèles en Logiques de Description pour l'aide à la décision par Martina Kullmann

**Thèse de l'Université Louis Pasteur à Strasbourg, soutenue en mai 2001 à Strasbourg.**

### Résumé.

En utilisant les Logiques de Descriptions (LD) pour la modélisation des situations complexes et dynamiques ainsi que des réactions appropriées, nous proposons un schéma basé sur les LD pour l'aide à la décision. Suivant ce schéma, pour une application donnée, des modèles de LD plus spécifiques peuvent être élaborés. Comme extension de ce schéma nous proposons l'utilisation du raisonnement LD pour construire un modèle de la base de connaissance considérée. En se basant sur la connaissance générale d'un domaine d'application considéré et sur une description spécifique d'une situation donnée, il est possible de déduire des actions concrètes adaptées à cette situation. Ces actions ne sont pas forcément prédéfinies d'une façon explicite. De plus, une mise à jour régulière de la base de connaissances permet d'effectuer le

raisonnement LD par rapport à l'état courant de la situation considérée.

En pratique, la construction de modèles à l'aide des LD nécessite une adaptation de l'algorithme conventionnel de raisonnement, c'est-à-dire au lieu de vérifier seulement l'existence ou l'inexistence d'un seul modèle l'algorithme doit calculer et stocker plusieurs modèles en entier. Pour résoudre le problème de mémoire ainsi que du temps de calcul, nous proposons d'utiliser un tableau optimisé en structure de treillis. Un deuxième avantage du stockage des modèles est la réutilisation des résultats des calculs précédents pour réaliser de nouveaux calculs, ce qui implique une réduction du temps de cal-

cul. En plus nous proposons une manière d'exploiter la nouvelle structure de tableau pour un raisonnement non-monotone et pour fournir des explications sur les origines des contradictions.

### Mots clés

représentation des connaissances, aide à la décision, logiques de descriptions, modélisation.

**Contact :** Martina Kullmann,  
Essenweinstr. 18, D-76131 Karlsruhe  
(Allemagne).

**Email :** [m.kullmann@gmx.net](mailto:m.kullmann@gmx.net)

Téléphone : (+49) 0721-7817344

---

## Extraction des connaissances basée sur les treillis de Galois: Méthodes et applications par Engelbert Mephu Nguifo

**Habilitation à diriger des recherches de l'Université d'Artois -- Faculté Jean Perrin à Lens, soutenue le 10 décembre 2001.**

### Résumé.

L'extraction de connaissances est un processus interactif et itératif d'analyse d'un grand ensemble de données brutes afin d'en extraire des connaissances exploitables, et où l'utilisateur-analyste joue un rôle central. Dans la perspective de conception de systèmes d'extraction de connaissances, nous présentons nos travaux d'élaboration de méthodes de prétraitement de données, et de classification par treillis pour la fouille de données.

La première partie introduit la problématique du prétraitement de données en traitant deux types de méthodes : la sélection et la construction d'attributs. Ensuite sont abordés plusieurs méthodes de classification s'appuyant sur les arbres de décision, les plus proches voisins, les treillis de Galois, et la logique floue.

Nous présentons dans la deuxième partie nos méthodes de conception de systèmes de classification basée sur le treillis de Galois et portant progressivement sur l'usage du vote majoritaire, sur l'introduction des mesures de la théorie de l'information, la transformation d'attributs, l'usage des techniques des plus proches voisins, la sélection des prototypes, et sur l'introduction de la théorie des sous-ensembles flous. Ces méthodes ont permis d'implanter plusieurs systèmes (LEGAL, Flexible-LEGAL, GLUE, IGLUE, CIBLE) qui ont été évalués sur un ensemble de jeu de données. Cette partie présente également une méthode de transformation d'attributs symboliques en attributs numériques.

Enfin la troisième et dernière partie concerne les aspects interaction et applications de nos contributions. Nous montrons comment une technique de contrôle par objections permet à l'utilisateur-analyste d'interagir avec le système afin de valider les connaissances produites. Nous présentons ensuite une étude comparative d'opérateurs utilisés en dialogue avec ceux utilisés en apprentissage automatique. Nous terminons par une présentation d'applications, notamment en biologie moléculaire où plusieurs problèmes ont été abordés parmi lesquels: la prédiction de sites de jonction d'épissage, l'alignement de séquences protéiques et l'analyse de codage d'acides aminés.

**Mots clés :** fouille de données, apprentissage automatique,

transformation d'attributs, interaction homme-machine, bioinformatique.

**Contact :** Engelbert Mephu Nguifo,  
CRIL - IUT de Lens - Université  
d'Artois, rue de l'Université SP 16, 62307 Lens Cedex.  
**Email :** [mephu@cril.univ-arts.fr](mailto:mephu@cril.univ-arts.fr)  
Téléphone: 03 21 79 32 73  
Télécopie: 03 21 79 32 72  
**Adresse Web :** <http://www.cril.univ-arts.fr/~mephu>

## **Le modèle satisfaction-altruisme : coopération et résolution de conflits entre agents situés réactifs, application à la robotique, par Olivier Simonin**

**Thèse de l'Université de Montpellier II, soutenue le 20 décembre 2001 au LIRMM.**

### **Résumé.**

Cette thèse se situe à la rencontre du domaine des systèmes multi agents (SMA) et de la robotique mobile. Le modèle d'agent développé permet d'introduire des comportements coopératifs intentionnels dans l'approche de résolution collective de problèmes.

Après avoir analysé les situations/actions coopératives entre agents situés, nous définissons deux états de satisfaction évalués par l'agent. D'une part, la satisfaction personnelle, qui est une mesure de la progression des actions de l'agent. D'autre part, la satisfaction interactive qui évalue les actions de son voisinage : gêne, aide (potentielle), indifférence. Les agents peuvent alors s'influencer en émettant des signaux attractifs ou répulsifs. En réception, si ces influences sont plus intenses que la satisfaction personnelle, elles déclenchent des réactions altruistes.

L'architecture bâtie sur ce principe est évaluée par le traitement, en simulation, de problèmes physiquement distribués. Le problème des robots fourrageurs et celui des conflits spatiaux sont résolus efficacement par notre modèle, combinant auto-organisation et comportements coopératifs intentionnels. Le modèle est ensuite appliqué à un système hétérogène, puis il est étendu par l'introduction d'un module d'apprentissage exploitant les mesures de satisfactions.

Dans le cadre d'une implémentation réelle, nous proposons un protocole de communication générique dédié aux SMAs situés. Celui-ci établit, par une méthode de synchronisation indirecte, un multiplexage temporel assurant une fréquence de communication constante aux agents, et ce, quel que soit leur nombre. Enfin, le modèle proposé, ainsi que le protocole de communication, sont validés expérimentalement par la résolution de conflits spatiaux entre robots mobiles autonomes.

**Mots clefs :** Systèmes multi-agents, agents situés réactifs, coopération, communication dans les SMA situés, robotique mobile distribuée

**Contact :** Olivier Simonin,  
LIRMM, 161 rue Ada, 34392 Montpellier Cedex 5.  
**Email :** [simonin@lirmm.fr](mailto:simonin@lirmm.fr)  
**Adresse Web :** <http://www.lirmm.fr/~simonin>  
Téléphone : 04.67.41.86.0  
Télécopie : 04.67.41.85.00

## **Automatisation du raisonnement et de la rédaction de preuves en géométrie de l'enseignement secondaire, par Jean-Pierre Spagnol**

**Thèse en intelligence artificielle de l'université René Descartes Paris V soutenue à Paris le 12 octobre 2001.**

### **Résumé.**

L'objectif de cette recherche est la conception d'un système à base de connaissances, de nom ARGOS, capable de résoudre de façon automatique des exercices de géométrie niveau collège ou lycée. Le système est ensuite capable de produire une solution rédigée en français adaptée au niveau de la classe considérée.

L'utilisateur donne les connaissances au système de façon déclarative. Il donne aussi et simultanément la manière dont lui-même les expliquerait s'il avait à les utiliser dans un exercice. ARGOS est alors chargé d'opérationnaliser les définitions et théorèmes qu'on lui a fournis. Le système construit automatiquement des règles modélisant différentes utilisations possibles des connaissances mathématiques. Il construit des règles traduisant directement la propriété, des règles guidées par le but permettant au système de conclure plus rapidement en cas de présence d'une condition suffisante de la conclusion, des règles de conjecture permettant d'aborder un même problème suivant plusieurs points de vue et des règles de création permettant à ARGOS d'introduire dans la démonstration des objets n'ayant pas été créés de façon immédiate à partir de l'énoncé.

Quand une démonstration a été trouvée, ARGOS filtre la base des faits déduits pour ne conserver que les déductions utiles constituant le graphe de la preuve. Le système construit alors une preuve rédigée en français de la démonstration adaptée aux desiderata de l'utilisateur lui ayant fourni l'énoncé et le niveau de résolution souhaité.

Après chaque nouvel exercice résolu, le système met à jour une liste de règles, appelée stratégie, constituée de toutes les règles ayant déjà été utiles pour résoudre le même type de conclusion. Ces règles seront alors chargées prioritairement en cas d'un nouvel exercice de même type de conclusion.

Une utilisation au sein d'un EIAO (Environnement Interactifs d'Apprentissage avec Ordinateur) est envisagée couplée avec un logiciel de construction géométrique. Elle permettra à un élève de développer ses capacités de raisonnement, de conjecture et de rédaction de preuve par une interaction avec le système. L'élève doit construire de lui-même, à partir d'un livre ou de recommandations du professeur, les connaissances, les modèles de rédaction correspondants ainsi que l'énoncé de l'exercice fourni à ARGOS. Il est donc à l'initiative et impliqué dans la démarche de résolution de problèmes. Ce système permet une grande souplesse d'utilisation par son adaptation à chaque type d'utilisateur expert ou novice. Une interface permettant des échanges conviviaux entre l'utilisateur et le système est en phase de conception.

### **Mots clefs :**

automatisation du raisonnement en géométrie, EIAO, démonstration automatique de théorèmes, planification de preuve, explication.

**Contact :** Jean-pierre Spagnol,  
Université Paris 5, Crip5, équipe SBC, 45 Rue des Saints-Pères 75270 Paris Cedex 06.  
**Email :** [spaj@math-info.univ-paris5.fr](mailto:spaj@math-info.univ-paris5.fr)

## Changements de représentation, abstractions et apprentissages, par Jean-Daniel Zucker,

Habilitation à diriger des recherches de l'Université Pierre et Marie Curie, soutenue le 17 Décembre à Paris (LIP6).

### Résumé.

Mes recherches sont principalement organisées autour des changements de représentation pour l'apprentissage automatique. La capacité à adapter sa représentation en fonction de la tâche à résoudre « reconnue comme étant au coeur de l'intelligence humaine » est l'un des problèmes fondamentaux de l'Intelligence

Artificielle. Ce problème se pose avec d'autant plus d'acuité que les langages de représentations des concepts sont expressifs car les algorithmes d'apprentissage associés sont d'autant plus complexes. Dans ma thèse de doctorat, j'avais proposé une approche de l'apprentissage relationnel par des changements de représentation permettant de limiter la complexité de l'appariement.

Depuis 1996, j'ai étendu mes recherches à des problèmes complexes d'apprentissage relationnel, d'apprentissage situé et d'apprentissage multi-agent. L'approche générale que je propose est de combiner une abstraction (changement de la granularité des connaissances dans un même formalisme) à une reformulation (changement de formalisme de représentation). Cette approche s'inscrit dans le cadre d'une théorie située de l'abstraction à laquelle j'ai contribué.

Les originalités de cette théorie sémantique de l'abstraction sont : un rôle explicite de la notion de simplicité et une classification des abstractions basées sur des opérateurs d'abstractions. Ce cadre a été depuis utilisé avec un relatif succès dans plusieurs tâches d'apprentissage dans les domaines de : la généralisation cartographique, l'apprentissage situé de l'ancrage dans un robot autonome, l'apprentissage non-supervisé de caractères chinois, l'analyse de profils d'expressions de gènes impliqués dans l'obésité, l'aide à la découverte en physique granulaire, etc. Ce travail est une étape vers l'automatisation des changements de représentation pour les problèmes complexes d'apprentissage.

**Mots clés :** abstraction, changement de représentation, apprentissage situé, reformulation, propositionnalisation, représentation multi-instance, apprentissage multi-agent, risque structurel, aide à la découverte scientifique.

**Contact :** Jean-Daniel Zucker,

LIP6 8, rue du Capitaine Scott 75015 Paris.

**Email :** Jean-Daniel.Zucker@lip6.fr

**Adresse Web :** <http://www-poleia.lip6.fr/~zucker/HDR>

Téléphone : 01 44 27 72 07

Fax: 01 44 27 70 00

**Document récupérable :**

<http://terk.lip6.fr/~jdzucker/HDR/HDRZuckerTemp.pdf>

## Habilitation à diriger des recherches

### Titre : Des Règles d'Interaction comme Modèle de Contrôle en Univers Multi agent Par Jean-Luc Koning,

Présentée le 27 juin 2001

### Composition du jury :

*Président :*

Philippe Jorrand, Directeur de Recherche, CNRS

*Rapporteurs :*

Jeffrey Bradshaw, Research Scientist, University of West Florida (USA)

Suzanne Pinson, Professeur, Université Paris Dauphine

*Examineurs :*

Yves Demazeau, Chargé de Recherche, CNRS

Pascal Estrailier, Professeur, Université de La Rochelle

Nicolas Halbwachs, Directeur de Recherche, CNRS

### Résumé :

Les systèmes multi agents peuvent être vus comme un ensemble d'agents autonomes (programmes éventuellement mobiles) formant un tout structuré et collaborant en vue d'atteindre un objectif commun : permettre au système de remplir sa fonction. Le contrôle d'un tel système fait clairement appel à des mécanismes de communication qui organisent et structurent la société d'agents. Il est acquis que la coopération entre agents intelligents requiert une communication explicite pour leurs échanges de connaissances.

Ce travail s'est concentré sur cette question de l'interaction entre agents et l'a abordé selon plusieurs thèmes. Dans un premier temps il a contribué à la formalisation d'un langage d'interaction entre agents. Il s'intéresse ensuite à la question de l'ingénierie des protocoles d'interaction qui structurent l'échange des connaissances entre agents et également leur modélisation selon une approche à base de composant. De manière plus récente, il aborde la notion de contrainte sur les protocoles d'interaction via un modèle de calcul appelé POS qui est basé sur la sémantique opérationnelle.

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

Envoyer à Amedeo Napoli (adresse page 2 de ce Bulletin) un fichier texte ou un document word attaché contenant les informations suivantes :

Titre de la thèse  
Prénom et nom du candidat  
Type de la thèse  
(Thèse de l'Université de xxx, ou Habilitation à diriger  
des recherches de l'Université de xxx)  
Date et lieu de soutenance de la thèse

Mots clefs (5 maximum)  
Adresse postale du candidat  
Mél  
Adresse Web  
Téléphone et télécopie (facultatif)

Résumé (une demi-page)  
Adresse où le document peut être récupéré (le cas échéant)  
L'ensemble de ces informations doit tenir en 400 mots environ.

# APPELS À DOSSIERS

## Dossier « L'Espace, le temps et le changement »

Date de parution prévue : Juin 2002 (bulletin numéro 50)

Coordinateurs : Philippe Balbiani, IRIT, [balbiani@irit.fr](mailto:balbiani@irit.fr) ; Salem Benferhat, IRIT, [benferhat@irit.fr](mailto:benferhat@irit.fr)

### Objectifs

La problématique du traitement intelligent de l'information soulève la question de la représentation et du raisonnement sur le temps, l'espace et le changement.

Depuis des travaux précurseurs comme ceux d'Allen (temps), de Cohn (espace) et de Gärdenfors (changement), cette question constitue le foyer d'une intense activité de recherche. La tenue à Marseille en Septembre 2000 de l'école thématique « Documents et Evolution » du groupement de recherche "Information, Interaction, Intelligence" témoigne de la vitalité de cette activité en France.

L'objectif de cet appel est d'offrir une synthèse des développements récents des activités de recherche portant sur : le raisonnement temporel, le raisonnement spatial, la révision et la mise-à-jour des croyances, la fusion des informations multi-sources, et les applications aux systèmes d'informations géographiques, aux bases de données, au diagnostic, et à la planification, etc.

Merci d'envoyer, dès que possible, votre intention de participer à ce dossier, à l'adresse des coordonnateurs.

### Présentation des contributions

Les contributions devront être envoyées aux coordonnateurs avant le :

**15 mai 2002,**

par courrier électronique au format Word (.doc) ou (.rtf). Aucune mise en forme particulière ne doit être faite sur les textes autre que le gras, l'italique et la taille des caractères.

Deux types de contributions sont sollicités :  
- les descriptions d'activités,  
- les articles courts

### Les descriptions d'activité pourront être organisées suivant le schéma suivant :

- identification du chercheur ou de l'équipe, adresse, téléphone, courriel
- cadre général, mission et composition de l'équipe,
- bref historique des travaux,
- types de problèmes traités,
- panorama des travaux effectués, présentation d'un ou deux projets représentatifs,
- méthodes et outils employés, résultats obtenus,
- position par rapport à d'autres travaux, nationaux ou internationaux,
- bibliographie.

### Les « articles courts » seront organisés suivant le schéma suivant :

1. titre
2. auteur (contact)
3. description
4. courte bibliographie

### Présentation des contributions

Les contributions feront 1400 mots maximum (en raison de l'organisation spécifique de ce dossier, merci de respecter scrupuleusement cette contrainte), que ce soit pour la description d'activités ou pour les « articles courts » et devront être envoyées à :

Salem Benferhat : [benferhat@irit.fr](mailto:benferhat@irit.fr)  
avant le **15 mai 2002,**

par courrier électronique au format Word (.doc) ou RTF. Aucune mise en forme particulière ne doit être faite sur les textes autre que **gras**, *italique* et la taille des caractères.

**Adhésion individuelle et abonnement** Demande Renouvellement

Nom : Prénom :  
 Affiliation :  
 Adresse postale :  
 N° de téléphone : N° de télécopie :  
 Adresse électronique :  
 Activités (à titre professionnel / à titre privé (*rayez la mention inutile*)) :

	Consultation du bulletin sur WEB (pour une personne)	Envoi du bulletin papier + un accès pour consultation du bulletin sur WEB
<input type="checkbox"/> Adhésion simple :	30 Euros	60 Euros
<input type="checkbox"/> Adhésion étudiant (sur justificatif) :	15 Euros	30 Euros
<input type="checkbox"/> Adhésion de soutien	Sans objet	90 Euros
<input type="checkbox"/> Abonnement au bulletin sans adhésion	Sans objet	55 Euros

- Adhésion au collège *IAD-SMA* : ajouter 7,5 Euros pour les étudiants, 15 Euros pour les autres.  
 Adhésion au collège *Cafe (Apprentissage)* : gratuit.

**Adhésion Personne morale** Demande Renouvellement

Organisme :  
 Adresse postale commune aux bénéficiaires couverts par cette adhésion :

Le tarif d'adhésion comprend une partie fixe et une partie par bénéficiaire  
 Coordonnées des bénéficiaires (10 maximum) :

NOM, prénom	Mél.	Tél.	Fax

	Tarif de base fixe :	Tarif par bénéficiaire :
<input type="checkbox"/> Laboratoires universitaires	100 Euros	30 Euros
<input type="checkbox"/> Personnes morales non universitaires	300 Euros	30 Euros
<input type="checkbox"/> Adhésion de soutien	600 Euros	Sans objet

- j'accepte que les renseignements ci-dessus apparaissent dans l'annuaire de l'AFIA.  
 j'accepte que les renseignements ci-dessus soient transmis à l'ECCAI pour constituer un fichier européen.

**Veillez trouver un règlement (à l'ordre de l'AFIA) de ..... Euros**

**Trésorier AFIA :** Marc AYEL, LIA-Université de Savoie, 73376 Le Bourget du Lac cedex.

**Mode d'adhésion :**

De préférence, en ligne via le site Internet de l'AFIA : <http://www.afia-france.fr>

A défaut, cette page doit être envoyée au trésorier.

**Modes de paiement :**

- 1) par chèque, à l'ordre de l'AFIA, envoyé au trésorier.
- 2) par bon de commande administratif, à l'ordre de l'AFIA, envoyé au trésorier.
- 3) Par virement bancaire sur le compte de l'AFIA : Société Générale, Résidence du Val de Seine, 78430 LOUVECIENNES. Code banque 30003, code guichet 01902, numéro de compte 00037283856 clef RIB 3.

**TVA non applicable, article 293B du CGI**

## Les dossiers du *Bulletin de l'AFIA*

<p>IA et Médecine ..... Bulletin n°48          Fouille de données ..... Bulletin n°46/47          IA et document..... Bulletin n°44          IA et connexionisme..... Bulletin n°43          IA et Vie Artificielle ..... Bulletin n°42          IA et CHM..... Bulletin n°41          IA et EIAH ..... Bulletin n°40          Plates-formes multi-agents ..... Bulletin n°39          IA et WEB ..... Bulletin n°38          Mémoires d'entreprises ..... Bulletin n°36          IA et logique ..... Bulletin n°35          Ingénierie des connaissances ..... Bulletin n°34          IA et Télécommunications ..... Bulletin n°33          IA et Terminologie..... Bulletin n°32          Décision et IA..... Bulletin n°31          Raisonnement IA et Image ..... Bulletin n°30          Raisonnement temporel et spatial..... Bulletin n°29          Systèmes Multi-agents..... Bulletin n°28          IA et robotique ..... Bulletin n°27          I.A. et biologie moléculaire ..... Bulletin n°26          A. et droit ..... Bulletin n°25</p>	<p>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 ..... Bulletin n°17          Planification et action ..... Bulletin n°16          Traitement automatique des langues..... Bulletin n°15          I.A. et médecine ..... Bulletin n°14          Diagnostic à base de modèles..... Bulletin n°13          Validation des SBC ..... Bulletin n°12          Le connexionisme..... Bulletin n°11          I.A. et jeux ..... Bulletin n°10          E.I.A.O..... Bulletin n°9          I.A. et gestion..... Bulletin n°8          Conception et I.A..... Bulletin n°7          Intelligence artificielle distribuée ..... Bulletin n°6          Acquisition des Connaissances ..... Bulletin n°5          IA et ordonnancement ..... Bulletin n°4</p>
--	--

## SOMMAIRE DU BULLETIN N° 48

La vie de l'A.F.I.A. ....	4
Débats : Améliorer ou supprimer les referees .....	8
L'IA au Québec .....	12
Dossier « IA et Médecine » .....	15
Livres .....	50
Compte rendu de conférences .....	50
Calendrier .....	56
Sommaire des revues .....	69
Comptes rendus habilitations et thèses .....	73
Appels à dossiers .....	68

## CALENDRIER DE PARUTION DU BULLETIN DE L'AFIA

<p><i>Hiver</i>          Réception des contributions:  <b>15 décembre</b>          Sortie le <b>31 janvier</b></p>	<p><i>Eté</i>          Réception des contributions:  <b>15 juin</b>          Sortie <b>31 juillet</b></p>
<p><i>Printemps</i>          Réception des contributions:  <b>15 mars</b>          Sortie le <b>30 avril</b></p>	<p><i>Automne</i>          Réception des contributions:  <b>30 septembre</b>          Sortie le <b>31 octobre</b></p>