

BULLETIN

DE

L'AFIA



ASSOCIATION FRANCAISE POUR
L'INTELLIGENCE ARTIFICIELLE

OCTOBRE 2002

N° 52

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 : Marc Schoenauer
Projet FRACTALES - I.N.R.I.A.
Rocquencourt
B.P. 105 - 78153 LE CHESNAY
Cedex - France
Tel : +33 (0)139 63 50 87
Fax : +33 (0)139 63 59 95
Marc.Schoenauer@inria.fr
<http://www-rocq.inria.fr/fractales/Staff/Schoenauer>

Contributions au bulletin
Jean-Paul SANSONNET
voir ci-contre

Serveur WEB
<http://www.afia-france.org>

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

Jacques Pitrat, Jean-Paul Haton, Marie-Odile Cordier

Jean-Marc David, Daniel Kayser, Claude Vogel, Alain Colmerauer

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

Marc SCHOENAUER, président
Serge DUPUY, secrétaire
Marc AYEL, trésorier

Comité :

Jean-Paul BAQUIAST, Gilles BISSON, Bertrand BRAUNSCHWEIG,
Alain CARDON, Jérôme EUZENAT, Christophe JACQUEMIN,
Eunika MERCIER-LAURENT, Sylvie PESTY, Jean-Paul SANSONNET,
Michèle SEBAG, Laurent SIKLÓSSY, Fabien TORRE,
Jean-Daniel ZUCKER.

Comité de rédaction

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

Isabelle ROBBA
Rédactrice adjointe
LIMSI — CNRS
B.P. 133, 91403 ORSAY Cedex
Isabelle.Robba@limsi.fr

Catherine BARRY-GRÉBOVAL
Rubrique « Présentation de laboratoires »
LaRIA, Equipe Ingénierie des
Connaissances,
Université de Picardie Jules Verne
5, rue du Moulin Neuf
80000 Amiens
barry@laria.u-picardie.fr

Brigitte GRAU
Rubriques « Sommaires des revues »
et « petites annonces »
LIMSI — CNRS
B.P. 133, 91403 ORSAY Cedex
grau@limsi.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. Numérocopie – 01 46 31 64 53
ISSN 1273-1323
Dépôt légal octobre 2002

De nouveaux dossiers spéciaux pour le Bulletin de l'AFIA

Ce numéro du Bulletin de l'AFIA de la rentrée 2002 clôt l'année sur une publication de dossiers très riche : en janvier nous avons produit un gros dossier sur « IA et médecine » qui devient un sujet important dans le cadre des relations interdisciplinaires. Ensuite, en avril nous avons publié un n° spécial du Bulletin à l'occasion de la tenue de la conférence ECAI 2002 en juillet à Lyon: cela a été l'occasion de faire un recensement des nombreuses équipes d'IA, ou qui se réclament partiellement de l'IA, dans les laboratoires de recherche et de développement en France. Ce numéro reflète la vivacité de la recherche dans ce domaine et aussi l'essaimage dans de nombreux centres de recherche de la thématique IA. Au mois de juin nous avons publié un dossier sur la « représentation et le raisonnement sur le temps, l'espace et l'évolutif » qui devient un sujet crucial dans le cadre de la prise en compte des contextes dans les systèmes d'information 'intelligents'. Dans ce numéro d'octobre, Marie-Odile Cordier (IRISA) ainsi que Louise Travé-Massuyès (LAAS), Michel Dojat (U438, Inserm-UJF) et Sylvie Cauvin (IFP) sont les éditeurs d'un dossier sur « IA et Diagnostic » qui fait le point sur le sujet plusieurs années après les dossiers « Diagnostic à base de modèles » (voir n° 13) et « I.A. et temps-réel » (voir n° 17).

Après le succès qu'a rencontré le n° spécial d'avril 2002, nous allons poursuivre dans cette voie en essayant de proposer un n° spécial par an et nous envisageons de publier un n° spécial qui doit paraître en janvier 2003 qui porte sur « L'IA dans le RNTL » et qui sera édité par Bertrand Braunschweig (IFP) aidé de Jean-Luc Dormoy (EDF) et Patrick Gallinari (LIP6). Vous trouverez donc un appel à dossier à ce sujet en fin de ce bulletin dont voici un extrait « Le Réseau National des Technologies Logicielles existe depuis trois ans. L'appel à propositions 2003 sera donc le quatrième du nom. Les technologies IA sont très présentes dans la thématique du RNTL, et de nombreux projets comportent un volet IA consistant. Le thème II, traditionnellement, est un peu le thème « IA » du RNTL. Mais des projets des autres thèmes mettent également en œuvre certains outils de l'IA, de manière explicite – dans les objectifs - ou bien cachée – dans les moyens. Le but de ce dossier est de faire le point sur la présence des méthodes et outils de l'IA dans les projets RNTL existants. Par extension, les projets de deux autres réseaux, le RNRT (réseau national de recherche en télécommunications), et RIAM (Réseau pour la recherche et l'innovation dans l'audiovisuel et le multimédia) sont également invités à présenter leurs volets IA ».

Cependant il nous faut aussi remarquer que dans ce numéro certaines rubriques sont absentes : Débats, Synthèses, Livres. Nous sommes toujours à la recherche de collègues qui souhaitent nous présenter des synthèses, sur le modèle de l'excellent dossier de synthèse du mois de juin 2002 que Antoine Cornuéjols (LRI-IIIE) à proposé sur « Une nouvelle méthode d'apprentissage : Les SVM. Séparateurs à vaste marge ». Par ailleurs, Amédéo Napoli ne souhaite plus s'occuper de la rubrique « Revue de livres ». Nous le remercions chaleureusement pour tout le travail qu'il a accompli dans le bulletin de l'AFIA depuis de nombreuses années. Actuellement, nous recherchons de manière urgente un remplaçant (un appel a été lancé à ce sujet dans AFIA-INFO) dont le travail consisterait tout d'abord à lister les livres nouveaux qui présentent un intérêt pour la communauté IA au sens large et ensuite pour en faire 'critiquer' certains d'entre eux par des collègues. L'information sur les nouveaux livres pourrait alors être mise en ligne et accessible directement via le site de l'AFIA. Si vous êtes intéressés, pour en savoir plus sur les modalités (acquisition des livres, relations avec les éditeurs) contactez Jean-paul Sansonnet (jps@limsi.fr).

Jean-Paul Sansonnet

Rédacteur en chef du Bulletin de l'AFIA

Nouvelles du Québec

Nicole Tourigny

Je commence ma rubrique en vous soumettant l'appel à communications du plus grand rassemblement multidisciplinaire scientifique francophone qui a lieu chaque année dans une université canadienne francophone.

Vous trouverez à la suite 3 descriptions de travaux de Master (M. Sc.) qui ont porté sur le domaine de l'intelligence artificielle et ont été récemment soutenus au département d'informatique et de génie logiciel de l'Université Laval (<http://www.ift.ulaval.ca>). Les résumés sont issus des mémoires des étudiants. Ceci vous donnera une idée plus précise des sujets traités au département.

APPEL À COMMUNICATIONS

Comme chaque année au mois de mai, se tient le Congrès de l'ACFAS (Association Francophone pour le Savoir). Du 19 au 23 mai 2003, le 71^e aura lieu à l'Université du Québec à Rimouski sur le thème « Savoirs partagés ». Ce congrès est le plus grand rassemblement multidisciplinaire scientifique francophone.

L'ACFAS invite la communauté scientifique à lui soumettre des propositions de communication libre (date de tombée : 22 novembre 2002), de colloque, de colloque/atelier ou de forum (date de tombée : 11 novembre 2002) pour son 71^e Congrès. Les scientifiques de tous les domaines du savoir ont une occasion unique de présenter leurs plus récents travaux et de prendre connaissance des derniers développements en recherche. Ce congrès annuel, qui réunit des personnes

de tous les continents, est l'événement scientifique multidisciplinaire le plus important de toute la francophonie. Les participants et participantes y découvrent de multiples champs de recherche, raffermissent des liens de collaboration ou en créent des nouveaux. C'est aussi l'occasion de faire connaître plus largement au grand public les récentes découvertes.

Le congrès de l'ACFAS, c'est enfin un Salon des exposants, les assises de nombreuses associations, des lancements, des conférences, des débats et de nombreuses autres activités.

Pour plus d'informations, consulter le site : www.acfas.ca/congres

Titre du mémoire :

Raisonnement qualitatif dans les systèmes multiagents basé sur les cartes causales

Auteur du mémoire : Fabiola Mata Avila

Laboratoire DAMAS

<http://www.damas.ift.ulaval.ca>

Directeur de recherche : Dr Brahim Chaib-draa, Département d'informatique et de génie logiciel, Université Laval

Rapporteurs : Dr Nadir Belkhiter, Département d'informatique et de génie logiciel, Université Laval, Dr Laurence Capus, Département d'informatique et de génie logiciel, Université Laval

RÉSUMÉ. Les cartes cognitives ou cartes causales ont été mises en avant par les politiciens pour visualiser leurs alternatives politiques. C'est un modèle pour analyser, prédire et comprendre comment les gens prennent leurs décisions. D'une façon générale, les cartes causales sont une représentation graphique des croyances d'une personne, d'un agent ou d'un domaine particulier. Formellement, les croyances d'une personne par rapport à un sujet sont représentées dans ce genre de cartes

comme un graphe dirigé avec des concepts et des relations causales reliant ces concepts. Ce mémoire propose un modèle formel pour le raisonnement qualitatif dans des environnements multiagents fondé sur les cartes causales. La sémantique utilisée pour la manipulation des cartes causales est fondée sur l'algèbre causale. Cette recherche a porté plus particulièrement sur le développement d'un outil informatique permettant la création et l'analyse de cartes causales d'une manière

graphique, nommé Ma-CMa (Multiagent Causal Maps). Actuellement, il existe peu d'applications conçues spécialement pour traiter la prise de décision qualitative et moins encore pour être utilisées dans un environnement multiagent. Les quatre aspects suivants

ont été pris en compte : (1) la prise de décision qualitative d'un agent, (2) le raisonnement des vues subjectives d'un agent, (3) la combinaison des cartes cognitives entre agents pour agrandir leur base de connaissances, (4) le renforcement de connaissances entre agents. Ces

Titre du mémoire :

Métra-moteur de recherche sur le Web basé sur l'usage d'expressions et d'une ontologie

Auteur du mémoire : Kaïs Hefaiiedh

Laboratoire d'informatique cognitive

<http://www.ift.ulaval.ca/~moulin/>

Directeur de recherche : Dr Bernard Moulin, Département d'informatique et de génie logiciel, Université Laval

Rapporteurs : Dr Mamadou T. Koné, Département d'informatique et de génie logiciel, Université Laval

Dr Sylvain Delisle, Département d'Informatique, Université du Québec à Trois-Rivières

RÉSUMÉ. Depuis la création d'Internet, la technologie pour l'accès et la diffusion de l'information électronique sur le Web n'a cessé d'évoluer. Le développement d'Internet a permis l'accès à d'énormes quantités d'informations avec l'inconvénient majeur de submerger les usagers avec beaucoup d'informations non pertinentes. La croissance phénoménale du Web, du nombre de sources d'informations disponibles et la diversité de la nature des documents accessibles par Internet empêche l'utilisateur de trouver facilement l'information dont il a besoin, et ce malgré l'utilisation d'outils de recherche avancés. La recherche par mots-clés, qui est utilisée par la plupart des outils de recherche généralistes ou spécialisés, tend à identifier un nombre considérable de documents non pertinents. Cette recherche porte sur les besoins des usagers corporatifs qui cherchent des informations dans le cadre de leurs activités professionnelles. Le domaine d'application est celui de la santé environnementale. Une ontologie a été choisie pour décrire l'ensemble des connaissances de ce domaine. Elle est constituée d'un ensemble de catégories thématiques (thèmes et sous-thèmes) et d'un vocabulaire qui est utilisé par les usagers du domaine. Dans ce

contexte, un outil de recherche mettant en œuvre des agents logiciels a été développé. Il permet la méta-recherche de documents sur le Web en interrogeant des moteurs de recherche et des répertoires. La requête peut être formulée sous forme d'expressions qui peuvent être fournies par l'utilisateur ou choisies dans l'ontologie. Les agents logiciels récupèrent les résultats de la recherche sous forme de fichiers de type html. Ces fichiers sont traités en parallèle par les agents, de manière à analyser le code html afin d'extraire les adresses URL des résultats et le texte descriptif correspondant. Pour présenter l'information pertinente aux usagers après une recherche, un algorithme de filtrage à base d'expressions a été proposé. Cet algorithme analyse le texte qui décrit le document. Si le texte contient les expressions recherchées et s'il y a une distance acceptable entre les différentes expressions alors le document est sélectionné, sinon il est rejeté. Les URL redondantes sont supprimées des résultats, les autres sont présentées à l'utilisateur. Pour chaque URL, l'utilisateur peut consulter le texte qui décrit le document et consulter le document sur le Web. Enfin, un module a été développé pour évaluer l'outil.

Titre du mémoire :**Formation de coalitions d'agents logiciels pour le groupement d'acheteurs**

Auteur du mémoire : Kaïs Hefaiiedh

Auteur du mémoire : Frédérick Asselin

Laboratoire DAMAS

<http://www.damas.ift.ulaval.ca>

Directeur de recherche : Dr Brahim Chaib-draa, Département d'informatique et de génie logiciel, Université Laval

Rapporteurs : Dr Mamadou T. Koné, Département d'informatique et de génie logiciel, Université Laval, Dr Sylvain Delisle, Département d'informatique, Université du Québec à Trois-Rivières

RÉSUMÉ. La généralisation est un processus long et coûteux et beaucoup de travaux de recherche sont en cours pour tenter de l'automatiser. Une des limites à l'automatisation vient de la difficulté à formaliser les connaissances des cartographes employant la généralisation. Ce travail de recherche avait pour objectif de modéliser plus particulièrement le processus de généralisation conceptuelle qui consiste à produire une nouvelle base de données géographiques à partir d'une base de données géographiques existante de plus grande échelle.

Pour cela, un cas pratique de généralisation conceptuelle a été étudié, sur lequel ont travaillé les cartographes du Ministère des Ressources Naturelles du Québec (MRNQ). À partir de cette analyse, un graphe causal a été établi pour modéliser la généralisation conceptuelle. Enfin, le processus de généralisation conceptuelle a été simulé avec le langage PROLOG+CG afin de montrer comment le modèle proposé est un pas vers l'automatisation.

En ce début d'automne, je vous souhaite une bonne rentrée. Quant à nous, nous sommes déjà rentrés depuis le 3 septembre et la session est bien entamée. Il me fait toujours plaisir de répondre à vos questions et vos commentaires sont grandement appréciés.

Nicole Tourigny, Ph.D., professeur
Département d'informatique et de génie logiciel
Pavillon Adrien-Pouliot
Université Laval
Québec, CANADA
G1K 7P4
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	Société ISOFT	Bulletin n°16
LRI, Orsay	Bulletin n°1	Dépt. d'Info de l'Université d'Ottawa	Bulletin n°16
Service Systèmes Experts, Renault	Bulletin n°1	Equipe CHM, Université du Colorado (USA) .	Bulletin n°17
CEDIAG,	Bulletin n°2	LIRMM, Montpellier	Bulletin n°19
CERT, ONERA, Toulouse	Bulletin n°2	Institut autrichien de recherches en I.A.	Bulletin n°20
IRIT, Toulouse	Bulletin n°2	ENST Bretagne	Bulletin n°21
LAAS, Toulouse	Bulletin n°2	LIA - Université de Savoie.....	Bulletin n°22
HEUDIASYC, UTC.....	Bulletin n°3	INRETS	Bulletin n°23
IFP, Rueil Malmaison	Bulletin n°3	IRIN Nantes.....	Bulletin n°24
DIAM, INSERM U194	Bulletin n°3	CRIN - INRIA Lorraine	Bulletin n°25
Lab. Math. Info., Fac Médecine de Marseille..	Bulletin n°4	DIRO - Université de Montréal.....	Bulletin n°26
GMD, St. Augustin (RFA)	Bulletin n°4	IRIT - Toulouse (1)	Bulletin n°28
ONERA, Chatillon	Bulletin n°4	IRIT - Toulouse (2)	Bulletin n°29
KSL, Université de Stanford (USA)	Bulletin n°5	LAAS - Toulouse (1).....	Bulletin n°30
Dépt Applications de l'IA au CNET, Lannion.	Bulletin n°5	Sony CSL	Bulletin n°31
LAFORIA, Univ. Pierre et Marie Curie.....	Bulletin n°6	LAAS - Toulouse (2).....	Bulletin n°32
L'institut FAW, ULM (RFA)	Bulletin n°6	LIMSI - Département CHM.....	Bulletin n°33
Institut IIIA, Compiègne	Bulletin n°6	LAMSADE.....	Bulletin n°34
LAIR, OHIO State University (USA).....	Bulletin n°7	Institut autrichien de recherches en I.A.	Bulletin n°36
ARAMIHS, Labo mixte MATRA-CNRS,	Bulletin n°7	LIP6 (Univ. Pierre et Marie Curie)	Bulletins n°37 & 38
CEA, Service SERMA, Saclay	Bulletin n°8	GREYC (Univ. de Caen).....	Bulletin n°40
Société ILOG.....	Bulletin n°8	LIFL (Univ. de Lille).....	Bulletin n°41
LAIAC, Université de Caen	Bulletin n°9	LRI (équipes IA et IASI)	Bulletin n°43
Institut Français du Pétrole	Bulletin n°10	IMAG - Grenoble	Bulletin n°44
DFKI (Centre allemand de recherches en IA) .	Bulletin n°11	PSI	
GRTC, Marseille	Bulletin n°11	(Perception, Système, Information - Rouen) ...	Bulletin n°45
Inst. d'Analyse des Systèmes, Ac. Russe.....	Bulletin n°12	INRIA – Sophia Antipolis	Bulletin n°46/47
Georges Mason Univ., Center for AI (USA)....	Bulletin n°13	LIH (Labo d'Informatique du Havre).....	Bulletin n°46/47
IRISA, INRIA et Université de Rennes	Bulletin n°13	Tech-CICO	
Société INGENIA	Bulletin n°14	(Univ. de Technologie de Troyes)	Bulletin n°51
LIPN, Université de Paris Nord	Bulletin n°14	LIFO (Univ. d'Orléans – Équipe Contraintes et	
Institut EURISCO	Bulletin n°15	Apprentissage)	Bulletin n°52
LRDC, Université de Pittsburgh (USA).....	Bulletin n°15	LIIA (École Nationale des Arts et Industries	
		de Strasbourg – ENSAIS)	Bulletin n°52

**Equipe Contraintes
et Apprentissage
LIFO – Laboratoire
d’Informatique Fondamentale
d’Orléans
Université d’Orléans**

Responsable : Christel Vrain,

Contacts :

Gérard Ferrand, Christel Vrain
{gerard.ferrand, christel.vrain}@lifo.
univ-orleans.fr

Mots clés : Apprentissage Automatique, Programmation par Contraintes, Programmation en Logique, Diagnostic Déclaratif d’Erreurs, Consistance, Itérations Chaotiques, Classification, Règles d’association, Regroupement, Caractérisation, Programmation Logique Inductive (PLI), Algorithme Génétique (AG), Extraction de Connaissances dans les Bases de Données (ECD), Fouille de Textes

Liste des membres :

Permanents : Jean-Claude Bassano, Thi Bich Hanh Dao, Ali Eddbali, Matthieu Exbrayat, Gérard Ferrand, Sylvie Guilloré-Billot, Arnaud Lallouet, Frédéric Moal, Lionel Martin, Alexandre Tessier, Christel Vrain, Viviane Clavier (membre associé)

Doctorants : Julien Arsouze, Agnès Braud, Guillaume Cleuziou, Isabelle Debourges, Gaël Dias, Willy Lesaint, Nicolas Roméro, Ansaf Salleb, Teddy Turmeaux

1. Thématique

Dans l’équipe *Contraintes et Apprentissage* du LIFO, on s’intéresse à la sémantique et à l’apprentissage de systèmes de règles exprimées dans un formalisme relationnel. Les domaines d’application sont, par exemple, la mise au point de programmes dans le cadre

de la Programmation par Contraintes ou encore l’extraction de connaissances dans des bases de données relationnelles ou avancées (travaux en cours d’une part, dans un projet RNTL - Réseau National des Technologies Logicielles, d’autre part, en collaboration avec le BRGM - Bureau de Recherche Géologique et Minière). Les méthodes développées reposent principalement sur des mécanismes de déduction ou d’induction étudiés dans un cadre logique où la notion de contrainte joue un rôle prépondérant.

La Programmation par Contrainte est un style de programmation déclaratif dans lequel le programmeur définit les objets du discours (les variables) et spécifie les relations entre ces objets (les contraintes). Le programmeur attend du système, appelé solveur, une solution, c’est à dire une affectation des variables satisfaisant toutes les contraintes. Prolongeant la vision déclarative qui a fait l’un des intérêts de la Programmation en Logique, la Programmation par Contraintes propose un champ d’action plus large aussi bien en ce qui concerne la modélisation que du point de vue de l’efficacité des techniques de résolution mises en oeuvre. La spécificité de la programmation par contrainte consiste à épauler les algorithmes de recherche arborescente traditionnels par des techniques puissantes d’élagage connues sous le nom de « consistances ». Leur efficacité est telle que de nombreux problèmes peuvent être résolus de façon exacte, c’est à dire en prouvant l’optimum ou le cas échéant l’absence de solution.

En ce qui concerne la mise au point de programmes dans le cadre de la Programmation par Contraintes, la notion classique de trace s’avère difficilement utilisable à cause de l’imbrication complexe des différents processus qui concourent à la résolution d’un système de contraintes. Un niveau d’abstraction convenable est indispensable, en particulier l’usage d’outils s’appuyant

sur une sémantique déclarative, c’est à dire indépendante d’une stratégie particulière de résolution.

Pour la résolution de contraintes dans les domaines finis, le calcul consiste en un entrelacement de recherche arborescente (par exemple labeling) et de réduction de domaine. Dans la réduction de domaine l’événement fondamental est le retrait de valeur c’est à dire l’élimination d’un élément d’un domaine, élément inconsistant en un certain sens. Nous avons étudié une représentation déclarative de cet événement sous la forme d’un arbre de preuve appelé explication de retrait de valeur. Cette notion d’explication doit ensuite être intégrée dans une modélisation complète de la résolution de contraintes fondée sur une formalisation convenable d’un arbre de recherche.

En Apprentissage, nos recherches se situent principalement en Programmation Logique Inductive (PLI), domaine qui regroupe les travaux sur l’apprentissage de connaissances exprimées dans un formalisme relationnel. Dans ce cadre, nous avons tout d’abord travaillé sur la tâche de classification, que l’on peut définir comme l’apprentissage de définitions en intention de concepts, à partir d’exemples étiquetés (apprentissage supervisé) et de connaissances de base sur le domaine. La collaboration débutée en 1999 avec le BRGM nous a conduit à nous intéresser aussi à l’apprentissage de règles d’association et de règles de caractérisation.

On peut envisager deux types d’application à la PLI : la synthèse de programmes logiques d’une part, l’apprentissage de connaissances exprimées dans un langage du 1er ordre d’autre part. Dans la mesure où la Programmation Logique a évolué vers la Programmation par Contraintes, les travaux menés en PLI doivent être étendus pour induire des programmes logiques avec contraintes. Pour la seconde application, il est primordial de

pouvoir traiter non seulement des données symboliques, mais aussi numériques, ce qui nécessite aussi de développer les méthodes existantes en PLI. Les extensions nécessaires ont, dans les deux cas, comme point commun, la prise en compte de domaines autres que celui de Herbrand. Les problèmes que nous avons abordés sont de deux natures : définir un cadre formel pour étudier les problèmes de généralisation et spécifier des systèmes hybrides couplant des techniques d'apprentissage classiques en Programmation Logique Inductive et des techniques d'induction numérique. Un cadre naturel pour étudier formellement les problèmes posés par l'extension de la PLI est celui de la Programmation Logique avec Contraintes. Parallèlement, dans la mouvance des recherches sur les contraintes, est apparu le concept de base de données contraintes, extension naturelle des bases de données relationnelles, mais aussi des bases de données déductives et spatiales : à la notion classique de tuple en base de données relationnelles, est substituée la notion de tuple généralisé, vu comme une conjonction de contraintes. D'un point de vue formel, ce concept permet de capturer dans un même modèle la notion de contraintes et celle de bases de données.

L'extraction de connaissances à partir de bases de données (ECD) est un domaine actuellement très actif dans lequel l'apprentissage relationnel joue un grand rôle, car il prend en compte la nature relationnelle des données, alors que beaucoup d'outils existants nécessitent un aplatissage préalable de la base de données, perdant ainsi toutes les informations sur sa structure. Cependant, il faut prendre en compte la taille de l'espace de recherche, le volume de données à traiter et le nombre important de requêtes à exécuter. Nous nous intéressons d'une part aux algorithmes évolutifs qui permettent une exploration stochastique de l'espace de recherche et à la parallélisation de ces

algorithmes qui apparaît comme une solution prometteuse.

En ce qui concerne les applications, nous collaborons actuellement avec le service Ressources Minérales du BRGM qui conçoit et développe des systèmes d'information géographiques (SIG) à vocation métallogénique, systèmes d'information qui doivent être exploités pour définir des cartes de potentiel métallogénique. Ce service dispose donc de grandes quantités de données, issues de programmes de recherche en cours (SIG Andes et SIG Afrique). Des outils ont déjà été développés, fondés principalement sur des méthodes statistiques. Les approches symboliques comme celles développées au LIFO permettent d'exploiter des données qualitatives et d'extraire des connaissances exprimées sous forme de règles et donc plus facilement interprétables par un expert ; elles apparaissent comme complémentaires des méthodes statistiques. Cependant, appliquer nos travaux aux SIGs pose de nouvelles contraintes, comme la prise en compte de données spatiales ou la complexité liée aux nombreuses couches thématiques.

Un autre domaine où l'apprentissage automatique peut jouer un rôle important est celui du traitement automatique de la langue et de la fouille de textes. Dans ce cadre, nous utilisons des méthodes de recherche statistiques de récurrences existantes dans un ensemble de textes et nous abordons aussi la tâche de regroupement.

Nous démarrons aussi un thème de recherche consistant à construire un solveur des contraintes par apprentissage des règles de propagation. En effet, la conception des solveurs est une activité difficile rendue de plus en plus complexe par l'émergence de contraintes globales adaptées à des problèmes spécifiques comme l'ordonnancement.

Enfin, on peut noter que cette année,

nous avons organisé à Orléans la conférence francophone d'apprentissage Cap 2002.

2. Détail des travaux

Modélisation des solveurs de contraintes sur les domaines finis utilisant les techniques de réduction de domaine. Il existe déjà une formalisation théorique de la réduction de domaine par les itérations chaotiques d'un ensemble d'opérateurs corrects, croissants et contractants. La réduction de domaine est décrite simplement par des notions de point-fixe et de clôture. Nous avons développé deux extensions de ce formalisme : les g-itérations chaotiques permettant de modéliser les algorithmes de propagation et les dérivations chaotiques qui décrivent l'entrelacement des étapes de recherche et de consistance dans les CSP [18]. Ce cadre théorique est commun aux travaux de l'équipe et plusieurs directions d'applications sont étudiées, concernant notamment le débogage déclaratif, le parallélisme ou l'apprentissage. Nous nous intéressons particulièrement au cas où les opérateurs sont spécifiés de façon déclarative par un système de règles (comme dans le système Gnu-Prolog), plutôt que programmés dans un langage impératif.

Débogage déclaratif de programmes avec contraintes. C'est l'un des volets du projet RNTL OADymPPaC (Outils pour l'Analyse Dynamique et la mise au Point de Programmes avec Contraintes) dans lequel nous sommes impliqués. Ce projet regroupe quatre partenaires académiques: l'INRIA-Rocquencourt, l'École des Mines de Nantes, l'INSA-IRISA-Rennes et le LIFO, et deux partenaires industriels: COSYTEC et ILOG.

En Programmation par Contraintes, le contrôle est dirigé par les interactions entre contraintes au cours de la propagation. La mise au point de programmes avec contraintes n'est donc pas aisée et les outils de débogage pratiquement

inexistants. Le travail effectué dans l'équipe a pour objectif principal la définition de techniques de mise au point reposant sur les explications de retrait de valeurs. Nous avons proposé une notion d'explication arborescente proche des justifications utilisées dans les TMS que nous allons utiliser comme fondement pour le diagnostic d'erreur [15]. Ces travaux ont pour but d'être intégrés dans le système Gnu-Prolog. En collaboration avec l'école des Mines de Nantes, nous avons également utilisé ces explications pour effectuer la preuve de correction d'algorithmes de retrait de contraintes utilisés en Programmation par Contraintes dynamique [10]. Les recherches actuelles se centrent autour de deux sujets. D'une part, il est nécessaire de prolonger la modélisation afin d'inclure un langage hôte du type CLP(FD). Ceci a été partiellement réalisé en modélisant la construction de l'arbre de recherche CLP comme un système de contraintes dont on recherche les solutions, qui sont en fait des arbres de preuve pour la partie programme logique [16]. D'autre part, la définition de notions de symptômes et d'erreurs sera utile afin d'utiliser les explications pour le débogage déclaratif de programmes avec contraintes. Le retrait inattendu d'une valeur (appelé symptôme) sera la racine d'une explication. Le diagnostic déclaratif consistera alors à remonter dans cet arbre jusqu'à atteindre l'erreur qui a causé ce symptôme.

Enfin, nous avons participé au développement de l'environnement de programmation hypertexte pour la programmation en logique et par contraintes, nommé HyperPro, dans le cadre d'un projet franco-brésilien [11].

Caractérisation en Programmation Logique de la réduction de domaines. La modélisation déclarative des opérateurs de consistance nous a amené à étudier le parallèle avec la Programmation en Logique. Nous avons proposé une transformation d'un CSP sur lequel on souhaite calculer l'arc-consistance en un

programme logique avec négation qui est telle que les états consistants sont exactement les modèles stables trivalués du programme logique [14].

Apprentissage de solveurs. La vision du calcul des consistances locales comme enchaînement de règles nous a naturellement conduit à la question de l'apprentissage automatique de ces règles, d'autant plus que la conception des solveurs est une activité difficile rendue de plus en plus complexe par l'émergence de contraintes globales adaptées à des problèmes spécifiques comme l'ordonnancement.

A partir d'une contrainte décrite en extension, nous avons proposé un algorithme d'apprentissage permettant de construire les règles utilisées dans Gnu-Prolog, appelées indexiques [6]. Pour cela, nous avons considéré l'ensemble de toutes les réductions possibles effectuées par la contrainte pour poser un nouveau problème de contraintes dont la solution optimale est un opérateur qui est retourné. Cet opérateur existe pour toute contrainte définie par des tuples. Notre système permet de retrouver les solveurs de Gnu-Prolog sur les exemples que nous avons testés et permet par exemple de construire un solveur ad hoc pour une contrainte globale définie par l'utilisateur. Nous étudions actuellement plusieurs algorithmes pour cette voie prometteuse.

Théorie des arbres et contraintes du 1^{er} ordre. En collaboration avec Alain Colmerauer (Laboratoire d'Informatique Fondamentale de Marseille), nous étudions des théories (logiques) complètes et décidables, qui sont fondamentales en informatique. Nous avons étudié la théorie des arbres finis ou infinis sur un ensemble infini de symboles de fonction. Nous nous intéressons ensuite à la théorie des arbres avec queues, la théorie des arbres binaires infinis. Parallèlement, nous étudions des algorithmes et leurs complexités pour résoudre dans ces théories des contraintes du premier ordre. Ces

contraintes sont écrites sous forme de formules du premier ordre, avec les connecteurs logiques « et », « ou », la négation, les quantificateurs existentiels et universels et des variables libres. Pour les contraintes du premier ordre avec l'égalité dans la théorie des arbres finis ou infinis, nous avons proposé un algorithme de résolution sous forme de règles de réécriture de formules [7]. Nous nous intéressons aussi au pouvoir d'expression de ces théories. Nous avons montré que dans la théorie des arbres, avec une formule de taille proportionnelle en k , on peut représenter un arbre fini dont le nombre de nœuds est une tour d'exponentielles de hauteur k , ce qui implique la complexité non élémentaire de la théorie des arbres [5].

Parallélisme et contraintes. Nous avons abordé le thème du parallélisme sous deux angles différents. Du côté langage, nous avons proposé le domaine de contraintes des vecteurs distribués permettant de construire un CLP data-parallèle. Dans ce domaine, une variable représente un vecteur distribué qui est manipulé globalement. La communication peut ainsi être vue comme une contrainte d'égalité à distance entre éléments de deux vecteurs situés sur des lieux différents [1]. Cette proposition de langage possède une interprétation intuitive dans le modèle de parallélisme BSP. Du côté algorithme parallèle de résolution, en collaboration avec L. Granvilliers de l'IRIN (Nantes), nous avons proposé un algorithme distribué symétrique de calcul de consistance incluant un équilibrage de charge [17].

Etude formelle de l'apprentissage dans le cadre de la Programmation par Contraintes. Afin d'étendre le cadre classique de la programmation logique inductive, nous avons défini un ordre de généralité dans l'espace des clauses contraintes, où une clause contrainte est une clause définie à laquelle on ajoute une contrainte sur un domaine bien défini, comme par exemple le domaine des entiers naturels.

A l'intérieur de cet ordre, nous nous sommes intéressés au calcul du plus petit généralisé et du plus grand spécialisé, notions classiquement employées dans les systèmes de PLI et nous avons défini des opérateurs de raffinement qui permettent de généraliser ou de spécialiser une hypothèse afin d'expliquer un ensemble d'exemples [26]. Nous avons aussi étudié l'apport des bases de données contraintes à cette problématique [25].

Apprentissage et Extraction de Connaissances dans les Bases de Données.

Tâche de classification. Le but est de trouver une définition en intention d'un concept représenté par un ensemble d'exemples étiquetés comme appartenant ou non à ce concept. Cette problématique, qui peut être abordée comme un problème de recherche, est en général particulièrement difficile dans un contexte relationnel [21]. Deux approches sont alors possibles : la réduction de la taille de l'espace de recherche en utilisant des biais de langage ou l'utilisation de méthodes approchées pour n'explorer qu'une partie de l'espace de recherche. Nous nous sommes intéressés à ces deux approches, en étudiant d'une part les formalismes de spécification de biais de langage dans un cadre relationnel, en proposant un nouveau formalisme fondé sur les grammaires d'arbres et un langage de contraintes, et en étudiant d'autre part l'utilisation de méthodes stochastiques de recherche, en particulier la programmation génétique et les algorithmes génétiques. Dans le premier cas, on peut représenter directement sous forme d'arbres des connaissances relationnelles [20], alors que dans le second cas, il est difficile de trouver une représentation adéquate des individus, car la forme de la règle recherchée n'est pas connue. Dans des systèmes d'apprentissage relationnels fondés sur les algorithmes génétiques, un modèle est donc préalablement construit pour la règle et donné au système, ce qui permet de

reformuler un problème du premier ordre en un problème attribut-valeur. Cette solution est aussi utilisée en PLI pour faire face aux problèmes de complexité rencontrés en logique du premier ordre et est appelée *propositionnalisation*. La propositionnalisation comporte deux phases, la première consiste à trouver un schéma intéressant pour la règle à apprendre et la seconde à apprendre des contraintes pertinentes pour ce schéma. Nous avons plus particulièrement travaillé sur la seconde étape et avons proposé une approche originale, inspirée des algorithmes génétiques. L'idée principale est de considérer l'ensemble des variables apparaissant dans le schéma et d'apprendre une partition de cet ensemble en classes d'équivalence regroupant les variables égales. Les contraintes sont alors placées directement sur ces classes d'équivalence. Nous avons proposé pour cela un nouveau codage et un ensemble d'opérateurs, inspirés des opérations ensemblistes, pour altérer une partition ou en combiner deux [2]. Nous avons conçu un système reposant sur ces idées, qui a été implanté en C++ et en utilisant le SGBD PostgreSQL.

Définition de distance. L'élaboration de mesures de similarité pour la classification et/ou le regroupement d'objets par similitude est un sujet important. L'idée sous-jacente de nos travaux est que la similarité entre deux objets ne dépend pas seulement des valeurs des attributs des objets, mais plus particulièrement du contexte d'apprentissage, c'est à dire de l'ensemble des définitions de concepts pertinentes pour le problème considéré. Dans cette approche, la mesure de similarité est spécifiée par un langage défini par l'utilisateur à l'aide d'une grammaire. Chaque terme du langage représente une propriété satisfaite ou non par les objets. La similarité entre deux objets est alors définie par la probabilité qu'une propriété (terme), tirée aléatoirement dans le langage, soit satisfaite ou rejetée simultanément par ces deux objets. Cette définition nécessite

donc de définir un espace probabilisé sur un langage, ce qui peut être difficile pour certains langages infinis. Pour cette raison, nous proposons d'approcher cette mesure en la restreignant à une partie finie du langage, générée de façon stochastique à partir de la grammaire [19]. Cette approche offre un cadre général pour différentes mesures de similarité et de distances usuelles, permettant de traiter des formalismes variés comme l'attribut-valeur (avec des valeurs numériques et/ou symboliques) ou la logique du premier ordre. Par ailleurs, elle intègre naturellement la notion de biais de langage au moyen de la spécification grammaticale du langage. De plus, elle ne nécessite pas la définition d'une mesure de similarité locale entre les attributs des objets. Cette méthode peut être appliquée à la fois dans un cadre supervisé comme la classification, et non supervisé comme le groupement par similitude. L'évaluation empirique sur des problèmes classiques de classification montre des résultats prometteurs.

Extraction de règles d'association. L'extraction des règles d'association est l'un des principaux problèmes de l'ECD, introduit par Agrawal et al. en 1993 pour l'analyse de bases de données de transactions de ventes. Chaque transaction est constituée d'une liste d'articles achetés (items). Le but est de découvrir les groupes d'articles (itemsets) achetés le plus fréquemment ensemble, et donc de mettre à jour des relations significatives entre les données. Le processus d'extraction des associations peut se décomposer en deux principales étapes : découverte des itemsets fréquents et génération des règles d'association. La phase de découverte des itemsets fréquents est la plus coûteuse en temps d'exécution car le nombre d'itemsets fréquents est exponentiel en fonction du nombre d'items de la base de données. Dans ce contexte, nous proposons un nouvel algorithme *BoolMiner*, qui permet d'extraire tous les itemsets fréquents maximaux [22,23]. Il est fondé sur une

approche booléenne du problème dans la mesure où la base de données est vue comme une table de vérité à n variables, n étant le nombre d'items possibles de la base de transactions. A cette table de vérité, on associe une fonction de sortie qui donne pour chaque ligne représentant une combinaison d'items possible, le nombre d'instances de cet itemset dans la base de données. La fonction de sortie représente alors une forme condensée de la base de données qui peut être chargée en mémoire en vue de l'exploration des itemsets fréquents. Elle est représentée par un graphe dirigé acyclique, emprunté aux résultats de compression des circuits logiques. Cette approche minimise le nombre de parcours coûteux de la base de données. Nous avons développé un prototype et nos premières expérimentations montrent qu'il est possible d'utiliser ce format condensé pour l'extraction des itemsets fréquents en général et maximaux en particulier sur divers types de bases de données en des temps raisonnables. Par ailleurs l'approche booléenne que nous proposons présente un autre intérêt, en effet, elle s'étend facilement à l'extraction des itemsets fréquents avec négation.

Parallélisation et apprentissage.

Les algorithmes d'apprentissage devant traiter des volumes de données de plus en plus importants, leurs performances en terme de temps s'en trouvent considérablement diminuées. La parallélisation de ces algorithmes apparaît comme une solution prometteuse à ce problème, et nous nous sommes intéressés à la parallélisation des algorithmes génétiques. Cette étude est fondée sur le modèle BSP (Bulk Synchronous Parallelism). Deux algorithmes parallèles ont été étudiés et implantés dans le cadre d'un problème simple d'ECD où une représentation attribut-valeur est suffisante. Dans la première la population des individus est répartie sur les processeurs [3] ; dans la seconde, visant plus spécialement les tâches d'ECD, c'est la base de données qui est répartie

[4]. Les formules de coût obtenues correspondent à des accélérations presque linéaires en fonction du nombre de processeurs. Ces études ont donné lieu à l'implantation en C et en utilisant la librairie BSPLib de deux prototypes, et les expérimentations réalisées confirment les résultats théoriques obtenus.

Application, en collaboration avec le BRGM, à l'Extraction de Connaissances dans des Systèmes d'Information Géographiques

Le but de cette collaboration avec le service Ressources Minérales du BRGM était de découvrir les relations de corrélation entre les diverses couches d'un Système d'Information Géographique (SIG). Nous avons tout d'abord implanté un prototype ArGIS qui permet d'apprendre de manière non supervisée des règles d'associations entre les couches [24]. Les règles trouvées ont été confrontées aux experts qui les ont jugées fort intéressantes tant au niveau de la compréhension même de leurs données que dans l'aide à la décision. Les règles obtenues ont été classées en trois types :

- *des règles statistiques* qui donnent la répartition des attributs parmi les objets géographiques ;
- *des règles de contrôle* qui permettent aux experts de vérifier grâce à leur expérience des corrélations connues dans leurs données et donc de vérifier leur cohérence ;
- *de nouvelles règles* qui donnent des corrélations surprenantes ou intéressantes que les experts n'auraient pas vues entre les couches thématiques.

Cependant, le processus a engendré un nombre important de règles, ce qui nous a poussé à orienter nos recherches vers d'autres types de règles dont le nombre est plus restreint, qui sont les règles de caractérisation.

Traitement automatique des langues. Après s'être intéressé à la recherche d'informations, nos travaux s'orientent vers l'extraction de connaissances à partir de textes, en nous fon-

dant principalement sur une recherche statistique de récurrences existantes dans un ensemble de textes : associations polylexicales [12,13], concepts et relations entre concepts. Dans ce dernier cadre, le problème que nous avons abordé est celui de la recherche de connaissances dans un corpus, recherche guidée par la requête d'un utilisateur. Nous avons d'abord cherché les concepts liés aux mots clefs de la requête, en nous fondant principalement sur la recherche de co-occurrences [8]. Nous étudions maintenant comment des techniques d'apprentissage pourraient être appliquées pour étiqueter les liens entre les concepts [9]. Actuellement, nous pensons que cette recherche de phénomènes récurrents dans un corpus de textes peut être étendue à différents indices linguistiques, soit de forme, soit plus « sémantique » et donner lieu à des études sur le style d'écriture, sur le genre des textes, le profil des lecteurs potentiels, etc. Dans ce cadre, il nous semble que les outils proposés par l'Apprentissage Automatique associés à des observations de corpus par des linguistes sont une bonne voie d'exploration.

Nous abordons aussi l'organisation thématique d'unités textuelles, sous ses deux principaux aspects que sont la définition puis l'étude de mesures de similarité et l'élaboration de techniques de regroupement (*clustering*). Dans une première phase, nous nous intéressons à l'organisation de mots, et nous devons proposer des méthodes capables de construire des groupes de mots proches sur le plan sémantique. Il existe de nombreuses définitions de mesure de similarité entre mots qui peuvent être regroupées en trois catégories : les méthodes symboliques (fondées sur une analyse syntaxique), des méthodes utilisant des connaissances extérieures (thésaurus) et des méthodes statistiques (fondées sur la notion de cooccurrence dans les documents). Nous avons effectué une étude comparative de ces différentes mesures pour sélectionner les plus pertinentes pour notre cadre d'étude. Ces travaux

nous ont permis d'étudier différentes techniques usuelles de regroupement et de mettre en évidence leurs limites dans le domaine du traitement d'unités textuelles [27]. Ainsi, nous avons proposé dans un premier temps une amélioration de méthodes agglomératives et dans un deuxième temps, nous avons proposé une méthode originale fondée sur la recherche de cliques dans un graphe de cooccurrence et permettant ainsi de construire des groupes non disjoints.

Actuellement, d'autres thèmes de recherche débutent en parallèle des travaux mentionnés dans la section précédente. On peut citer :

La recherche de règles de caractérisation. La caractérisation est une tâche supervisée de fouille de données permettant de résumer de manière succincte et concise un ensemble de données. Etant donné un ensemble cible d'objets à décrire (par exemple des gisements d'or), le but de la caractérisation est de trouver toutes les caractéristiques communes de ces objets, en se basant non seulement sur leurs propriétés propres mais aussi sur les propriétés de tous leurs voisins. Cette notion de voisinage est naturellement à définir, elle peut être définie par des notions topographiques ou par des distances dans le cas de bases de données géographiques. Mais, elle peut aussi être définie par des associations dans le cadre de bases de données relationnelles. Une réflexion sur l'extraction des caractérisations a déjà été entamée et un prototype est en cours de réalisation, implantant un cadre général d'extraction de règles de caractérisation et s'appliquant aux données géographiques.

Haute performance. Finalement, nous nous orientons vers la définition de techniques d'extraction de connaissances hautes performances, permettant d'améliorer l'interface SGBD / algorithmes d'apprentissage. En particulier, comme le nombre de requêtes exécutées est très important, des optimisations

peuvent être envisagées : évaluation groupée de requêtes relationnelles (technique notamment intéressante dans le cadre d'algorithmes génétiques accédant à des bases de données), optimisation des requêtes ou encore utilisation de clusters de stations.

3. Références

- [1] O. Ballereau, G. Hains and A. Lallouet. BSP Constraint Programming. Chapter 7, pp.95-111, *Advances in Computation: Theory and Practice*, S. Gorlatch, and C. Lengauer ed., Nova Science Books and Journals, 2002.
- [2] A. Braud and C. Vrain. A Genetic Algorithm for Propositionalization. *Proceedings of the 11th International Conference on Inductive Logic Programming*, C. Rouveirol and M. Sebag ed., Vol.2157, LNAI, pp.27-40, Springer-Verlag, sep, 2001.
- [3] A. Braud and C. Vrain. A Study of the Scalability of a Data Mining Genetic Algorithm on a Distributed Database. *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001)*, L. Spector, E. D. Goodman, A. Wu, W.B. Langdon, H.-M. Voigt, M. Gen, S. Sen, M. Dorigo, S. Pezeshk, M. H. Garzon and E. Burke ed., pp.753, San Francisco, California, USA, Morgan Kaufmann, 7-11 July, 2001.
- [4] A. Braud and C. Vrain. A Parallel Genetic Algorithm based on the BSP Model, in *Proceedings of the Genetic and Evolutionary Computation Conference Workshop Program*, A. S. Wu ed., pp.160-162, Orlando (Florida), USA, July, 1999.
- [5] A. Colmerauer and T. B. H. Dao. Expressiveness of Full First Order Constraints in the Algebra of Finite or Infinite Trees, *Principles and Practice of Constraint Programming*, ed. R. Dechter, LNCS 1894, 2000.
- [6] T. B. H. Dao, A. Lallouet, A. Legtchenko and L. Martin, Indexical-based Solver Learning, in *International Conference on Principles and Practice of Constraint Programming*, Pascal Van Hentenryck ed., Incs, Ithaca, NY, USA, Springer-Verlag, Sept. 7 - 13, 2002.
- [7] T. B. H. Dao. Résolution de contraintes du premier ordre dans la théorie des arbres finis ou infinis, *Journées Francophones de Programmation Logique et de Programmation par Contraintes*, JFPLC 2000, éd. Touraivane, pp. 225-240, Hermes Science Publications
- [8] I. Debourges, S. Guilloré-Billot and Ch. Vrain. Learning verbal relations in text maps, in *JADT 2002 - 6èmes Journées Internationales d'Analyse Statistique de Données Textuelles*, A. Morin and P. Sébillot ed., Vol.1, pp.245-256, St Malo, March, 2002.
- [9] I. Debourges, Ch. Vrain and S. Guilloré. TEXT MAPPING: A New Approach for Exploring Large Corpora, in *Proceedings of the Fourth International Workshop on Computational Semantics (IWCS-4)*, pp.447-450, Tilburg, January 2001.
- [10] R. Debruyne and G. Ferrand and N. Jussien and W. Lesaint and S. Ouis and A. Tessier. Correctness of constraint retraction algorithms. LIFO, Rapport de recherche RR-LIFO-2002-09.
- [11] A. Ed-Dbali, P. Deransart, M. Bigonha, J. de Siqueira and R. Bigonha. HyperPro - An Integrated Documentation Environment for CLP. *Proceedings in WLPE'2001 (11th Workshop on Logic Programming Environment at CP and ICLP'2001)*, A. J. Kusalik, M. Ducassé and G. Puebla ed., Paphos - Cyprus, november, 2001.
- [12] G. Dias, S. Guilloré, J.-C. Bassano and J. G. Pereira Lopes. Extraction automatique d'unités lexicales complexes : Un enjeu fondamental pour la recherche documentaire. *Revue T.A.L. (Le traitement automatique des langues) - "Traitement automatique des langues pour la recherche d'information"*, Hermès, Vol.41, No.2, C. Jacquemin ed., 2000.
- [13] G. Dias, S. Guilloré and J. G. Pereira Lopes. Extraction automatique d'associations textuelles à partir de corpora non traités. *JADT 2000 - 5èmes Journées Internationales d'Analyse Statistique de Données Textuelles*, M. Rajman and J.-C. Chapelier ed., Vol.2, pp.213-220, Lausanne, École Polytechnique Fédérale de Lausanne, 22-24 March, 2000.
- [14] G. Ferrand and A. Lallouet. A Logic Program characterization of domain reduction approximations in finite domain CSPs.

International Conference on Logic Programming, P. J. Stuckey ed., LLCS, Springer-Verlag, 2002

[15] G. Ferrand, W. Lesaint and A. Tessier. Theoretical Foundations of Value Withdrawal Explanations for Domain Reduction. 11th International Workshop on Functional and (Constraint) Logic Programming, Grado, Italy, June, 2002.

[16] G. Ferrand, A. Lallouet and J. Arsouze. The Constraint System of the Grammatical View of (Constraint) Logic Programming. International Workshop on Functional and (Constraint) Logic Programming, Moreno Falaschi ed., Grado, Italy, University of Udine, June 20-22, 2002.

[17] A. Lallouet and L. Granvilliers. Stratégies parallèles pour la relaxation de contraintes. Workshop Résolution parallèles de Problèmes NP-complets, Daniel Singer ed., Hammamet, Tunisie, April, 2002.

[18] A. Lallouet, G. Ferrand and J. Arsouze. Chaotic derivations : an operational semantics for complete constraint solving. RCORP'01, post ICLP/CP'01 workshop on Rule-Based Constraint Reasoning and Programming, Slim Abdennadher and Thom Frühwirth ed., Paphos, Cyprus, 2001.

[19] L. Martin and F. Moal. A Language-Based Similarity Measure. Proceedings of the 12th European Conference on Machine Learning, Luc De Raedt and Peter Flach ed., Vol.2167, LNAI, pp.336-347, Berlin, Springer, September, 2001.

[20] L. Martin, F. Moal and C. Vrain. Declarative expression of biases in Genetic Programming. GECCO-99: Proceedings of the Genetic and Evolutionary Computation Conference, W. Banzhaf, J. Daida, A. E. Eiben, M. H. Garzon, V. Honavar, M. Jakiela and R. E. Smith ed., Orlando (Florida), USA, Morgan Kaufmann, 13-17 July, 1999.

[21] F. Moal, T. Turmeaux and C. Vrain, Mining Relational Databases. Proceedings of the 4th European Symposium on Principles of Data Mining and Knowledge Discovery, D.A. Zighed, J. Komorowski and J. Zytkow ed., Vol.1910, LNAI, pp.536-541, Berlin, Springer, sept 2000.

[22] A. Salleb, Z. Maazouzi and C. Vrain. Mining Maximal Frequent Itemsets by a Boolean Approach. ECAI, F. van Harmelen and IOS Press Amsterdam ed., Lyon, France, Proceedings of the 15th European

Conference on Artificial Intelligence, July 21-26, 2002.

[23] A. Salleb, Z. Maazouzi and C. Vrain. Search Space Splitting for Mining Maximal Frequent Itemsets. International Conference on Information and Knowledge Engineering IKE'02, Las Vegas, USA, CSREA Press ed., June, 2002.

[24] A. Salleb and C. Vrain. An Application of Association Rule Discovery to Geographic Information Systems. PKDD'2000, 4th European Conference on Principles and Practice of Knowledge Discovery in Databases, Lyon, France, September, 2000.

[25] T. Turmeaux and C. Vrain. Learning in Constraint Databases. Proceedings of the 2nd International Conference on Discovery Science (DS-99), S. Arikawa and K. Furukawa ed., Vol.1721, LNAI, pp.196-207, Berlin, Springer, dec 6-8, 1999

[26] T. Turmeaux and C. Vrain. Generalization and Specialization for Constrained Clauses. C. Rouveirol and M. Sebag, editors, Proceedings of the Work-in-Progress Track at the 11th International Conference on Inductive Logic Programming, pages 95-106, September 2001.

[27] V. Clavier, G. Cleuziou, L. Martin. Organisation conceptuelle de mots pour la recherche d'informations sur le Web. Conférence d'Apprentissage CAP'2002, Presses Universitaires de Grenoble, pp. 220-235, Orléans, Juin 2002.

Laboratoire d'Informatique et d'Intelligence Artificielle de l'Ecole Nationale des Arts et Industries de Strasbourg (ENSAIS).

Directeur :

Bernard Keith
Keith@liia.u-strasbg.fr
LIIA-ENSAIS
24 Bd de la Victoire
67084 Strasbourg-CEDEX

Statut

Équipe reconnue par le Ministère

pour la durée du plan quadriennal 2001-2004

Composition

L'équipe compte actuellement une dizaine de personnes

5 enseignants chercheurs

Keith Bernard, Migault Bernard, De Beuvron François, Rousselot François, Smigiel Eddie.

6 doctorants

Gagean Nicolas, Gannard Julien, Luna Francisca, Schneider Etienne, Picioraga Florentin, Schwab Hervé.

Collaborations actives

Laboratoire LRPS (ENSAIS), Institute for Process Control (IPR) de Karlsruhe, équipe READ du LORIA Nancy, IAR (Institut franco-allemand des Applications de la Recherche), CRG (centre de recherche Gabriel Lippmann) à Luxembourg, groupe de recherche national TIA, université de Iasi (Roumanie), université de Pereslavl-Zaleski (Russie), Université d'Etat d'Energétique de Ivanovo (Russie). Une convention de recherche est en cours de signature avec l'IREAM de Moscou (Institut de Radiotechnique, Électronique et Automatique de Moscou)

L'équipe étant donnée sa situation géographique est résolument tournée vers des collaborations avec les pays voisins et plus spécialement l'Allemagne. Etant hébergée dans une école d'ingénieur : l'ENSAIS, elle est résolument tournée vers les problèmes de l'industrie et à ce titre entretient des contacts permanents avec des industriels, notamment Siemens et Bosch.

Le LIIA est membre fondateur de l'IAR (Institut franco-allemand pour les Applications de la Recherche) comprenant des partenaires dans le Grand Est, (Strasbourg, Nancy, Mulhouse, Grenoble) et en Allemagne (Karlsruhe et Duisbourg). Le LIIA fait partie du groupe de travail de l'IAR intitulé ICD (Intelligent Control and Diagnostic).

Une collaboration suivie est en cours avec le laboratoire IMM8 de

l'Université de Erlangen-Nuremberg.

Une partie de l'équipe participe à l'ERTI (équipe de Recherche Technologique Informatique) avec des membres des équipes IA et réseaux du LSIT de l'Université Louis Pasteur.

Projet scientifique

L'équipe s'intéresse à l'acquisition des connaissances et à la modélisation de celles-ci selon deux axes principaux de recherche : acquisition de connaissances à partir de textes et logiques de description. Elle participe à ce titre activement aux communautés d'acquisition des connaissances et d'ingénierie linguistique: comité de programme d'IC (Ingénierie des Connaissances), de TIA (Terminologie et Intelligence Artificielle), d'EKAU (European Knowledge Acquisition Workshop) et à la communauté DL (Description Logic) par des publications régulières aux workshops DL. A ces deux axes s'est ajouté depuis peu, un axe orienté vers des recherches autour des textes numérisés.

La modélisation des problèmes ou des systèmes en Intelligence Artificielle passe par une phase de repérage des concepts essentiels d'un domaine. On part en général de textes décrivant ce domaine. Il s'agit de constituer des ressources conceptuelles informelles avant d'aboutir à une représentation formalisée du modèle conceptuel.

Le LIIA travaille ainsi sur une méthode d'acquisition de concepts et de création d'ontologie. Il travaille au développement d'outils informatiques permettant son application avec efficacité.

L'originalité du LIIA est qu'il a développé parallèlement à des outils d'acquisition à partir de textes (LIKES voir paragraphe 1.1), une logique de description (CICLOP voir §2.2). Celle-ci constitue un système de représentation des connaissances qui garantit une bonne cohérence des modèles conceptuels représentés.

L'équipe tend donc maintenant à fusionner les outils résultant de ces deux

approches pour implanter une station complète d'aide à la création d'ontologie à partir de textes.

L'axe ingénierie linguistique s'est étendu. Il ne concerne plus uniquement l'utilisation d'outils dans l'aide à la création de terminologie, mais également pour l'aide à l'indexation de documents et pour la constitution de dictionnaires électroniques.

Les outils d'ingénieries linguistiques joints à ceux conçus dans l'axe représentation des connaissances ont ouvert de nouvelles perspectives en indexation et requêtes sur des documents indexés.

L'axe logique de description s'est étendu également vers des applications des DL à des problèmes réels le plus souvent industriels: configuration et réponses à des situations de crise, aide à la maintenance.

L'axe document électronique s'est développé autour des nouvelles possibilités offertes aux lecteurs face à l'évolution des nouveaux systèmes matériel ou logiciel de manipulation des textes : affichages, mémoires, compactage et également indexation et étiquetage sémantique.

La volonté de résoudre des problèmes industriels a constitué un **axe transversal** où les travaux appliqués à des problèmes industriels utilisent des techniques de l'IA. Cet aspect sera plus précisément décrit dans la dernière partie (voir §4.)

1. Thème ingénierie linguistique

1.1 Extraction d'informations à partir de textes

Les recherches du laboratoire portent sur les problèmes théoriques liés à l'utilisation d'outils informatiques pour l'analyse des langues et pour l'extraction des informations.

Ces travaux se situent à l'intersection des disciplines informatique et linguistique. En effet, de nombreux travaux d'ingénierie linguistique reposent sur l'utilisation de textes étiquetés par des programmes qui donnent des résultats

entachés d'erreurs. On ne connaît pas précisément les capacités ni les limitations de ces outils. Il est intéressant d'étudier et de faire la typologie de tous les problèmes que peuvent rencontrer les outils informatiques, puis d'établir l'importance de ceux-ci quant aux résultats.

Un autre problème fréquent est l'usage de distributions syntaxiques pour étudier les corpus de textes. En effet, il est admis que des syntagmes partageant les mêmes distributions syntaxiques ont des similitudes sémantiques. A l'usage, on s'aperçoit que la constitution des classes à la main (ou en utilisant des outils élémentaires comme les concordanciers) est longue et fastidieuse, il s'agit donc de fournir des outils de construction automatique des distributions les plus efficaces possibles. Le problème est le même que précédemment, les outils partant généralement d'étiqueteurs ou d'analyseurs peu paramétrables.

D'autres recherches concernent le collectage de motifs de recherche liés à certains phénomènes linguistiques : exemples, pour rechercher des déverbaux (c'est à dire des verbes nominalisés) rechercher des syntagmes nominaux qui ont des compléments temporels ou locatifs.

Certaines régularités syntaxiques utiles à la localisation de certains phénomènes linguistiques s'expriment bien sous forme d'expression régulières, nous nous penchons sur le problème de l'inférence semi-automatique de ces expressions : comment généraliser à partir d'exemples et obtenir des motifs pertinents.

Les documents peuvent être analysés de manière robuste favorisant ainsi une indexation plus fine. Les capacités inférentielles des logiques de description sont utilisées ici non seulement pour représenter les concepts repérés mais pour effectuer les appariements sémantiques compositionnels.

L'outil LIKES

Le laboratoire travaille depuis quelques années à la construction d'une

station d'ingénierie linguistique dédiée à l'analyse de corpus, à la construction de terminologies et d'ontologies.

Cette station diffusée en logiciel libre à l'adresse http://www-ensais.u-strasbg.fr/liia/LIIA_Products_Installers/Install.htm a été réalisée pour servir et aux linguistes et aux chercheurs informaticiens linguistes pour analyser des textes. Elle comporte un certain nombre d'outils de bas niveaux, découpage en phrase, liste des mots par fréquence, automates récursifs pour la recherche de motifs.

Elle possède un module d'aide au repérage de termes basé sur le repérage de segments répétés : celui-ci fonctionne déjà en plusieurs langues : français, anglais, espagnol, bulgare et roumain (d'autres langues sont en cours : le portugais notamment)

Elle a été dotée récemment de nouveaux modules permettant la création de terminologies multilingues.

Elle possède également un module d'application de transducteurs pour notamment procurer une aide à l'étiquetage sémantique.

Extensions futures

Deux grandes extensions sont en cours, la première consistant à inclure dans LIKES les fonctionnalités du système CICLOP de logique de description déjà développé au laboratoire afin de gérer la partie conceptuelle des terminologies (ou des ontologies) construites.

La seconde consiste à intégrer des fonctionnalités d'analyse linguistique basée sur des ressources syntaxiques dictionnaires voir paragraphe suivant

1.2 Création de ressources

Un problème posé par un besoin particulier : l'étude de la langue luxembourgeoise a mis en évidence le manque de ressources lexicales utilisables par un ordinateur.

L'étude consiste ici à créer une typologie des classes de désinences morphologiques d'une langue pour laquelle rien n'existe : le luxembourgeois.

L'outil DICOMANAGER

On a profité de ce travail pour mettre au point un logiciel d'aide à la création et à la manipulation de dictionnaires DICOMANAGER. Le but est d'assurer le maximum de cohérence à l'entrée des données.

1.3 Applications

Dans une application opérée dans le cadre de notre collaboration avec le LRPS de l'ENSAIS, concernant la modélisation de la méthode TRIZ, une méthode de normalisation des termes et du modèle d'un domaine résolvant une situation spécifique a été abordée.

En effet, suite à un manque de formalisation, les tentatives de modélisations informatiques de la méthode se sont dans le passé soldées par des échecs, malgré la ressemblance des règles de TRIZ avec les méta-règles qu'on trouve en Intelligence Artificielle comme dans Eurisko ou dans AM (LENAT) par exemple. Nous avons commencé à travailler sur des textes en collaboration avec des experts enseignant la méthode, à relever par une étude linguistique appropriée toutes les ambiguïtés venant de définitions imprécises ou redondantes. Le but est autant d'élaborer une méthode d'aide à la détection d'ambiguïté que de formaliser TRIZ afin de fournir par exemple une interface d'aide à l'utilisation de la méthode.

2. Thème logique de description

2.1 Les Logiques de Description

Les logiques de description ou en anglais Description Logics (DL) sont une famille de langages logiques permettant de formaliser les langages à base de « frame » : réseaux sémantiques, représentation centrée objet, KL-ONE... Les logiques de description distinguent deux niveaux de représentation des connaissances : le niveau terminologique (Tbox) où sont définis les concepts du domaine représenté, et les relations entre ces concepts (rôles), et le niveau assertionnel (Abox) où sont défini-

nies les instances (individu) de ces concepts. Dans la partie terminologique, les concepts et les rôles sont définis grâce à la combinaison d'opérateurs. Les opérateurs disponibles (conjonction, disjonction, négation, existence d'un rôle, restriction sur le domaine, ou le co-domaine d'un rôle, restriction de cardinalité d'un rôle, rôle inversible ...) fixent l'expressivité du langage. La famille des logiques de description englobe l'ensemble des langages spécifiques définis en utilisant tel ou tel ensemble d'opérateurs. Les travaux en logique de description attachent une importance particulière à l'étude de la complexité des différents langages proposés (complexité en particulier du test de cohérence). Des algorithmes complets de vérification de cohérence ont été implémentés, d'abord pour des langages dont la complexité maximale est polynomiale (CLASSIC), puis pour des langages dont la complexité théorique est exponentielle, mais dont le comportement effectif sur des bases de connaissances réelles est satisfaisant (RACER, DLP, FaCT, CICLOP).

2.2 Le système CICLOP.

CICLOP est un logiciel de représentation des connaissances fondé sur les Logiques de Description (DL) et développé au sein du LIIA. Le développement de ce système a toujours été guidé par les fonctionnalités requises par des applications pratiques. Dès le départ, les logiques de description offrent deux avantages par rapport aux représentations centrées objet (RCO) pour la définition de bases de connaissance flexibles :

- Les relations taxonomiques entre un concept et ses sous-concepts (classe sous-classes en RCO) ne sont pas nécessairement définies explicitement dans la base de connaissance. Un algorithme de classification est capable de calculer toutes les relations implicites entre les concepts définis par le concepteur de la base de connaissance.
- De même, la relation entre individus et concepts (instances et classes en RCO)

n'est pas obligatoirement spécifiée a priori.

Ces caractéristiques permettent à la DL de raisonner sur des connaissances incomplètes ou semi-structurées. De plus, les DL sont fondées sur une syntaxe et une sémantique formellement bien définie, ce qui permet la définition d'algorithmes corrects et complets pour déterminer :

- si la description d'un concept est cohérente (test de consistance).
- la hiérarchie des concepts (classification)
- si un individu est instance d'un concept (test d'instance)
- l'ensemble des individus instances d'un concept donné (recherche)

Le système CICLOP inclut ces services de base, mais également des extensions qui sont apparues comme nécessaires dans des domaines comme l'indexation de documents, la configuration, ou la constitution d'ontologies à partir de bases de données textuelles (corpus de textes). En particulier, CICLOP :

- permet de définir plusieurs Tbox et Abox . Chaque Tbox peut posséder une expressivité spécifique pour répondre aux besoins du domaine à modéliser. Bien que les domaines associés à chaque Tbox soient disjoints, ils peuvent être connectés à l'aide de relations inter-domaine. Cette possibilité a par exemple été utilisée dans le domaine de la configuration pour modéliser les domaines disjoints des fonctionnalités et des composants : un/des composants sont liés à une/des fonctionnalités par la relation « implante », bien que leur domaines soient disjoints (un objet ne peut être à la fois composant et fonctionnalité).
- des Tbox spécifiques sont disponibles pour représenter des objets particuliers. En particulier, une Tbox numérique est définie pour les nombres. Il est possible d'y définir des contraintes sous forme d'un ensemble d'équations et d'inéquations linéaires. Une Tbox de chaînes de caractères permet de

définir des concepts correspondant à des types énumérés.

- une gestion optimisée des disjonctions dans le raisonneur permet de tester efficacement la cohérence lors de l'ajout incrémental de nouveaux individus et/ou de contraintes dans la ABox.

2.3 Applications

2.3.1 Aide à la construction d'ontologies

L'ontologie est modélisée en Logique de Description, choisie comme formalisme de représentation des connaissances pour ses propriétés de manipulation des données incomplètes et ses mécanismes logiques d'inférence. Le système permet de vérifier que les informations données sur les concepts (synonymie, antonymie, hyperonymie...) et les relations (domaine, co-domaine...) sont cohérentes.

Ces fonctionnalités sont complémentaires de celles offertes par le système LIKES qui permet de retrouver les concepts et relations d'un domaine à partir des textes du domaine. L'intégration des fonctionnalités de LIKES et CICLOP au sein d'un système complet d'aide à la création d'ontologies est actuellement à l'étude au sein du LIIA.

2.3.2 Aide à la décision

En utilisant les Logiques de Description (LD) pour la modélisation des situations complexes et dynamiques ainsi que des réactions appropriées, un schéma d'aide à la décision basé sur les LD a été proposé. En se basant sur la connaissance générale du domaine et sur la description spécifique d'une situation donnée, il est possible de déduire des actions concrètes adaptées à cette situation. Le domaine d'application considéré était la gestion des catastrophes naturelles, et plus particulièrement des crues. Une extension temporelle fondée sur la logique temporelle d'Allen a été développée.

2.3.3 Configuration

Intuitivement, la configuration consiste à affiner la définition incomplète d'un objet jusqu'à ce que chacun de ses constituants soit totalement défini. Le processus commence avec une instance très abstraite et générale de l'objet à configurer. Cette instance initiale est alors complétée avec de nouvelles assertions, jusqu'à ce que plus aucun choix ne soit disponible. Un système fondé sur une logique de description légèrement étendue peut vérifier à chaque étape que tous les choix sont compatibles avec les contraintes du domaine (stockées dans la Tbox). De plus, le système peut également proposer à l'utilisateur un ensemble de choix possibles compatibles avec ses précédentes décisions.

2.3.4 Aide à la maintenance

Ce projet a pour objet de modéliser les spécifications de maintenabilité sous formes de contraintes à satisfaire. On concevra dans un deuxième temps un algorithme d'analyse de la cohérence des spécifications de maintenabilité s'appuyant sur les services d'inférences procurés par la logique de description. Ce projet est effectué en collaboration avec des chercheurs du LRPS (Laboratoire de Recherche en production de Strasbourg) de l'ENSAIS .

3. Thème document numérisé

Lecture électronique et représentation de documents

La recherche porte sur les documents numériques, des collections anciennes numérisées aux documents natifs dans les formats structurés récents (HTML, XML, ...).

3.1 Personnalisation multi-support

Un premier axe de recherche est basé sur le problème de la personnalisation multi-support de la lecture (PDA, WAP, ...) où l'objectif est de répondre à des demandes personnalisées de lecture. Concernant les documents structurés, la personnalisation revient à un

réagencement de ses structures qui pose deux types de problèmes :

- un problème de ré-édition ou de reformulation : transformation à étudier par rapport à la spécificité originale du document, tout en conservant ses qualités originelles ;
- Un problème de cohérence entre ce que l'utilisateur demande, ce que le support permet et ce que le ou les usages autorisent.

La reformulation se décompose en deux étapes qui correspondent à la réponse à deux questions distinctes :

- quoi afficher, par la prise en compte des choix de l'utilisateur uniquement sur le contenu ;
- comment l'afficher en tenant compte à la fois des demandes de l'utilisateur et des contraintes de l'usage et du support.

Les contraintes ergonomiques et technologiques du support sont multiples et évolutives : leur prise en compte rationnelle et durable passe par la modélisation générique de ces contraintes. La feuille de style, la méthode de composition et l'affichage deviennent le résultat de l'application d'un document au modèle du support cible particulier.

De même, la modélisation générique des usages (y compris les aspects linguistiques, étiquetage sémantique, indexation, gestion des concepts) du document permet l'enrichissement de la feuille de style. La recherche ici concerne l'extension des outils de traitement des langues du laboratoire pour indexer sémantiquement les textes et manipuler les ontologies spécifiques à l'utilisateur et personnalisables. Les outils seront adaptés et étendus pour aider à résoudre ce nouveau problème.

3.2 Traitement graphique d'images mots

Un second axe de recherche concerne le traitement graphique d'images obtenues par numérisation de collections anciennes. Nous développons des méthodes de valorisation de ces collections basées sur leur représentation et sur la reconnaissance des formes qui le

composent. Le format hypertiff, proposé par le laboratoire, permet d'encapsuler des liens de navigation dans un fichier tiff multipage en exploitant des tags privés, c'est-à-dire en restant compatible avec le standard tiff en vigueur. Par ailleurs, le laboratoire maîtrise les principales méthodes de traitement d'images et de reconnaissance de formes pour exploiter et valoriser les documents des collections.

Le LIIA est en collaboration active avec l'équipe READ du LORIA/INRIA à Nancy, spécialiste national reconnu en reconnaissance de caractères.

4. Axe Transversal

Alors que certains travaux de l'équipe du LIIA visent à approfondir tel ou tel aspect théorique, d'autres complémentaires visent à appliquer les méthodologies disponibles à des problèmes industriels concrets. De telles applications ont aussi leur utilité sur le plan théorique : elles servent quelques fois d'aiguillon pour susciter de nouvelles investigations, elles permettent aussi de vérifier l'importance pratique de certaines limitations théoriques (par exemple sur le temps de calcul de certains algorithmes).

Projets appliqués :

- modélisation de la fiabilité de moteurs électrique (thèse financée par la société Bosch)
- applications de l'IA à des systèmes embarqués de l'industrie pétrolière (2 thèses financées par la société Schlumberger)

Transferts de technologie

Dans le cadre de l'ERTI Equipe de Recherche Technologique en Informatique de Strasbourg (regroupement de chercheurs transversal aux laboratoires) des chercheurs du LIIA coopèrent avec différents groupes du LSIIT de Strasbourg (Laboratoire des Sciences de l'Image, de l'Informatique et de Télédétection-UMR7005 CNRS-ULP)

Les chercheurs du LIIA travaillent dans le cadre du projet général de l'ERTI intitulé Réseaux Intelligents dans la partie « Gestionnaire de bases de connaissances ».

Dans le cadre de la même ERTI, le LIIA participe également au projet « Terminal intelligent pour des services boursiers » qui procure des fonctionnalités de veille en temps réel sur les informations boursières : il s'occupe plus particulièrement de la veille concernant les informations boursières textuelles.

Exemples de publications récentes

BECHINA A., BRINKSCHULTE U., PICIOROAGA F., SCHNEIDER E., Real Time middleware for industrial embedded measurement and control application, IPR, Institute for Process Control - Karlsruhe & LIIA, ENSAIS, Laboratoire d'Informatique et d'Intelligence Artificielle, Strasbourg, 2000.

TODIRASCU A., de BEUVRON F., ROUSSELOT F. - "Using Description Logics for Indexing Documents", Proceedings of the 14th IAR Annual Meeting (18-19 November 1999, Strasbourg, France), pp. 109 - 116, ISSN 0947-0727.

de BEUVRON F., ROUSSELOT F., KULLMANN M., RUDLOFF D., SCHLICK M.. "CICLOP." DL'99 workshop (STOCKHOLM, SWEDEN, 1999).

GRATHWOHL M., de BEUVRON F. et ROUSSELOT F. "A New Application for Description Logics: Disaster Management." DL'99 workshop (STOCKHOLM, SWEDEN, 1999).

TODIRASCU A, de BEUVRON F., ROUSSELOT F. - "An Ontology-Based Approach for Information-Retrieval", 10th European-Japanese Conference on Information Systems and Knowledge Modelling, Saariselka, Finland, 8-11 May 2000.

ZEGHIB Y., de BEUVRON F., KULLMANN M. "Using Description Logics for Case-Based Reasoning in Hybrid Diagnosis.", In Proceedings of the Thirteenth International Conference on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems,

PRÉSENTATION DE LABORATOIRES

June 19-22, 2000, New Orleans, Louisiana, USA.

TODIRASCU A, de BEUVRON F., GALEA D., KEITH B., ROUSSELOT F. " Using Semantics for Efficient Information Retrieval", NLDB'2000, Versailles, 25-29 June 2000.

TODIRASCU A, de BEUVRON F., ROUSSELOT F., "Integrating Information from Different Channels in Multi-Media-Contexts" Workshop, ESSLLI'00, Birmingham, 6-10 August 2000.

de BEUVRON F., KULLMANN M., RUDLOFF D., SCHLICK M., and ROUSSELOT F. "The Description Logic Reasoner CICLOP (Version 2.0)." In Proceedings of the 7th Workshop on Automated Reasoning 2000, London, Great Britain, 2000.

KULLMANN M., RUDLOFF D., de BEUVRON F. "A Disaster Management System for High Water Response Based on Description Logics." In Proceedings of the 2nd ECAI Workshop on Binding Environmental Sciences and Artificial Intelligence 2000, Berlin, Germany, 22 August 2000.

TODIRASCU A., de BEUVRON F., GALEA D. and ROUSSELOT F. "Using Description Logics for Ontology Extraction." In Proceedings of ECAI'2000 Workshop on Ontology Learning, pp.31-37, Berlin, 22 August 2000.

KULLMANN M., de BEUVRON F., ROUSSELOT F. "A Description Logic Model for Reacting in a Dynamic Environment." In Proceedings of the International Workshop on Description Logics 2000, Aachen, Germany, 2000.

TODIRASCU A., de BEUVRON F. "Using Description Logics for Ontology Extraction ." In Proceedings of ROMAND'2000 (Workshop on Robust Parsing of TALN'2000), Lausanne, 16-18 October 2000, pp. 89-105, 2000.

KULLMANN M., and KEITH B. "Description Logic Model of Disaster Reaction Processes." In Proceedings of the IFAC/IFIP/IEEE Second Conference on Management and Control of Production and Logistics 2000, Grenoble, France, 6-9 July 2000.

BECHINA A., KULLMANN M., and KEITH B. "A Description Logic Model for the Management of Automated Guided Vehicles.", In Proceedings of the IFAC Conference on Control Systems Design 2000, Bratislava, Slovak Republic, 2000.

BECHINA A., KULLMANN M., and KEITH B. "A Traffic Management System Based on Description Logics." In Proceedings of the IAR Annual Meeting 2000, Nancy, France, 16-17 November 2000, ISSN 0947-0727, pp 21-26.

FRATH P., OUESLATI R., ROUSSELOT F. "Identification de relations sémantiques par repérage et analyse de cooccurrences de signes linguistiques.", In J.Charlet, M.Zacklad, G.Kassel & D.Bourigault, eds, Ingénierie des Connaissances, évolutions récentes et nouveaux défis, Editions Eyrolles 2000, ISBN: 2-212-09110-9.

TODIRASCU A., KEITH B., ROUSSELOT F. "Handling Erroneous and Incomplete Queries." In Proceedings of IAR2000, Nancy, France, 16-17 novembre 2000, ISSN 0947-0727, pp 57-62.

TURENNE N. "Learning of the pre-structured concept set of a domain : the Galex tool." *Mathématique, Informatique & Sciences Humaines*, n°148, pp 41-71, ISSN 0995-2314, 1999.

TODIRASCU A., de BEUVRON F., KEITH B., ROUSSELOT F., GALEA D., Using Description Logics for Semantic Indexing, *Romanian Journal Of Information Science And Technology*, Volume 4, Numbers 3-4, 2001, pp 311-323..

BECHINA A., BRINKSCHULTE U., KULLMANN M., KEITH B., A knowledge representation framework for mobile robot. In Proceedings of the Robotics and Manufacturing Conference (RM 2001) Cancun, Mexico, 2001.

BECHINA A., BRINKSCHULTE U., PICIORAGA F., SCHNEIDER E., Real Time middleware for industrial embedded measurement and control. In proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications, PDPTA'2001, Las Vegas, USA, 2001.

BECHINA A., BRINKSCHULTE U., PICIORAGA F., SCHNEIDER E., A temporal description Logic approach for transportation system. In proceedings of the International Conference on Artificial Intelligence IC-AI'2001, Las Vegas, USA, 2001.

BECHINA A., BRINKSCHULTE U., KEITH B., Une nouvelle plate-forme dédiée à l'enseignement à distance, 4èmes Journées des Réseaux. JRES2001, Lyon, France, 10-14 décembre 2001, pp177-184.

Dossier « IA et Diagnostic »

Coordonnateurs :

**Marie-Odile Cordier (IRISA), Louise Travé-Massuyès (LAAS),
Michel Dojat (U438, Inserm-UJF), Sylvie Cauvin (IFP)**

Projet DREAM (Diagnostic, Recommandations d'Actions et Modélisation)

IRISA/Université Rennes I
Campus de Beaulieu
35042 Rennes Cedex
<http://www.irisa.fr/dream/site>

Personne à contacter :

Marie-Odile Cordier,
Professeur Rennes I
cordier@irisa.fr
Tel.: 02 99 84 71 00
Fax : 02 99 84 71 71

Membres de l'équipe :

Philippe Besnard, Christine Largouët,
Véronique Masson, Yves Moinard,
Yannick Pencolé, Dominique Py, René
Quiniou, Sophie Robin, Laurence Rozé

Description des travaux

Le thème principal de recherche du projet DREAM est l'aide à la surveillance et au diagnostic de systèmes ou activités complexes évoluant dans le temps. Il s'agit d'inférer l'état d'un système à partir d'observations issues de capteurs afin de détecter et de caractériser un éventuel dysfonctionnement de ce système. L'approche que nous privilégions s'appuie sur les modèles du système (modèle de bon fonctionnement ou modèle de pannes). Ces modèles, de type modèles à événements discrets, sont qualitatifs (automates, automates communicants, graphes causaux temporels, chroniques) et intègrent une dimension temporelle.

Trois points sont particulièrement étudiés : i) l'acquisition automatique de modèles, en utilisant en particulier des techniques d'apprentissage symbolique tel que la PLI (Programmation Logique Inductive); ii) le développement d'algorithmes efficaces utilisant la technique de diagnostiqueur (inversion et compi-

lation du modèle permettant de passer directement des observations aux pannes), et en particulier le diagnostiqueur décentralisé et le diagnostiqueur générique, ainsi que l'utilisation des techniques de model-checking; iii) l'interaction entre diagnostic et décision dans un univers incertain.

Les approches étudiées sont appliquées dans trois grands domaines :

i) le monitoring cardiaque (analyse en ligne d'électrocardiogrammes, étude de prothèses cardiaques *intelligentes* : pacemakers et défibrillateurs);

ii) la protection de l'environnement (amélioration de la classification de parcelles agricoles à partir d'images, modélisation qualitative du transfert de polluants tels que pesticides et nitrates);

iii) le diagnostic industriel (en particulier la supervision de réseaux de télécommunications et de distribution d'électricité.

Deux projets représentatifs

Projets MAGDA <http://magda.elibel.tm.fr/> et MAGDA2 http://www.telecom.gouv.fr/rnrt/index_net.htm

Les projets MAGDA et MAGDA2 (projets RNRT / 98-03) portent sur la corrélation d'alarmes, la surveillance et le diagnostic de réseaux de télécommunications. Les partenaires en sont France-Telecom R&D, Alcatel, Ilog, le LIPN et l'IRISA. Le projet MAGDA (Modélisation et Apprentissage pour une Gestion Distribuée des Alarmes) s'est terminé en décembre 2001. Dans ce cadre, nous avons en particulier proposé et développé des algorithmes décentralisés pour le diagnostic en ligne de systèmes à événements discrets. Ils ont été expérimentés sur les réseaux de type SDH. Le projet MAGDA2 (Modélisation et l'Apprentissage pour une Gestion Distribuée des Alarmes De bout En boUt) a débuté en décembre 2001. Il poursuit le projet MAGDA et rassemble les mêmes partenaires. Il a

pour objectif de trouver des solutions avancées aux problématiques de la gestion des réseaux hétérogènes et de la prise en compte d'événements faisant intervenir l'interaction entre la couche réseau et la couche service. Nous sommes plus particulièrement concernés par i) l'étude et le développement d'algorithmes de diagnostic robustes et distribués, capables de supporter l'incomplétude des modèles, la perte d'alarmes et la reconfiguration en ligne du réseau ; ii) l'apprentissage de chroniques caractéristiques des pannes par analyse des fichiers de logs d'alarmes.

Projets PISE (ACI « Télé médecine et Santé ») et CEPICA (RNRTS)

Ces projets se font en collaboration avec le LTSI (laboratoire de traitement du signal), le CHU de Rennes/département de cardiologie et la société ELA Medical, spécialiste des prothèses cardiaques implantables tels que les pacemakers et défibrillateurs. L'objectif de ces projets est la détection en ligne d'arythmies cardiaques par analyse des signaux issus des capteurs tels que les électrocardiogrammes. Il vise le développement de prothèses implantables « intelligentes », c'est à dire capables d'analyser en ligne les signaux et de réagir de manière appropriée. Nous étudions en particulier l'approche reconnaissance de chroniques et l'apprentissage automatique d'une base de chroniques dédiées aux arythmies.

Position par rapport à d'autres travaux, nationaux ou internationaux

• Les travaux de l'équipe autour des approches décentralisées pour le diagnostic de systèmes à événements discrets sont originaux et se font en bonne concertation avec les équipes internationales travaillant sur ce sujet. C'est le cas aussi pour nos travaux autour de l'utilisation des techniques de model-checking qui sont originales dans le domaine du diagnostic et se font en

bonne concertation avec les travaux similaires actuels en planification.

- Nous sommes très impliqués dans les travaux du groupe Imalaïa visant à rapprocher les travaux en diagnostic issus des domaines de l'automatique et de l'intelligence artificielle ainsi que dans le réseau d'excellence MONET2 sur le raisonnement à base de modèles et les approches qualitatives du diagnostic.
- Dans le domaine médical, notre point fort est la bonne collaboration que nous avons avec les chercheurs en traitement du signal du LTSI ainsi qu'avec le département cardiologie du CHU de Rennes. Ceci nous permet d'envisager un système complet et des expérimentations réalistes.
- Nos travaux dans le domaine de la protection de l'environnement se font en collaboration et avec le support financier de l'INRA.

Activités d'animation dans le domaine

En France :

- Animation du groupe Imalaïa (voir groupe de travail AFIA). Participation active au GDR I3
- Au niveau international
- Participation active au groupe de travail Bridge (groupe de travail du réseau d'excellence MONET2)
- Responsabilités à divers comités éditoriaux ainsi que dans les comités de programme des conférences spécialisées telles que DX, IDAMAP, BESAI ou générales telles que RFIA, ECAI, IJCAI.

Courte bibliographie

Diagnostic de réseaux de télécommunications :

L. Rozé, M.-O. Cordier *Diagnosing discrete-event systems : extending the "diagnoser approach" to deal with telecommunication networks*, Journal on Discrete-Event Dynamic Systems : Theory and Applications (JDEDS), Kluwer Academic Publishers, vol. 12, No. 1, 43-81, 2002

Y. Pencolé, M.-O. Cordier, L. Rozé, *A decentralized model-based diagnostic tool for complex systems*, International Journal on Artificial Intelligence Tools, World Scientific Publishing, vol. 11, No. 3, 327-346, 2002

Monitoring cardiaque et apprentissage de chroniques :

G. Carrault, M.-O. Cordier, R. Quiniou, M. Garreau, J.-J. Bellanger, A. Bardou *A model-based approach for learning to identify cardiac arrhythmias*, AIMDM'99 : Artificial Intelligence in Medicine and Medical Decision Making, Springer Verlag, Aalborg, Denmark, vol. 1620, 165-174, 1999

Model-checking et diagnostic :

M.-O. Cordier, C. Largouët *Using model-checking techniques for diagnosing discrete-event systems*, Proceedings of the Twelve International Workshop on Principles of diagnosis (DX'01), 39-46, 2001

Graphes causaux temporels et interaction causale :

I. Grosclaude, M.-O. Cordier, R. Quiniou *Causal interaction: from a high-level representation to an operational event based representation*, IJCAI'01, 565-572, 2001

Surveillance de parcelles agricoles :

C. Largouët, M.-O. Cordier *Improving the Landcover Classification using Domain Knowledge*, AI Coms, special issue on Environmental Sciences and Artificial Intelligence, IOS Press, vol. 14, No. 1, 35-43, 2001

C. Largouët, M.-O. Cordier *Timed Automata Model to Improve the Classification of a Sequence of Images*, ECAI'2000 (European Conference on Artificial Intelligence), Berlin, Allemagne, 156-160, 2000

R. Quiniou, M.-O. Cordier, G. Carrault, F. Wang *Application of ILP to cardiac arrhythmia characterization for chronicle recognition*, ILP'2001, Springer-Verlag, vol. 2157, 220-227, 2001

M.-O. Cordier, F. Lévy, J. Montmain, L. Travé-massuyès, M. Dumas, M. Staroswiecki, P. Dague, *AI and Automatic Control Theory approaches of model-based diagnosis : links and underlying hypotheses*, Proceedings of Safeprocess'2000, 274--279, 2000

Adresse d'un site web où l'on peut trouver l'ensemble des références et articles : <http://www.irisa.fr/dream/site>

Groupe Connaissances et Complexité. LGI2P – École de Mines d'Alès (EMA)

Site EERIE
30000 Nîmes Cedex1
<http://www.lgi2p.ema.fr/>

Personne à contacter :

Jacky Montmain, Ingénieur CEA à

l'Unité Mixte de Recherche entre le CEA et l'EMA

jacky.montmain@site-eerie.ema.fr

Tel.: 04 66 38 70 58

Fax : 04 66 38 70 59

Description des travaux

Le thème général du groupe est d'élaborer des modèles pour **aider l'opérateur humain à décider et agir en situation complexe** faisant appel à ses capacités cognitives. Afin qu'un opérateur humain puisse agir de façon pertinente, il est nécessaire aujourd'hui d'élaborer des systèmes d'information qui soient de véritables aides au raisonnement et à la compréhension de situations complexes.

Nos travaux dans le domaine des systèmes d'information évolués pour l'aide à la décision concernent deux grands domaines d'application. D'abord, cette évolution a été perçue en contexte industriel, où le rôle de l'opérateur d'exploitation a évolué de la conduite à la **supervision** avec des tâches à plus haut niveau décisionnel, l'accent est mis sur l'aide informatisée au raisonnement et à la prise de décisions adéquates lors de dysfonctionnements (depuis 1987). L'aide à la décision peut également prendre une dimension collective : dans le cadre d'un grand projet, la collectivité des acteurs doit partager un modèle commun de la situation, un référentiel des connaissances, qui est utilisé comme support pour élaborer une aide à l'argumentation et à l'évaluation collective qui est vue comme une activité de **supervision non plus de systèmes techniques mais de systèmes organisationnels ou stratégiques** (depuis 1999).

Supervision de systèmes techniques

En 1987, le CEA de Marcoule lance un premier programme de R&D, DIAPASON, dans le domaine de la supervision des procédés continus. Le retraitement nucléaire, par la taille, le degré d'automatisation et la complexité de ses installations, est typiquement un centre d'intérêt pour des applications de supervision. Si DIAPASON a pour objectif de

promouvoir l'intérêt du **diagnostic à base de modèles**, le second programme SUP-XX, lancé en 1993 pour une usine de retraitement français doit, lui, montrer la faisabilité de fonctionnalités avancées d'aide à l'opérateur en contexte industriel.

Pour expliquer ou aider à mener un raisonnement, il faut que les exploitants et le système informatique partagent une même représentation de la situation de conduite. A cet effet, l'équipe s'oriente vers le **Raisonnement Qualitatif** : les modèles qualitatifs sont des outils pour comprendre et expliquer le comportement de systèmes physiques. De ces études naît un simulateur causal événementiel qui privilégie le critère de pertinence à celui de précision de la simulation. L'utilisation de cette simulation comme image du comportement de référence du procédé est à la base de notre rapprochement avec la communauté du diagnostic. Une méthode utilisant la connaissance des **dépendances causales et dynamiques (Raisonnement Causal)** entre les variables du graphe est proposée pour rechercher la déviation première capable d'expliquer l'ensemble des autres alarmes constatées. La puissance de localisation des méthodes numériques est acquise sans avoir recours à des manipulations algébriques pour obtenir un jeu de résidus structurés. Enfin au raisonnement causal pour la supervision, s'adjoint le raisonnement approché : le principe de base en est que la résolution pratique en termes d'ordres de grandeur ou qualitative de problèmes complexes correspond à l'idée que raisonner symboliquement sur des données numériques est une façon de procéder courante chez l'être humain. Ce dernier point explique l'intérêt porté au **Raisonnement Flou** dans l'approche choisie et à son utilisation pour l'aide à la décision. Enfin, nous avons montré la nécessité d'avoir un **Raisonnement Multi-points** de vue pour un diagnostic pertinent dans le temps. En effet, le véritable objet de la supervision, c'est le suivi et la compréhension de la situation de conduite et pas seulement du processus physique mis en jeu : le triplet {opérateur, installation, conduite} détermine une situation complexe qui doit être déclinée selon plusieurs points de vue (change-

ments de mode de fonctionnement, de configuration, d'activité, de cadence, de configuration ou de stratégie) qui ne relèvent pas seulement de la connaissance des principes premiers utilisés dans les modèles comportementaux.

Supervision de systèmes organisationnels et stratégiques

En 1999, l'orientation de nos recherches concerne l'extension de la problématique de la supervision à des systèmes non plus techniques mais organisationnels ou stratégiques, les grands projets en particulier. Dans le cadre de projets de longue durée faisant intervenir un ensemble important et dynamique d'acteurs de cultures variées, notre équipe développe des outils pour le suivi de la production scientifique de ces projets autour de serveurs intranet : les référentiels de connaissances (depuis 1997). La fonction technique d'un référentiel est de gérer l'évolution du corpus de connaissances afin de faciliter la constitution de dossiers de synthèse et de fournir un argumentaire à l'évaluation des solutions et décisions du projet. Sur la base du référentiel des connaissances est alors développée une activité de supervision de projet ou de management de projet par pilotage des risques associés aux décisions accompagnant le projet. L'analogie avec le **contrôle ou la supervision** au sens de l'automatique est proposée : *l'objet de la surveillance* est la base de connaissances gérée par le référentiel et la *variable régulée* est l'entropie de la base de connaissances ou le risque décisionnel associé. Le couplage d'un système de gestion dynamique des connaissances avec un système informatisé d'aide à la décision (SIAD) a naturellement conduit à doter le SIAD de **fonctionnalités d'argumentation ou d'élucidation**, bouclant ainsi sur le thème principal de notre groupe de recherche.

Position par rapport à d'autres travaux, nationaux ou internationaux

- Les travaux de l'équipe pour le diagnostic ont souvent été développés en

collaboration avec le LAG (aspect approches causales et floues) mais aussi avec le LAIL de Lille (aspect FDI) et le LAMIH de Valenciennes (aspect Intégration Homme-Système). Nous avons toujours eu le souci de comparer nos travaux à ceux de la communauté internationale : de Iri en 1979 sur les Signed Directed Graphs (SDG) pour le diagnostic jusqu'au Fault Monitoring avec un SDG de Biswas et Mostermann en 1998.

- Nous sommes impliqués dans les travaux du groupe IMALIA visant à rapprocher les travaux en diagnostic issus des domaines de l'automatique et de l'intelligence artificielle.
- Nos travaux dans le domaine du contrôle du risque pour le management de projet se font en grande partie en collaboration avec le LAMII (LIS-TIC) d'Annecy.

Activités d'animation dans le domaine

En France : Participation au groupe IMALIA (voir groupe de travail AFIA).

Au niveau international : Responsabilités à divers comités de programme des conférences spécialisées.

Courte bibliographie

Supervision de systèmes techniques :

- Evsukoff, A., Gentil, S., Montmain, J. (2000) Fuzzy Reasoning in co-operative supervision systems, *Control Eng. Practice*, 8 (2000), 389-407.
- Montmain, J., Gentil, S. (2000) Dynamical causal model diagnostic reasoning for online technical process supervision, *Automatica*, 36 (2000), 1137-1152.
- Montmain, J. (2000) Supervision applied to nuclear fuel reprocessing, *AI Communications The European Journal of AI*, 13 (2), 61-81
- Cordier, M-O., Dague, P., Lévy, F., Dumas, M., Montmain, J., Staroswiecki, M., Travé-Massuyès, L. (2000). A Comparative Analysis of AI and Control Theory Approaches to Model-based Diagnosis. *ECAI'2000*, Berlin, Germany.
- Vaudry, C., Crampes, M. (2001) AMEBICA : an adaptive user interface for process control. *IHM-HCI'01, Conférence Interaction Homme-Machine*, Lille, France
- S. Cauvin, M-O. Cordier, C. Dousson, P. Laborie, F. Lévy, J. Montmain, M. Porcheron, I. Servet, L. Vaudry, C., Crampes, M. (2002) *Composition adaptati-*

ve dans la supervision industrielle, Revue IHM

S. Cauvin, M-O. Cordier, C. Dousson, P. Laborie, F. Lévy, J. Montmain, M. Porcheron, I. Servet, L. Travé-Massuyes. (1998) Monitoring and alarm interpretation in industrial environments, *AI Communications The European Journal of AI*, IOS Press, 11 (3/4), 139-173.

Supervision de systèmes organisationnels ou stratégiques :

Akharraz, A., Montmain, J., Mauris, G. (2002). Decision acceptability in project management : argumentation and risk control. IAR ICD Workshop Intelligent Control and Diagnosis, Grenoble, France

Akharraz, A., Montmain, J., Mauris, G. (2002). Fonctionnalités explicatives d'un système de fusion de connaissances collectives par intégrale de Choquet 2-additive. *LFA'2002, Rencontres Francophones sur la Logique Floue et ses Applications*, Montpellier, France.

Akharraz, A., Montmain, J., Mauris, G. (2002). A project decision support system based on an elucidative fusion system. *Fusion 2002, 5th International Conference on Information Fusion*, Annapolis, Maryland, USA.

Penalva, J-M. and Montmain, J. (2002). Travail collectif et intelligence collective : les référentiels de connaissances. *IPMU'2002, 9th International Conference on Information processing and Management of uncertainty in Knowledge-Based Systems*, Annecy, France.

Montmain, J., Akharraz, A. and Mauris, G. (2002). Knowledge management as a support for collective decision-making and argumentation processes. *IPMU'2002, 9th International Conference on Information processing and Management of uncertainty in Knowledge-Based Systems*, Annecy, France.

Projets les plus représentatifs

Projet DIAPASON (1987-1994)

Il s'agit d'un projet CEA dont l'objectif est de démontrer l'intérêt du raisonnement à base de modèle dans le cadre de l'aide à la conduite.

Les procédés nucléaires présentent des caractéristiques qui les rendent tout à fait susceptibles de justifier une assistance à la supervision :

- Les milieux traités sont extrêmement radioactifs, il est donc difficile d'équiper ces procédés avec des capteurs à la fois sophistiqués et durables. L'équipement de mesure est alors élaboré à partir de techniques rudimentaires, de façon à nécessiter le moins

d'entretien possible. Les capteurs de mesures chimiques sont particulièrement délicats à mettre au point ; ainsi, de nombreuses informations sur les phénomènes chimiques sont obtenues hors ligne après prises d'échantillons.

- Les phénomènes physico-chimiques sont relativement complexes, et la conduite des procédés nécessite une grande expérience des opérateurs, et de nombreux essais effectués sur des prototypes pour cerner l'influence des différents paramètres sur le comportement.

- La conduite est de plus très délicate, puisque, afin de permettre le recyclage en continu des produits utilisés, les différentes unités d'une installation sont interconnectées, et les différents appareils d'une même unité fortement couplés.

DIAPASON offre à l'opérateur une représentation synthétique du procédé grâce à une modélisation inspirée de la physique qualitative et de l'automatique ; il permet de prédire l'évolution du procédé, pour une conduite par anticipation. En situations incidentelles, il permet la détection et la localisation de dysfonctionnements, le diagnostic de défaillances pour une maintenance préventive ou une conduite en fonctionnement dégradé.

Les trois raisonnements mis en œuvre dans DIAPASON sont :

- le raisonnement causal ;
- le raisonnement flou ;
- le raisonnement temporel.

Projet SUP-XX (1994-1999)

Le projet SUP-XX consiste à concevoir et réaliser un système de supervision pour une usine de traitement française. L'objectif de ce prototype est de démontrer la faisabilité de la mise en œuvre, en contexte industriel, de fonctionnalités avancées d'aide à l'opérateur. Les fonctionnalités retenues dans le système d'aide à l'exploitation se divisent en deux classes : celles dédiées à la surveillance et celles dédiées à la synthèse.

Fonctionnalités de Surveillance

La surveillance est un ensemble de fonctionnalités dont l'objectif commun est d'automatiser la vérification du bon fonctionnement de l'installation, de détecter et d'expliquer au plus tôt les

éventuels dysfonctionnements qui peuvent se manifester.

Détection

La détection permet de déceler la perte ou la dégradation d'une fonction élémentaire. C'est dans ce sens qu'il a été choisi d'élaborer une détection que l'on dira fonctionnelle, parce qu'elle permet de signaler la dégradation d'une fonction sans chercher à l'expliquer par un problème dans la structure du système.

Diagnostic

L'opération de diagnostic suit l'étape de détection et attribue le défaut à des modules du type capteurs, organes de commande, processus, unité de commande. La détection permet seulement de détecter la perte ou la dégradation d'une fonction élémentaire du procédé alors que le diagnostic donne l'explication et la cause de la perte de cette fonction.

Le diagnostic permet de déduire à partir de plusieurs défauts détectés lequel est le défaut origine, c'est-à-dire celui capable d'expliquer à lui seul, l'ensemble des autres défauts constatés et enfin, d'en fournir la cause physique.

Fonctionnalités de synthèse

La synthèse regroupe plusieurs fonctionnalités dont l'objectif commun est de fournir une représentation pertinente et synthétique du procédé.

Gestion des modes de fonctionnement

La reconnaissance du mode de fonctionnement de l'installation à surveiller rentre dans le cadre des fonctionnalités de synthèse : elle a pour objet de rappeler, en "temps réel", quelle est la phase d'exploitation en cours, en se basant sur des acquisitions sur le procédé uniquement.

La gestion des modes est en premier lieu une fonctionnalité de synthèse, mais elle est également essentielle pour le contrôle du fonctionnement d'un système d'aide à l'exploitation : ce contrôle doit être mis en œuvre pour gérer l'activation des fonctionnalités de surveillance du système aux contextes de production et de fonctionnement de l'installation.

Gestion des configurations de fonctionnement

Comme les appareils de l'installation effectivement utilisés peuvent varier au cours de l'exploitation, la reconnaissance de la configuration en service est une fonction indispensable du système de supervision : elle permet de sélectionner la connaissance à utiliser par les modules de surveillance et de l'interface homme-machine, mais également de commander la mise en veille de ces modules.

Les configurations à détecter sont relatives aux appareils utilisés dans le cas de transferts doublés, d'ajustages, ainsi qu'à la mise en service de transferts optionnels.

Restitution de données

Cette fonctionnalité permet de consulter une synthèse des événements pertinents s'étant produits lors du dernier poste ou depuis une date donnée. La technique de segmentation utilisée permet la transformation du signal échantillonné d'acquisition sur le procédé - représentation informatique -, en une suite de variations significatives, représentation plus explicite pour l'observateur.

Interface Homme/Machine

L'interface opérateur a pour objectif de présenter les résultats fournis par le système d'aide à l'exploitation pour chacune des fonctionnalités, mais également de rendre le plus explicite possible les raisonnements utilisés par celui-ci. Le caractère explicatif de cette interface est une condition nécessaire si l'on veut que l'opérateur accorde sa confiance aux conclusions du système. Il paraît utile de distinguer les vues de suivi de fonctionnement et les vues dédiées à une fonctionnalité donnée.

Projet AMEBICA (2001)

AMEBICA signifie Auto-adaptive Multimedia Environment Based on Intelligent Collaborating Agents. AMEBICA est un projet ESPRIT avec pour partenaires : Alcatel, IFE, Loughborough, Softeco, Iberdrola, LABEIN, ELSAG. L'objectif du projet est la modélisation et l'expérimentation d'un framework 'agents' pour la réalisation d'interfaces multimédia industrielles adaptatives. L'objectif de l'interface est d'assister l'opérateur en lui

présentant des informations pertinentes, sous une forme adéquate, au moment approprié.

Projet Gestion des Risques (2000)

L'objectif du projet « Gestion des Risques » était de fournir un prototype de système d'aide au diagnostic pour limiter les impayés de facture pour un opérateur de télécommunications. Il s'agissait à partir des caractéristiques et des consommations d'un client dans les deux premiers mois suivant son activation, de prévoir si celui-ci présentait des risques élevés de ne pas payer ses factures par la suite. Le système de classification fournissait les degrés d'appartenance d'un usager à chacune des classes {Très mauvais client, mauvais client, moyen client, bon client, très bon client} après analyse des variables explicatives décrivant le client et son comportement lors des deux premiers mois suivant son activation.

Deux techniques ont été comparées :

- l'identification des modèles de classification basée sur le principe de « fuzzy clustering » ou regroupement flou ;
- les réseaux de neurones proposent une autre alternative pour la synthèse expérimentale de classificateurs par les algorithmes courants d'identification des poids du réseau. Dans le cas de la synthèse expérimentale d'un classificateur « l'inférence floue » précédemment proposée est remplacée par un réseau de neurones. L'aspect dynamique de l'analyse de la consommation des clients a été un des éléments qui nous a poussés à tester des réseaux récurrents.

Société :

THALES Airborne Systems

THALES Airborne Systems
1 Bd Jean Moulin
78852 - ELANCOURT
+33134596000

Responsable Équipe : Patrick Taillibert
patrick.taillibert@fr.thalesgroup.com
+33134597993

Les premiers travaux de THALES

Airborne Systems – à l'époque Dassault électronique- visant à l'automatisation du diagnostic des circuits électroniques analogiques remontent à 1982! Le diagnostic n'était pas chose nouvelle pour l'entreprise dans la mesure où des programmes automatiques étaient déjà disponibles pour des circuits digitaux de plus en plus complexes mais, pour les modules « hybrides analogiques », le diagnostic était fait entièrement « à la main » - ou, plus exactement grâce à l'expérience des dépanneurs. 1982 vit également l'émergence des systèmes experts, tout au moins en France. La première idée venant à l'esprit d'une équipe commune avec IBM constituée pour résoudre le problème de diagnostic des circuits analogiques, fut naturellement de s'appuyer sur cette technologie « pleine de promesses ». Ils durent rapidement déchanter: les systèmes experts n'étaient décidément pas la bonne solution au problème. Les raisons de cette inadéquation sont maintenant bien connues (petites séries, très grand nombre de fautes possibles, présence quasi systématique de boucles de rétroaction...). La résolution du problème posé ne devait donc pas être si facile, comme l'historique qui va suivre permet de s'en rendre compte.

La première étape fut DEDALE [Dague et al., 1987], déjà une approche à base de modèles, s'appuyant sur la notion qualitative d'ordre de grandeur développée par O. Raiman [Raiman, 1986]. Bien qu'encourageants, les résultats n'ont pas néanmoins permis d'envisager une utilisation du système dans un contexte industriel. La limitation principale de DEDALE était que seuls les aspects statiques des circuits étaient pris en considération alors que - et cette propriété est bien comprise maintenant- de nombreux défauts ne peuvent être détectés et encore moins localisés sans raisonner sur le comportement dynamique du circuit.

Il fallait donc poursuivre les recherches. C'est pourquoi, toujours en collaboration avec le centre scientifique IBM de Paris, une nouvelle étude fut lancée en 1988 avec pour objectifs de permettre la prise en compte des aspects dynamiques des systèmes à diagnosti-

quer tout en considérant des approches plus quantitatives que les précédentes afin d'améliorer la finesse de diagnostic. Il en résulta le programme CATS [Dague et al., 1991] utilisant des algorithmes de propagation d'intervalles, où les données -c'est à dire les signaux électriques- y sont représentés par des tableaux d'intervalles. Cette fois fut la bonne et en 1992, une évaluation générale du système CATS menée sous contrat de la Délégation Générale à l'Armement (DGA/SEFT) prouva le bien fondé de l'approche retenue. Bien entendu, le système CATS est complètement indépendant du domaine d'application. Ainsi il fut utilisé avec succès pour le diagnostic du système d'alimentation électrique d'un satellite [Devès et al., 1991] -ce qui n'est bien sûr pas très éloigné de l'électronique- mais également pour le diagnostic des capteurs d'une usine pétrochimique espagnole [Alarcon et al., 1994] application pour laquelle les grandeurs manipulées étaient des flux, pressions, températures et non plus seulement des tensions et des courants électriques. Finalement, l'industrialisation complète du système CATS a démarré en 1997. Le projet -dénommé CDIAG/ DIAMANT- a permis la construction d'une bibliothèque d'objets et d'interfaces de programmation (API) destinée à intégrer les algorithmes de diagnostic dans n'importe quel système et en particulier dans des équipements de test automatiques développés pour les circuits électroniques (ATE). Le sous-ensemble du projet appelé DIAMANT avait pour but, pour sa part, de démontrer la capacité de cette bibliothèque d'objets à s'adapter à différents environnements en construisant un système complet de diagnostic automatique sur les bancs de test SESAR 400 de l'Aérospatiale.

L'aventure aurait pu s'arrêter là, tout au moins en ce qui concerne la recherche, mais les nombreuses expérimentations conduites avec le système CATS ont mis en évidence que de nouvelles améliorations étaient envisageables, en particulier en ce qui concerne les facilités offertes pour la modélisation des composants des systèmes à diagnostiquer. Dans CATS, la représentation des données n'est pas très

économique, les modèles de composants sont limités à des équations différentielles du premier ordre et, ce qui n'est pas le moindre des inconvénients, une évaluation numérique de la dérivée des observations est nécessaire. Ce dernier point n'est pas sans poser problème et cela d'autant plus que n'importe quel signal peut apparaître en raison justement de la ou des pannes présentes dans le système. Un nouveau projet fut donc lancé en 1993, cette fois en collaboration avec l'Institut Supérieur d'Electronique du Nord (ISEN) [Loiez and Taillibert, 1997; Taillibert 1998]. Les grandeurs sont modélisées par des « Suites de Bandes Temporelles » - approximations polynomiales par morceaux des signaux- et les composants par des équations de la forme $f+g'+h''+\dots=0$, où f, g, h,... sont des combinaisons des grandeurs mesurables du système à diagnostiquer.

De 1996 à 1999, l'entreprise a participé au projet européen « Vehicule Model-Based Diagnosis » [VMBD, 1996] destiné à étudier l'applicabilité des techniques de diagnostic à base de modèles dans le contexte de l'industrie automobile (diagnostic embarqué et à l'atelier).

De puis 2001, nous sommes partie prenante dans le projet CHEM [CHEM, 2001] (Advanced decision support system for Chemical/Petrochemical manufacturing processes), afin d'y intégrer nos algorithmes d'évaluation de résidus construits à partir d'une représentation des mesures sous forme de suites de bandes temporelles et évitant tout recours au calcul numérique de la dérivée des mesures. La recherche menée se focalise sur les techniques de développement des modèles des composants mis en jeu dans ce type d'installation.

Références

- [Alarcon et al., 1994] I. Alarcon et al., « Heterogeneous Integration Architecture for Intelligent Control Systems », *Int. Journal of AI in Engineering*, 1994.
- [Dague et al., 1987] P. Dague, P. Devès, and O. Raiman, « Troubleshooting: when Modeling is the Trouble. », *Proceedings of AAAI Conference, Seattle, July 1987*.

- [Dague et al., 1991] P. Dague, P. Devès, P. Luciani, O. Jehl and P. Taillibert, « When oscillators stop oscillating. », *Proceedings of the twelfth IJCAI, Sydney, August 1991*.
- [Devès et al., 1991] P. Devès, C. Fischer, P. Taillibert, « An Alternative to Expert Systems for Electrical Diagnosis », in *Future Generation Computer Systems*, Vol 7 91/92.
- [Loiez and Taillibert, 1997] E. Loiez, P. Taillibert, « Polynomial Temporal Band Sequences for Analog Diagnosis », *Int. Joint Conf. on Artificial Intelligence IJCAI'97, August 23th-29th, 1997, Nagoya, Japan*.
- [Taillibert, 1998] P. Taillibert. « Can Chemical Process Industry benefit from Analog Electronics Diagnosis » *3rd IFAC Workshop on on-line Fault Detection and Supervision in the Chemical Process Industry. LYON SOLAIZE (France) - 4-5 Juin 1998. - à paraître -*
- [Taillibert, 1998] P. Taillibert. « Various Improvements to Diagnosing with Temporal Bands » *9th Int. Workshop on Principles of Diagnosis (DX 98) CAPE COD (USA) - 24-27 Mai 1998*.
- [Taillibert, 1998] P. Taillibert « Residual Evaluation without Derivative Computation » *Symposium on fault detection, Supervision and safety for Technical Processes SAFEPROCESS'2000 BUDAPEST (Hongrie) - 14-16 Juin 2000*
- [Raiman, 1986] O. Raiman, « Order of Magnitude Reasoning » *Proceedings of AAAI Conference, Philadelphia, July 1986*.
- [VMBD, 1996] <http://www.aber.ac.uk/~nst/vmbd/>
- [CHEM, 2001] <http://www.chem-dss.org/chem-dss>

Équipe ILOG Business Rules R&D

9, rue de Verdun
94253 Gentilly
+33 1 49 08 35 00
Responsable : Nitsan Sèniak
Mél: nсениак@ilog.fr
http://www-leibniz.imag.fr/AFIA/ECAI/Fiches/ILOG-ILOGoJRules-Nitsan_Seniak.html

Thème général de l'équipe

L'équipe conçoit et développe le produit ILOG JRules, qui regroupe d'une part un moteur d'inférence fondé sur l'algorithme RETE, d'autre part un environnement de développement et de mise au point de programmes à base de règles.

Membres de l'équipe concernés par le thème du diagnostic

Changhai Ke (ke@ilog.fr)

Bruno Berstel (berstel@ilog.fr)

Travaux en lien avec le diagnostic

Les règles traitées par ILOG JRules peuvent porter sur des événements datés, et inclure des contraintes temporelles : leur traitement par ILOG JRules permet entre autre la reconnaissance de chroniques. Le formalisme des chroniques peut être utilisé dans le diagnostic pour représenter des séquences d'événements caractéristiques d'un symptôme.

Projets

Le développement de la fonction de reconnaissance de chroniques intégrée au produit ILOG JRules a été effectué dans le cadre du projet RNRT Magda. Le responsable de ce projet était France Telecom R&D, représenté par Christophe Dousson.

Le moteur d'inférence d'ILOG JRules étant fondé sur l'algorithme Rete, l'intégration a consisté à étendre cet algorithme pour lui permettre de manipuler les événements datés et les contraintes temporelles, et de reconnaître les occurrences de chroniques dans un flux de données.

Le résultat du projet a été le transfert de la technologie de reconnaissance de chroniques dans le produit industriel ILOG JRules.

Position par rapport à d'autres travaux

Le formalisme des chroniques a été introduit par Dousson en 1994. La reconnaissance de chronique a été implementée par lui dans l'outil CRS (Chronicle Recognition System). L'intégration de la reconnaissance de chroniques à ILOG JRules est évidemment fondée sur sa description du formalisme des chroniques, mais elle n'est pas fondée sur les mêmes techniques

que CRS puisqu'elle a consisté en l'extension de l'algorithme Rete. L'algorithme Rete a été décrit par Forgy en 1982.

Activité d'animation dans le domaine

ILOG prévoit de co-organiser la seconde édition des Journées sur les Chroniques. La première édition a eu lieu en avril 2002 ; son site web est à <http://crs.elibel.tm.fr/seminaire/>

Bibliographie

Les publications de l'équipe sur le sujet sont :

- Bruno Berstel. Intégration de la reconnaissance de chroniques à un algorithme incrémental de pattern matching. RFIA 2002.

- Bruno Berstel. Extending the RETE Algorithm for Event Management. TIME 2002.

On trouve de nombreuses références sur les chroniques sur :

<http://crs.elibel.tm.fr/fr/biblio.mhtml>

ARDANS - Consulting & Knowledge Management Équipe Mémoire en continu pour le diagnostic

ARDANS - Consulting & Knowledge Management

2, rue Hélène Boucher

78286 GUYANCOURT CEDEX

www.ardans.fr

tél. : 33 (0)1 39 30 99 00

Directeur : Pierre Mariot

pmariot@ardans.fr

Choix d'un système d'aide au diagnostic pour un industriel : initier une mémoire en continu

Ardans est une jeune entreprise de conseil & ingénierie spécialisée en système d'information & gestion des connaissances ("Mémoire en continu") qui regroupe aujourd'hui 15 ingénieurs.

Ardans a été créée par des Consultants issus de Decan (TDL & Ingénia et antérieurement Framentec-Cognitech).

Ardans intervient actuellement principalement auprès de grands comptes (PSA, EdF, FT, VAG) et appuie aussi des PME (Soprel).

Ses actions liées à la problématique « Diagnostic » sont centrées sur la valorisation

des connaissances des métiers pour rendre les entités clientes plus performantes :

- i) en détectant les connaissances critiques, formalisant les savoir-faire et les bonnes pratiques
- ii) en aidant à la diffusion de ces connaissances dans l'entité via les outils informatiques,
- iii) en préparant l'organisation destinée à faire vivre ce capital pour mettre en œuvre la "mémoire en continu" de l'entité (équipe, service, département, direction, société).

Nos interventions sont particulièrement pertinentes (en terme de ROI) dans les domaines de la conception, du diagnostic, de la maintenance. Dans l'industrie, la technologie s'applique tant au manufacturing qu'au process qu'il soit discret ou continu.

Présentation des travaux :

Voir l'article complet disponible sur www.ardans.fr

Les enjeux qui président à la décision de mettre en place un système d'aide au diagnostic sont multiples. Nous distinguons trois grands types d'utilisation d'un système d'aide au diagnostic.

- le diagnostic et la maintenance d'une ligne de production ;
- le diagnostic et la maintenance d'un produit manufacturé ;
- le diagnostic pour centres d'appels pour le support informatique.

Pour devenir opérationnelle, la mise en œuvre de solutions d'aide au diagnostic et à la maintenance doit tenir compte de différents facteurs : l'enjeu pour la société industrielle, sa maturité en termes de gestion des connaissances, le niveau technique des utilisateurs, le retour sur investissement attendu.

Après un rappel des principes d'une application de diagnostic nous proposons différentes solutions de complexité croissante pour chacun des types d'utilisation. En sachant que dans notre expérience les deux grandes difficultés à bien prendre en compte sont :

- l'intégration de ces applications à la vie opérationnelle des utilisateurs,
- l'actualisation ou la mise à jour des bases de connaissances.

Des évaluations de différents délais de réalisation sont présentées à titre indicatif.

Via ce choix de système d'aide au diagnostic, la société initie concrètement et s'engage durablement dans une démarche de « mémoire en continu ».

Laboratoire TIMC - IMAG, Institut Bonniot Équipe SIC-Équipe AFIRM

Faculté de Médecine -
Domaine de la Merci
38706 La Tronche Cedex - France

Équipe SIC

Contacts : Catherine Garbay,
04 76 54 94 85
Catherine.Garbay@imag.fr

Équipe AFIRM

Contact : Norbert Noury,
04 76 63 71 11
Norbert.Noury@imag.fr

Cadre Général : Au sein du laboratoire TIMC, deux équipes s'intéressent au diagnostic médical, l'équipe SIC (Systèmes Intégrés Cognitifs) animée par Catherine Garbay et l'équipe AFIRM (Acquisition, Fusion d'Informations et Réseaux en Médecine) animée par Norbert Noury. Un séminaire mensuel sur le thème de la Fusion d'Informations Médicales réunit ces deux équipes et est ouvert aux intervenants extérieurs (contact : Ahlame Douzal Ahlame.Douzal@imag.fr).

Équipe SIC

Membres de l'Équipe concernés
Catherine Garbay (DR2, CNRS), Ahlame Douzal (MC, UJF), Michel Dojat (IR1, INSERM), Anne-Sophie Silvent (Thésarde).

Projet Scientifique : 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 appli-

qués au traitement d'informations spatio-temporelles (signaux, images) avec une application privilégiée au domaine médical. 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.

Panorama des Travaux en Cours :

Dans le cadre du diagnostic, ces travaux sont appliqués à la supervision des malades hospitalisés dans des unités à haute dépendance technique de type service d'anesthésie ou de soins intensifs. Au sein du projet OSSCAR (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]. L'extraction graduelle de paramètres à différents niveaux d'abstraction, à partir de la fusion d'informations de bas niveau, de type signaux physiologiques ou alarmes, nécessite la mise en place d'architecture conceptuelle spécifique où des connaissances hétérogènes, validées par l'expert, sont insérées et manipulées [Silvent02]. Des modèles de similarité entre des séries temporelles sont étudiés pour prendre en compte la spécificité des données de monitoring, tel que le bruit ou l'incomplétude, tout en gardant une complexité de calcul minimale pour permettre une reconnaissance au plus tôt (processus incrémentaux, raisonnement à profondeur variable, etc ...).

Bibliographie :

voir <http://www-u438.ujf-grenoble.fr/people/mdojat/orderedpublications.htm> et <http://www-timc.imag.fr/sic/francais/publications.html>

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).
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.

M. Dojat, N. Ramaux, D. Fontaine: Scenario Recognition for Temporal Reasoning in Medical Domains, Artificial Intelligence in Medicine 14:139-155,1998.

A. Silvent, C. Garbay, P. Carry, and Dojat, M., Data abstraction for medical scenario construction, Proceedings of Workshop on Knowledge Discovery from (Spatio-) Temporal Data. ECAI-2002, Lyon, Fr (21-26 July), 2002.

Équipe AFIRM :

Membres de l'Équipe Concernés :

Norbert Noury (MCU1, UJF), Vincent Rialle (MCU-PH, UJF), Florence Duchêne (Doctorante), Gilles Virone (Doctorant).

Projet Scientifique : Le domaine de recherche de l'équipe AFIRM (Acquisition Fusion d'Information et Réseaux de Médecine) 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 et de la décision d'autre part. Cette multidisciplinarité est mise au service d'un troisième monde, celui de la médecine, abordé essentiellement à travers sa dimension « télé » : télémédecine, télésurveillance médicalisée, aide au maintien 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 d'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 non médicaux concernés.

Panorama des Travaux en Cours :

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, à la détection de situations critiques, à la modélisation de l'activité du patient. L'équipe bénéficie donc, en tant qu'utilisatrice de ces techniques, de la création au sein du laboratoire TIMC de l'axe transversal « Fusion de Données » [Dem 2002]. Au sein du projet TISSAD (RNTS 2000) nous avons créé une architecture « d'Habitat Intelligent pour la Santé » [Nou 2001] [Ria 2001] [Tho 2001] et développé des outils pour l'analyse automatique des rythmes circadiens de l'activité du patient à son domicile [Vir 2002], ainsi que pour la détection automatique d'alarmes en cas de comportements inhabituels [Duc 2002]. Dans le cadre du projet RESIDE-HIS (IMAG 2001) nous nous intéressons, en collaboration avec le laboratoire CLIPS-IMAG, à la fusion des informations de détection spatiale avec les signaux sonores élevés dans l'habitat [Cas 2001] afin d'élaborer des alarmes en cas de détresse de la personne. Le projet VTAMN (RNTS 2001) nous a permis de mettre au point des algorithmes de fusion de capteurs dans le cadre du développement d'un capteur autonome pour la détection des chutes de la personne [Nou 2002].

Bibliographie :

(voir le site : <http://www-timc.imag.fr/AFIRM/>)

[Cas 2001] Castelli E., Istrate D., Rialle V., Noury N., Information extraction from speech in stress situation : application to the medical supervision in a smart house. In : C. Cavé, I. Guaïtella et S. Santi (eds.), *Oralité et gestualité: interactions et comportements multimodaux dans la communication* (Actes du colloque ORAGE), Aix-en-Provence, 18-22 juin 2001, Edition l'Harmattan, pp 362-371.

[Dem 2002] J Demongeot, G Virone, F Duchêne, G Benchetrit, T Hervé, N Noury, V Rialle, « Multi-sensors acquisition, data fusion, knowledge mining and alarm triggering in health smart homes for elderly people », CR Académie des Sciences, Editions scientifiques et médicales Elsevier SAS, 325 (2002) 673-682.

[Duc 2002] Duchêne, F., V. Rialle et N. Noury (2002). Télésurveillance médicale à domicile : proposition d'une architecture pour un système de détection de situations critiques et de décisions sur l'état du patient. Neuvièmes Journées Francophones d'Informatique Médicale (JFIM), Québec, 6-7 mai, Springer-Verlag (sous presse).

[Nou 2001] Noury N., Rialle V., Virone G., 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.

[Nou 2002] Noury N., A smart Sensor for the remote follow up of activity and fall detection of the elderly. In Proc. IEEE-MMB2002, Madison-USA, May 2002, pp 314-317.

[Ria 2001] Rialle V., Noury N., Hervé T., 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.

[Tho 2001] 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. Steenkeste, G. Virone, 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.

[Vir 2002] G. Virone, N. Noury, J. Demongeot, A system for automatic measurement of circadian activity deviations in telemedicine, IEEE Transactions on Information Technology in Biomedicine 2002 (accepté).

Laboratoire d'Automatique de Grenoble LAG Équipe Sûreté, Supervision, Surveillance Diagnostic

ENSIEG- INPG

BP 46

38402 Saint Martin d'Hères Cedex

04 76 82 62 39/62 29

Contact : Jean Marie Flaus/ Sylviane Gentil

Jean-Marie.Flaus@inpg.fr

Sylviane.Gentil@inpg.fr

Thème général : Sûreté, Supervision, Surveillance Diagnostic des systèmes automatisés, avec applications dans le cadre des procédés électriques, chimiques, nucléaires, agro-alimentaires et dans le secteur de la santé.

Membres de l'équipe : O. Adrot, S. Charbonnier, J. M. Flaus, S. Gentil, S. Lesecq, S. Ploix.

Le thème S³D recouvre une thématique développée dans le cadre du Laboratoire d'Automatique de Grenoble. Compte tenu de la complexité croissante des procédés, de l'importance de leur sûreté et du coût de leur indisponibilité, l'activité traditionnelle des automaticiens consistant à concevoir une commande pour un système dynamique aux contours bien délimités, s'est enrichie. Le besoin nouveau consistant à gérer la commande pour prendre en compte des modes de fonctionnement autres que le mode nominal a vu le jour et a donné naissance au développement de techniques de diagnostic pour détecter les défauts et les identifier, et à celles de supervision pour gérer ces situations de crise ou plus simplement anormales.

En effet, un défaut diagnostiqué n'est pas forcément suivi d'un arrêt pour maintenance; il faut souvent continuer à conduire l'installation malgré les anomalies, d'où découlent les problèmes d'aide à l'opérateur. Ces techniques s'appuient sur de nouvelles méthodes de modélisation, permettant d'avoir une vue plus globale du système : comment, en effet, représenter un procédé qui obéit à différents modes de marche, dont la structure peut changer, sans avoir à refaire le modèle mathématique correspondant à chaque situation? C'est un enjeu important si l'on veut convaincre les industriels de l'intérêt des méthodes fondées sur des modèles. Les changements de structure envisagés dans ces situations diverses conduisent tout naturellement vers l'étude des systèmes hybrides. Enfin la nécessité d'une approche formelle pour la validation du comportement du système automatisé est évidente de façon à garantir sa sûreté.

Les actions de l'équipe dans le domaine concernent à la fois la méthodologie et les applications.

Le premier aspect concerne la représentation des systèmes en vue de la sûreté et du diagnostic.

• Le laboratoire a proposé une méthode

originale de **modélisation causale** de procédés continus, fondée sur un graphe représentant le fonctionnement normal de l'installation, dont les nœuds représentent les variables et les arcs les relations entre elles. Les nœuds intègrent l'historique de la variable, la liste des moyens d'action sur la variable, et l'explication de son comportement. Les arcs relient les évolutions des causes à celles des effets, par des relations temporelles inspirées de l'automatique. Le graphe peut être utilisé pour la simulation, l'explication, le conseil d'action.

- Un deuxième volet concerne la modélisation automatique. Un formalisme, nommé YASMA, a été développé pour représenter les systèmes de façon déclarative, et un ensemble d'algorithmes et d'outils ont été mis au point pour générer de façon automatique diverses représentations du système, comme par exemple, le modèle de simulation ou le graphe causal. Cette approche, orientée objet, permet de traiter des systèmes complexes, à dynamique hybride.
- Enfin, pour le diagnostic des systèmes difficiles à modéliser, un axe de recherche a permis de mettre en évidence, parmi les diverses méthodes de traitement du signal, l'intérêt de l'approche fondée sur les ondelettes.

Le deuxième aspect traité au laboratoire concerne le raisonnement de diagnostic.

- La première approche s'appuie sur la représentation par graphe causal. La notion de causalité est utilisée pour relier entre elles des variables suspectes. Une procédure de détection détermine l'ensemble des nœuds en défaut. Une procédure de localisation construit le sous-graphe de propagation du défaut source jusqu'à l'ensemble de ses conséquences, avec gestion du temps. Le graphe est utilisé de plus comme moyen d'explication à l'opérateur dans une interface de supervision.
- La seconde approche de diagnostic développée dans l'équipe est une approche logique dont l'objectif est d'étendre aux systèmes dynamiques le raisonnement de diagnostic originellement introduit par Reiter, en le combinant avec des résultats de l'approche

issue de l'automatique. Il s'agit en particulier de séparer la détection de symptômes, de la localisation de défauts afin de pouvoir raisonner sur tout type de tests. Des architectures logicielles orientées objets et agents sont développées dans ce cadre.

Enfin, un dernier aspect méthodologique concerne l'approche ensembliste pour représenter l'incertitude au niveau du modèle. Deux directions sont explorées.

- La première consiste à représenter l'incertitude par un intervalle au niveau de l'erreur de sortie et des algorithmes de type « erreur de sortie inconnue mais bornée » pour l'estimation des paramètres du modèle ont été développés.
- La seconde voie étudiée concerne l'estimation non linéaire ensembliste, dont l'objectif est de reconstruire l'ensemble des états possibles pour un système, à partir de l'observation d'un vecteur de mesures à erreur bornée.

Ces travaux sont à la base d'une détection de défaut prenant en compte les incertitudes de modèle. Par ailleurs, nous explorons la possibilité d'utiliser ces techniques ensemblistes pour développer des algorithmes de preuve de sûreté.

Ces concepts et techniques ont été appliqués et validés dans différents domaines comme le retraitement nucléaire, la pétrochimie, la chimie, la pharmacie et les procédés agroalimentaires ainsi que dans le domaine médical, dans lequel deux actions sont conduites dans le cadre du projet Medidiag : la première porte sur le monitoring de patients et la réduction du taux de fausses alarmes en unités de soins intensifs et la seconde porte sur le diagnostic automatique de l'urgence de l'évacuation de blessés au combat, à partir de mesures à distance de variables biologiques.

Projets en cours :

- **Projet européen MAGIC** : Multi-agents based diagnostic data acquisition and management in complex systems (Janvier 2002-Décembre 2005). Ce projet a pour objectif de construire un système réparti de diagnostic, reposant sur le concept d'agents. Il vise les installations complexes et intègre donc

des outils d'aide à la décision pour les opérateurs humains ainsi que des interfaces avancées.

• **Projet AGROSUR région Rhone-Alpes** (Septembre 2001- Aout 2004) .

L'objectif de ce projet est de développer une méthode d'analyse des risques systématique, intégrant les approches traditionnelles du secteur comme HACCP, et capable de générer un modèle de dysfonctionnement utilisable pour le raisonnement de diagnostic.

- **Projet Sage** : L'objectif de ce projet supporté par une équipe multidisciplinaire est de fédérer des recherches dans le domaine de la sûreté et de la sécurité des procédés agroalimentaires et environnementaux. Ce domaine concerne des processus particulièrement mal connus. Ce projet nécessite donc des compétences multiples, à l'intersection de plusieurs disciplines. Les axes de recherche proposés sont à la fois méthodologiques et applicatifs.

• **Projet MEDIDIAG (Novembre 2000 - Novembre 2003)** :

L'objectif de ce projet est le développement d'un système de surveillance de patients hospitalisés dans les unités de réanimation, sous assistance respiratoire, intégrant la forme des signaux biologiques et leur correspondance. Ce projet est réalisé dans le cadre d'un projet du Réseau National de Technologies pour la Santé, intitulé OSSCAR (Optimisation des Stratégies à partir des Signaux Cardio-respiratoires en Anesthésie Réanimation). Son objectif est de détecter rapidement une dégradation de l'état de santé du malade, tout en étant robuste aux perturbations pouvant affecter les variables enregistrées, comme des artefacts sur le signaux, l'agitation du patient ...

Notre équipe est reliée aux réseaux nationaux se préoccupant du diagnostic dans le cadre de l'automatique (GDR-RTP *Fiabilité, Diagnostic, Tolérance aux fautes*) et de l'Intelligence Artificielle (*groupe de travail IMA-LAIA*).

Bibliographie

1. B. Heim, S. Cauvin, S. Gentil, *FCC process diagnosis based on a causal and heuristic approach*, IFAC Workshop On line fault detection and supervision in the chemical process industries, Seoul (Corée), 8-9 June 2001.
2. C. Garcia-Beltran, S. Gentil, *Causal reasoning for human supervised process reconfiguration*, IEEE International Symposium on Intelligent Control ISIC2001, Mexico (Mexico), 5-7 septembre 2001.
3. Valdes-Gonzalez H., Flaus J.M., 2001, *State Estimation In a Bioprocess Described By a Hybrid Model*, Proceedings of the IEEE Joint International Conference on Control Applications & International Symposium on Intelligent Control (ISIC), Mexico city, (Mexico), September 2001.
4. Valdes-Gonzalez H., Flaus J.M., 2001, *A Globally Convergent State Estimation Approach Based On Interval Analysis*, Proceedings of the European Control Conference (ECC'2001), Porto (Portugal), September 2001.
5. I. Fagarasan, S. Ploix, S. Gentil, *Bounding approach for fault detection and diagnosis*, IFAC Symposium Large Scale Systems LSS 2001, Bucarest (Roumanie), 18-20 July 2001.
6. Ploix S., Follot C., 2001, *Fault Diagnosis Reasoning For Set-Membership Approaches And Application*, IEEE International Symposium on Intelligent Control (ISIC), Mexico City (Mexico), 5-7 septembre 2001.
7. Charbonnier S., Becq G., Biot L., Carry P.Y., Perdrix J.P., 2001, *On Line Segmentation Algorithm For Icu Continuously Monitored Clinical Data*, 15th World IFAC Congress, Barcelone (Espagne).
8. Suzanne Lesecq, Siana Petropol, Alain Barraud, *Asynchronous Motor Parametric Faults Diagnosis using Wavelet Analysis*, SDEMPED'2001, Grado-Gloriza, Italy, 1-3 sept 2001.

France Télécom R&D
Site de Lannion

Christophe Dousson

France Télécom R&D, DTL/TIC
Technopole Anticipa
22, avenue Pierre Marzin
22307 Lannion cedex
Tél: (+33) 2 96 05 21 99
Christophe.dousson@francetelecom.com

Historiquement, la recherche sur le diagnostic à France Télécom R&D remonte à l'étude des systèmes experts pour la supervision de réseaux de télécommunications. Depuis le début des années 90, l'activité en ce domaine s'est orientée vers le diagnostic à base de modèles afin de prendre en compte l'évolution rapide des équipements ainsi que la multiplication de nouveaux types de réseaux. Le réseau télécom a néanmoins jusque là toujours été considéré comme un tout, avec un centre de supervision unique. Plus récemment, la supervision du réseau a évolué vers une solution répartie afin de mieux correspondre à la réalité et de mieux s'adapter à la taille croissante des réseaux. Aussi, les nouvelles techniques actuellement étudiées par France Télécom R&D se situent dans la problématique du diagnostic de systèmes répartis.

Parallèlement, le métier d'opérateur de télécoms évolue fortement vers un métier de fournisseur de services plutôt que de simple « transporteur » de voix ou de données et, de fait, l'activité de diagnostic à France Télécom R&D est passée progressivement du niveau des équipements à un niveau logiciel (ou middleware) : celui des services offerts par le réseau.

Les activités de l'équipe portent sur l'étude d'une chaîne complète de supervision d'un réseau de télécommunications en intégrant des outils d'aide à l'acquisition d'expertise, à la modélisation de système ainsi que des outils de supervision (corrélation d'alarmes) et de diagnostic.

L'approche principale étudiée jusqu'alors est fondée sur un formalisme bien adapté à la représentation des évolutions d'un système observé [3]. Un outil de suivi de systèmes dynamiques s'appuyant sur ce formalisme a été développé (CRS pour Chronicle Recognition System, <http://crs.elibel.tm.fr>) et est régulièrement amélioré [2]. Du point de vue de l'apprentissage pour l'acquisition d'expertise, un effort particulier est mené sur l'analyse fréquentielle des journaux d'alarmes dans lesquels il s'agit d'isoler des phénomènes récur-

rents afin de focaliser l'attention des experts [4].

France Télécom R&D est impliqué dans divers projets collaboratifs. France Télécom R&D a notamment dirigé le projet RNRT Magda axé sur le diagnostic de réseaux de télécommunications (<http://magda.elibel.tm.fr>, 1998-2001) : il s'agissait de développer et d'expérimenter de nouvelles méthodes de gestion des alarmes et, plus précisément, de permettre une meilleure compréhension des défaillances ou des pannes à l'aide d'outils d'acquisition d'expertise, puis de reconnaître en ligne des situations à risques par des outils de corrélation d'alarmes et de diagnostic. Un nouveau projet, Magda2, est chargé d'étudier les répercussions des problèmes réseaux sur la qualité de service (2002-2003).

Membres concernés par le diagnostic:

- Christophe Dousson
(christophe.dousson@francetelecom.com)
- Irène Grosclaude
(irene.grosclaude@francetelecom.com)
- Bruno Guerraz, doctorant
(bruno.guerraz@francetelecom.com)
- Hélène Verrière
(helene.verriere@francetelecom.com)

Bibliographie

- (<http://perso.rd.francetelecom.fr/dousson>):
- [1] Modeling Fault Propagation in Telecommunications Networks for Diagnosis Purposes
A. Aghasaryan, C. Dousson, E. Fabre, A. Osmani, Y. Pencolé.
18th World Telecommunications Congress (WTC'02)
Septembre 2002.
 - [2] Extending and Unifying Chronicle Representation with Event Counters
C. Dousson.
Proc. of the 15th ECAI
F. van Harmelen, IOS Press, pp. 257--261, Juillet 2002.
 - [3] Alarm Driven Monitoring Based on Chronicles
M. O. Cordier, C. Dousson.
Proc. of the 4th Symposium on Fault Detection Supervision and Safety for Technical Processes (SAFEPROCESS)
IFAC, A.M. Eldemayer., pp. 286--291, Juin 2000.
 - [4] Discovering chronicles with numerical

time constraints from alarm logs for monitoring dynamic systems
C. Dousson, T. Vu Duong.
Proc. of the 16th IJCAI
pp. 620-626, 1999

Équipe SFSD – Sûreté de Fonctionnement des Systèmes Dynamiques – Université de Lille

LAIL – UMR 8021
Ecole Polytechnique Universitaire de
Lille
Université des Sciences et Technologies
de Lille
59655 Villeneuve d'Ascq cedex
<http://www.univ-lille1.fr/lail/>

Personne à contacter :

Marcel Staroswiecki, Professeur
marcel.staroswiecki@univ-lille1
Tel. secrétariat : 03 20 43 45 65

Membres de l'équipe concernés par le thème du diagnostic

Mireille Bayart, Professeur
(EUDIL). Marcel Staroswiecki,
Professeur (EUDIL). Denis Pomorski,
Professeur (IUT A). Abdel Aitouche,
Enseignant HEI.
Cyrille Christophe, MdC (IUT A).
Vincent Cocquempot, MdC (IUT A).
Blaise Conrard, MdC (EUDIL). Anne-
Lise Gehin, MdC (EUDIL). Bin Jiang,
CDD (CNRS). Belkacem Ould
Bouamama, MdC (IUT A).

Description des travaux

L'équipe développe des modèles et des méthodes d'analyse en vue de la conception et de la mise en œuvre d'algorithmes destinés aux couches « surveillance, diagnostic et tolérance aux fautes » des machines et systèmes intelligents. Les principaux thèmes développés dans l'équipe sont les suivants :

1. Surveillance avec modèle (applications en chimie, pétrochimie, systèmes énergétiques, systèmes électromécaniques, systèmes de transport, ...)

1.1. Modèles numériques : sont particulièrement étudiés les systèmes dont le

comportement est décrit par des équations différentielles et algébriques non linéaires polynomiales. Les méthodes de redondance analytique permettent de détecter et localiser les défaillances internes (défaillances des composants du système) alors que la surveillance des contraintes inégalité permet de détecter et localiser des défaillances externes (non conformité des conditions opératoires).

1.2. Modèles structurels, Bond-Graphs et Graphes causaux : les modèles structurels sont une abstraction des modèles de comportement, qui permettent une analyse fondée sur la théorie des graphes. Les Bond-graph sont des modèles qui décrivent les transferts de puissance et d'énergie sous une forme accessible aux analyses graphiques. Les graphes causaux sont issus de l'analyse structurelle et permettent, entre autres, la localisation des défaillances initiales, c'est à dire le filtrage des alarmes.

1.3. Décision : les méthodes de décision doivent permettre de détecter et localiser les défaillances en minimisant les erreurs (fausses alarmes, non détections, retards à la détection, erreurs de localisation, etc.), malgré les incertitudes concernant le modèle et son environnement. Leur optimisation relève des approches statistiques (lorsque les incertitudes sont décrites par des modèles stochastiques) ou des approches ensemblistes (lorsque les incertitudes sont simplement bornées).

2. Surveillance sans modèle

2.1. Aide au diagnostic par des méthodes d'apprentissage supervisé, utilisant des critères informationnels et les arbres de décision (applications pour l'aide au diagnostic et au pronostic en soins intensifs).

2.2. Optimisation des systèmes de détection, par maximisation de la quantité d'information pertinente pour la détection. Applications en fusion d'informations, par optimisation du compromis entre la quantité d'information transmise à l'opérateur de fusion et les performances demandées au système de décision.

3. Tolérance aux fautes

3.1. Surveillance distribuée, instru-

mentation intelligente et placement de capteurs/actionneurs : L'implémentation des algorithmes de surveillance peut être centralisée ou répartie, justifiant ainsi les capteurs et actionneurs intelligents. L'équipe a développé un modèle générique d'instrument intelligent, fondé sur les notions de services (et de leurs versions) et de modes d'utilisation, qui permet une analyse systématique des possibilités de reconfiguration lorsque certains services sont perdus suite à des défaillances. Les outils d'analyse structurelle permettent alors de proposer la mise en œuvre d'ensembles de capteurs / actionneurs à même de respecter des spécifications données de tolérance aux fautes.

3.2. Accomodation des défaillances / reconfiguration du système : On s'intéresse ici aux conditions sous lesquelles les capteurs (actionneurs) du système peuvent continuer à fournir les estimations (appliquer les commande), soit avec les performances nominales, soit en mode dégradé, en cas de défaillances. La reconfigurabilité d'un système est analysée selon deux approches. La première repose sur une modélisation au moyen d'un graphe structurel enrichi par une interprétation causale. On peut ainsi interpréter en termes graphiques les propriétés classiques d'observabilité et de commandabilité, et l'analyse de la reconfigurabilité consiste alors à voir dans quelle mesure ces propriétés restent vérifiées en présence de défaillances

Deux projets représentatifs

Projet européen CHEM : Decision support systems in CHEMical / petrochemical processes
www.chem-dss.org

Le projet CHEM regroupe 15 partenaires européens, industriels et laboratoires, dans le but de développer et intégrer des boîtes à outils pour la conception et la mise en œuvre de systèmes de supervision de grands processus de production à risques, tels qu'on en trouve dans les industries chimiques, pétrochimique, de production d'énergie, etc.

Le laboratoire réalise dans ce cadre

un logiciel d'aide à la modélisation des systèmes complexes, fondé sur une bibliothèque de composants industriels, auxquels sont associés leurs modèles de comportement et des méthodes d'interconnexion. Le système complet est d'abord validé d'un point de vue physique grâce à des approches de validation issues de la théorie Bong-graph, et ensuite analysé d'un point de vue structurel pour :

- 1) vérifier la faisabilité d'un cahier des charges donné de détection et localisation des défaillances ;
- 2) construire automatiquement les indicateurs de défaut qui devront être implémentés et analysés en temps réel ;
- 3) fournir les graphes causaux permettant de remonter aux défauts initiateurs.

Les boîtes à outils développées dans le cadre de ce projet sont testées sur plusieurs sites pilotes dont un générateur de vapeur disponible au laboratoire.

Réseau européen DAMADICS :
Development and Application of Methods for Actuator Diagnosis in Industrial Control Systems

<http://www.eng.hull.ac.uk/research/control/damadics/>

Ce réseau regroupe 8 universités autour du thème du diagnostic des actionneurs. Le projet vise à développer et mettre en œuvre des algorithmes de surveillance et de tolérance aux fautes, sur un ensemble de vannes de régulation du transport de jus de betteraves, dans une sucrerie polonaise (Lubin). L'équipe SFSD apporte sa compétence en conception d'algorithmes de surveillance utilisant la redondance analytique, et en architecture fonctionnelle d'actionneurs intelligents.

Position par rapport à d'autres travaux, nationaux ou internationaux

- l'équipe a une activité originale dans le domaine de l'analyse structurelle et de l'approche Bond-graph pour la surveillance et le filtrage d'alarmes,
- elle appartient au groupe leader pour les algorithmes de surveillance par redondance analytique, et pour l'analyse de la tolérance aux fautes

Activités d'animation dans le domaine

En France :

- Animation du groupe Sûreté de fonctionnement de l'Action Nouvelle de Recherche sur les transports et leur sécurité dans le Nord Pas-de-Calais
- Animation du Réseau Thématique Pluridisciplinaire : Fiabilité, Diagnostic et Tolérance aux fautes (RTP n°20) du CNRS

- Animation du groupe de travail Composants Intelligents pour l'Automatisation et la Mesure de la SEE

Au niveau international

- Participation à trois comités techniques de l'IFAC (Safeprocess, Components and Instruments, Low Cost Automation)

Courte bibliographie

1) Ouvrages

In *Diagnostic : Approches par modèles numériques et statistiques*, B. Dubuisson (coord.), Hermès, Paris, 2001

- M. Staroswiecki : Redondance analytique (chapitre 1)
- B. Ould Bouamama, M. Staroswiecki, R. Litwak : Surveillance d'un générateur de vapeur (chapitre 10)

In *Control of Complex Systems*, K. J. Aström, P. Albertos, M. Blanke, A. Isidori, W. Schaufelberger, T. Sanz ed., Springer 2000.

- M. Blanke, C. Frei, F. Kraus, R. J. Patton and M. Staroswiecki : Fault Tolerant Control Systems
- Isidori, M. Kinnaert, V. Cocquemot, C. De Persis, P. M. Frank and D. N. Shields : Residual Generation for Fault Detection and Isolation in Non-linear Systems.
- J. Lunze, J. Askari-Marnani, A. Cela, P. M. Frank, A.-L. Gehin, B. Heimig, M. Lemos, T. Marcu, L. Rato and M. Staroswiecki : Three Tank Control Reconfiguration
- R. Izadi-Zamanabadi, P. Amann, M. Blanke, V. Cocquemot, G. L. Gissinger, E. C. Kerrigan, T. F. Lootsma, J. M. Perronne, G. Schreier : Fault Detection and Reconfigurable Control for a Ship

2) Revues avec comité de lecture

M. Staroswiecki (2000), Quantitative and Qualitative Models for Fault Detection and Isolation, *Int. J. of Mechanical Systems and Signal processing*, 14 (3) :301-325.

B. Darkhovski, M. Staroswiecki, On hypotheses testing problem under unknown parameter, *Publications de l'Institut de recherche en Mathématiques de Lille*, Vol. 55, n° V, pp. 1-18.

M. Staroswiecki, A.-L. Gehin (2001), From

Control to Supervision, *Annual Reviews in Control*, 25 : 1 – 11.

G. Hoblos, M. Staroswiecki, A. Aïtouche (2001), Sur la Tolérance aux Fautes de Capteurs et d'Actionneurs, *Journal Européen des Systèmes Automatisés*, Vol. 35(3) : 331-352.

Staroswiecki, M., Comtet-Varga, G. (2001), Analytic redundancy relations for fault detection and isolation in algebraic dynamic systems (Regular paper), *Automatica* Vol. 37(5) : 687-699.

3) Conférences internationales

Bin Jiang, Marcel Staroswiecki, Vincent Cocquemot (2002), Fault identification for a class of time-delay systems, *American Control Conference ACC'02*, Anchorage (Al).

Boris Darkhovski and Marcel Staroswiecki (2002), FDI decision making using a game formulation, *IFAC World Congress*, Barcelona (Sp).

Marcel Staroswiecki (2002), On reconfigurability with respect to actuator failures, *IFAC World Congress*, Barcelona (Sp).

B. Ould Bouamama (2001). Model builder for thermofluid systems using Bond Graph and Functional modelling, *Work shop on Bond graphs. Part of 13th European Simulation Symposium and Exhibition in Industry*. Marseille

Jiang, B., Staroswiecki, M. and V. Cocquemot (2001), Fault Diagnosis for a Class of Nonlinear Systems with Unknown Parameters, *4th IFAC Workshop on On-line Fault Detection and Supervision in the Chemical Process Industries (CHEMFAS-4)*, Seoul, Korea.

B. Ould Bouamama, A.L. Gehin, M. Staroswiecki (2001), Alarm Filtering by Component Modelling and Bond Graph approach, *4th IFAC Workshop on On-line Fault Detection and Supervision in the Chemical Process Industries (CHEMFAS-4)*, Seoul, Korea.

M. Blanke, M. Staroswiecki and N. E. Wu (2001), Concepts and Methods in Fault-tolerant Control, (Tutorial), *American Control Conference, ACC'01*, Arlington, (Va)

Adresse d'un site web où l'on peut trouver l'ensemble des références et articles.

<http://www.univ-lille1.fr/lail/>

Laboratoire d'Informatique de Paris-Nord (LIPN-UMR 7030)

LIPN, Institut Galilée, Université Paris13, 99 av. J.-B. Clément, 93430 Villetaneuse
<http://www-lipn.univ-paris13.fr>

Équipe : Apprentissage, diagnostic et Agents (ADAge)

Responsable : Philippe Dague
(01 49 40 36 17,
dague@lipn.univ-paris13.fr)

L'équipe ADAge est fédérée autour de trois axes de recherche en forte synergie : l'apprentissage symbolique, connexionniste et statistique ; le diagnostic à base de modèles de systèmes complexes ; l'étude des aspects distribués et coopératifs dans les systèmes d'agents autonomes.

Participants (permanents) au thème diagnostic :

Diagnostic à base de modèles : Philippe Dague, François Lévy (équipe Représentation des Connaissances et Langage Naturel).

Apprentissage symbolique et diagnostic : Aomar Osmani.

Apprentissage numérique et diagnostic : Younès Bennani, Emmanuel Viennet.

Diagnostic et agents pour les télécommunications : Francine Krief.

Résultats nouveaux et perspectives de recherche

Nous pouvons mettre en avant les contributions suivantes de ces deux dernières années et dresser les projets de recherche qui les prolongent, dans un avenir proche ou plus lointain.

1. Conception et diagnostic : transformation de modèles et intégration

Le projet européen Brite-Euram III VMBD, « *Vehicle Model-Based Diagnosis* », (02/1996-03/1999), auquel participaient pour le LIPN Philippe Dague et Jean-Philippe Robles (post-doc), avait démontré la faisabilité des techniques de diagnostic à base de modèles pour le diagnostic à bord de *véhicules automobiles*. Lui faisant suite, le travail entrepris dans le cadre du projet européen du 5^{ième} PCRDT IDD, « *Integrated Design process for on-board Diagnosis* », (02/2000-01/2003), qui réunit CRF (Fiat), Daimler-Chrysler, PSA, Renault, Magneti-Marelli, OCC'M et les universités de

Munich, de Turin et de Paris 13, a pour objectif d'intégrer l'approche du raisonnement à base de modèles au processus de conception et de développement, afin de prendre en compte dès cette phase les nécessités du diagnostic. Le LIPN (Philippe Dague et Yuhong Yan, post-doc), a réalisé une maquette d'un logiciel qui extrait et discrétise automatiquement un modèle de CAO (en Matlab/Simulink) pour obtenir un modèle qualitatif de diagnostic qui alimente l'outil diagnostic RAZ'R. Ce travail se poursuit pour résoudre les problèmes posés par cette abstraction de modèles, comme le choix automatique des seuils et la prise en compte de la dynamique, ainsi que pour l'étude de la diagnosticabilité (déteabilité, discriminabilité, placement de capteurs). Plus largement, une EoI (« *Expression of Interest* ») SAFECAR (*Tools for Improving Vehicle Availability, Reliability and Safety*) de projet intégré a été soumise dans le cadre du 6^{ième} PCRDT.

2. Utilisation de techniques d'apprentissage

Une collaboration très étroite a commencé depuis un an avec le thème apprentissage de l'équipe ADAge pour des problèmes de diagnostic.

Dans le domaine de la *supervision de réseaux de télécommunications*, le LIPN (Philippe Dague, Francine Krief, François Lévy et les doctorants Serge Bibas puis Aomar Osmani) avait participé à deux contrats successifs (11/1993-10/1996 et 11/1997-10/2000) avec le CNET, en collaboration avec l'IRISA (et le LAAS pour le premier), sur le thème « *Modélisation, Diagnostic et Supervision de réseaux de télécommunications* », puis au projet RNRT (*Réseau National de Recherches en Télécommunications*) exploratoire MAGDA « *Modélisation et Apprentissage pour une Gestion Distribuée des Alarmes* » (11/1998-11/2001), en collaboration avec trois équipes de l'INRIA/IRISA (PAMPA, SIGMA2, AIDA), deux de France Télécom R&D, Alcatel et ILOG. En continuation de MAGDA, il participe actuellement (Philippe Dague, Francine Krief, Aomar Osmani, Emmanuel Viennet et Larbi Aït Mohamed, doctorant) au projet RNRT

exploratoire MAGDA2 « *Modélisation et Apprentissage pour une Gestion Distribuée des Alarmes de bout en bout* » (11/2001-11/2003) avec les mêmes partenaires. MAGDA2 étend la problématique de MAGDA à la gestion de réseaux hétérogènes et à la prise en compte de l'interaction entre la couche réseau et la couche service. L'effort du LIPN porte sur l'utilisation de techniques d'apprentissage, tant statistique (corrélation d'alarmes ou d'indicateurs de qualité de service pour du diagnostic prédictif) que symbolique (apprentissage de chroniques).

Il en est de même dans le projet RNTL (*Réseau National de recherche et d'innovation en Technologies Logicielles*) pré compétitif RESEDA, « *Réalisation d'un Environnement Système Evolutif pour le Diagnostic Automobile* » (06/2001-12/2003), en collaboration avec Renault, Silicomp et Trialog, qui a pour but de développer, aux fins de déploiement sur les véhicules et dans les stations de dépannage, une nouvelle infrastructure de composants logiciels de diagnostic (utilisant les techniques d'objets distribués et la technologie de code mobile Java), qui soit évolutive avec le cycle de vie des véhicules. Le LIPN (Younès Bennani, Philippe Dague et Christine Choppy, de l'équipe Logique, Calcul et Raisonnement) aborde la problématique du *diagnostic évolutif* avec des techniques d'apprentissage connexionniste (analyse de séquences temporelles), voire symbolique, à côté des techniques à base de modèles. La coopération de ces techniques pourrait être envisagée dans un futur plus lointain, comme l'utilisation de méthodes d'apprentissage pour l'acquisition ou la mise à jour de modèles explicites ou la cohabitation de modèles explicites et de modèles boîtes noires appris.

3. Comparaison et complémentarité des approches à base de modèles issues de l'Intelligence Artificielle et de l'Automatique

Cette recherche est menée depuis 1999 au sein du groupe de travail IMA-LAIA « *Intégration de Modèles ALiant l'Automatique et l'Intelligence Artificielle* » des PRC-GDR Automatique et I3 ainsi que de l'AFIA, qui réunit

Philippe Dague, François Lévy (de l'équipe RCLN), Marie-Odile Cordier (IRISA-INRIA), Jacky Montmain (EMA-CEA), Marcel Staroswiecki (LAIL-CNRS) et Louise Travé-Massuyès (LAAS-CNRS). Les hypothèses parfois implicites des approches fondées sur la cohérence logique en IA et sur les relations de redondance analytique en Automatique ont été explicitées et confrontées et des conditions suffisantes d'équivalence des résultats ont été fournies selon ces hypothèses, dans un cadre atemporel. Ce travail de recherche va à présent se poursuivre dans le cadre temporel : une proposition a été soumise au Réseau Thématique Pluridisciplinaire « Fiabilité, Diagnostic et Tolérance aux Fautes des systèmes complexes » ainsi qu'une *EoI BRIDGE* (« *Diagnosis and Monitoring of complex systems: Bridging the methodologies of the AI and Control Communities* ») de projet intégré dans le cadre du 6ième PCRDT.

4. Autonomie : diagnostic à bord

Elle est étudiée depuis plus de trois ans par Philippe Dague, avec Louise Travé-Massuyès du LAAS-CNRS, en ce qui concerne les satellites. D'abord au cours d'un travail de consultance pour le CNES (12/1998-3/1999) dans le cadre d'un dossier exploratoire sur le thème de l'autonomie des satellites, prolongé par une étude sur l'architecture logicielle pour les satellites autonomes (06/2000-11/2001). Ensuite au sein d'un projet de l'ESA concernant l'autonomie des futures constellations de satellites, réunissant le LAAS-CNRS (avec le LIPN comme sous-contractant), le CERT-Onera et Astrium, dans lequel Philippe Dague et Louise Travé-Massuyès ont été responsables en 2001 du lot concernant le diagnostic à bord (détection, localisation de fautes et reconfiguration). Enfin, dans le cadre d'une thèse du CNES avec un co-financement d'Astrium (10/1999-10/2002), où Emmanuel Bénazéra, sous la direction de Louise Travé-Massuyès et co-encadré par Philippe Dague, a réalisé une extension des travaux de la NASA pour le suivi d'états d'un système, du cas discret au cas hybride. Ce travail se prolonge à présent par le traitement autonome de la reconfiguration, par

couplage avec un module de planification. Une intégration sur des bancs de test du CNES est envisagée et une *équipe projet multi-laboratoires* est en cours de soumission sur le thème « *Diagnostic at applications spatiales* ». A un horizon plus lointain, des modules de coopération, d'aide à la décision, voire d'apprentissage seraient nécessaires dans l'architecture d'agent autonome.

5. Etude du diagnostic distribué dans un cadre multi-agents

Cet axe est plus prospectif. La *supervision de réseaux de télécommunications* a déjà nécessité une approche décentralisée, puisque l'aspect distribué a été le point le plus novateur du projet MAGDA, dans lequel le LIPN a été chargé de la modélisation des équipements et d'un réseau prototype. Ces modèles, développés sous forme d'automates ou de pièces, alimentent les algorithmes distribués développés à l'INRIA/IRISA, qui ont été testés sur la plate-forme de gestion de réseau SDH ALMAP-IF d'Alcatel. Le caractère distribué du diagnostic sera sans doute aussi à l'œuvre dans des applications comme les *futures constellations de satellites*. Au-delà du diagnostic distribué, un axe qui nous semble porteur est celui du diagnostic dans un cadre de systèmes multi-agents. Une étude préliminaire, associant les deux thèmes SMA (Amal El Fallah-Seghrouchni et Guillaume Vauvert, doctorant) et Diagnostic (Philippe Dague et Aomar Osmani) de l'équipe ADAGE, a eu lieu (09/2000-03/2001) avec Thomson-CSF Detexis, maintenant Thales Systèmes Aéroportés (Patrick Taillibert), qui a conduit à la spécification en termes d'*agents logiciels autonomes* d'un système de diagnostic de cartes électroniques. Nous espérons poursuivre plus avant cette investigation vers une architecture multi-diagnostiqueurs ou multi-superviseurs.

Bibliographie

Théorie logique du diagnostic à base de modèles, P. Dague, in *Diagnostic, Intelligence Artificielle et Reconnaissance de Formes*, sous la direction de Bernard Dubuisson, traité IC2, Hermès, ISBN 2-7462-0249-2, pp 17-105, 2001.
IDD: Integrating Diagnosis in the Design of

automotive systems, C. Picardi, R. Bray, F. Cascio, L. Console, P. Dague, O. Dressler, D. Millet, B. Rehfus, P. Struss et C. Vallée, *15th European Conference on Artificial Intelligence ECAI-02*, Lyon, juillet 2002.

Distributed model for an SDH telecommunication network diagnostics, P. Dague, F. Krief, F. Lévy et A. Osmani, International Conference on Information Systems ICIS'2001, Arlington, VI, juin 2001.

Modélisation d'un anneau SDH pour les besoins de la gestion de fautes, F. Krief, A. Osmani, *Annales des télécommunications*, tome 57, pp. 125-149, janvier/février 2002.

Modular Connectionist Modelling and Classification Approaches for Local Diagnosis in Telecommunication Traffic Management, Y. Bennani ET F. Bossaert, *International Journal of Computational Intelligence and Applications*, World Scientific Publishing Company, 2001.

A comparative analysis of AI and control theory approaches to model-based diagnosis, M. -O. Cordier, P. Dague, M. Dumas, F. Lévy, J. Montmain, M. Staroswiecki et L. Travé-Massuyès, *14th European Conference on Artificial Intelligence ECAI-00*, Berlin, août 2000 (et aussi *11th International Workshop on Principles of Diagnosis DX'00*, Morelia, Mexique, juin 2000).

State Tracking of Uncertain Hybrid Concurrent Systems, E. Benazera, P. Dague et L. Travé-Massuyès, *13th International Workshop on Principles of Diagnosis DX'02*, Semmering, Autriche, mai 2002.

IFP

Institut Français du Pétrole

1-4 avenue de Bois Préau,
92852 Rueil Malmaison Cedex,
33 1 47 52 70 00 et CEDI BP3 69 370
Vernaison, 33 4 78 02 20 20
www.ifp.fr

Projet CHEM : Advanced decision support systems for Chemical/Petrochemical manufacturing processes

Responsable : Sylvie Cauvin (Chef de projet pour la partie Européenne) : Sylvie.Cauvin@ifp.fr, tél : 33 1 47 52 70 17
Bertrand Braunschweig (Chef de projet pour la partie internationale IMS) : Bertrand.Braunschweig@ifp.fr, tél : 33 1 47 52 66 48

Autres personnes concernées : Benoît Celse : Benoit.Celse@ifp.fr,

tél : 33 4 78 02 21 76

Bruno Heim : Bruno.Heim@ifp.fr,

tél : 33 4 78 02 21 86

Thème général : L'activité en diagnostic de l'IFP relève de projets de recherche auxquels contribuent des chercheurs et ingénieurs ayant des compétences dans le domaine. Il n'y a pas d'équipe dédiée à cette activité ; les principales compétences sont regroupées au sein de la division informatique scientifique et mathématiques appliquées, une division de recherche de l'IFP. Depuis une dizaine d'années, nous développons des applications pour le métier pétrolier : de l'exploration pétrolière à la combustion dans les moteurs en passant par l'étude des réservoirs, le forage, les systèmes de production offshore, le transport, le raffinage et la pétrochimie, l'économie du pétrole et du gaz.

Actuellement, outre un système de diagnostic d'alarmes de forage, l'activité porte essentiellement sur la supervision de procédés de raffinage et de pétrochimie dans le cadre du projet CHEM que l'IFP coordonne. L'IFP a fait partie du groupe de travail « Alarmes » qui s'est prolongé par le groupe « IMALAIA » soutenu par le GDR IA et le GDR Automatique.

Le projet Chem : L'automatisation croissante des procédés conduit à une forte augmentation de la quantité d'information présentée aux opérateurs qui sont amenés à effectuer des diagnostics complexes, très rapidement pour décider des actions correctrices à entreprendre. L'objectif du projet CHEM qui s'inscrit dans le 5ème programme cadre de la commission Européenne sur la période 2001-2004, est de développer et de faire collaborer des outils de diagnostic et d'aide à la décision qui mettent en œuvre des technologies variées -automatique, statistique, graphes causaux, logique floue, à base d'intervalle, etc. - afin de donner une information ciblée aux opérateurs et d'améliorer le fonctionnement des unités. Le projet réunit 15 partenaires industriels et universitaires européens de 8 pays différents : Espagne, Finlande, France, Grande-Bretagne, Norvège, Pays-Bas, Pologne et Suède. 22 boîtes à outils ou « tool-

boxes » seront développées dans le cadre du projet et des ensembles de « toolboxes » seront testés sur 7 unités industrielles ou pilote tournant des procédés aussi différents que le FCC (Procédé de craquage catalytique) ou un procédé de fabrication de papier. Une méthodologie de développement de nouvelles applications de supervision est mise au point, ainsi que la communication entre les « toolboxes » sous forme modulaire utilisant les technologies XML, Corba, Java ainsi que G2 qui sert de plate-forme d'intégration. Des informations détaillées sur le projet se trouvent sur www.chem-dss.org. Il sera étendu par un volet international dans la structure IMS (International Manufacturing Systems).

Rôle de l'IFP dans le projet Chem :

- L'IFP est chargé :
- de la coordination européenne et internationale du projet,
 - de la coordination du « work package 1 : Méthodologie d'intégration et de déclinaison » : notre objectif est de bien définir comment construire rapidement de nouvelles applications de supervision de procédés, et notamment, de pouvoir utiliser les outils sur plusieurs de nos procédés,
 - de la réalisation de 2 boîtes à outil mettant en œuvre des techniques de diagnostic à base de modèles causaux semi-quantitatifs exploités par des raisonnements utilisant de la logique floue et de la logique à base d'intervalles. Ceci se fait dans le cadre d'une thèse avec le LAG à Grenoble et le LAAS à Toulouse
 - d'une partie des tests industriels, l'IFP intégrera les boîtes à outils adéquates pour construire un système de supervision adapté au pilote de FCC.

Bibliographie

- [Heim01] HEIM B., CAUVIN S., GENTIL S. FCC process diagnosis based on a causal and heuristic approach
Pre-prints of the 4th Ifac Workshop on « Online Fault detection and supervision in the chemical process industries. June 7-8, 2001, Jeju Island, Korea.
- [Heim02] HEIM B., GENTIL S., CAUVIN S., TRAVE-MASSUYES L., BRAUNSCHWEIG B.

Fault diagnosis of a chemical process using causal uncertain model.

15th European Conference on Artificial Intelligence, Prestigious Applications of Intelligent Systems (PAIS-2002), Lyon July 24-26 2002.

[Cauvin98] CAUVIN S., CORDIER M-O., DOUSSON C., DEFLANDRE G., LABORIE P., LEVY F., MONTMAIN J., PORCHERON M., SERVET I., TRAVE-MASSUYES L. Monitoring and alarm interpretation in industrial environments. AI Communications 11 (1998) 139-173, IOS Press

IRIT Équipe RPDMP – Raisonnements Plausibles, Décisions et Méthodes de Preuves

UPS – 118 route de Narbonne

31062 Toulouse Cedex 09

Tél. : 05 61 55 63 16

<http://www.irit.fr>

Membres de l'équipe concernés par le thème « Diagnostic et IA » :
Didier Dubois (DR), dubois@irit.fr
Olivier de Mouzon (Doctorant),
mouzon@irit.fr
Henri Prade (DR), prade@irit.fr

Thème général :

L'équipe "Raisonnements Plausibles, Décision, et Méthodes de Preuves" (RPDMP) a une activité importante en matière de modélisation quantitative ou qualitative de l'incertitude ou des préférences, et de formalisation de différentes formes de raisonnement et de décision.

Travaux liés au thème du dossier :

L'équipe RPDMP a un intérêt ancien pour les questions de raisonnement abductif et les problèmes de diagnostic. Son activité récente s'est développée autour de la question du traitement de l'incertitude dans le diagnostic.

Méthodologie :

Une modélisation qualitative de l'incertitude, dans le contexte de la théorie des possibilités, a été développée pour représenter la connaissance, selon une approche principalement causale (effets imprécis plus ou moins certains, valeurs plus ou moins possibles des observa-

tions). L'approche s'applique à des attributs binaires ou numériques.

Des distributions de possibilités restreignent les valeurs (plus ou moins attendues, certaines ou possibles) des attributs quand une panne (unique) est présente. Les observations sont également représentées à l'aide de distributions de possibilités, ce qui permet aussi une prise en compte d'observations imprécises ou incertaines.

Le processus de diagnostic est fondé sur l'exploitation de 2 indices.

Un indice de cohérence (CONS sur fig. 3) permet d'éliminer de la liste de toutes les pannes possibles celles qui sont incohérentes avec une ou plusieurs observations.

Un indice abductif, de pertinence, (PER sur fig. 3) permet d'indiquer les pannes qui sont le plus vraisemblablement en relation directe avec les effets observés.

Ces indices graduels permettent d'ordonner les pannes plus ou moins envisageables, selon les observations effectuées.

Cette approche est développée dans [Dubois et al. 2001], où l'on trouvera tous les détails sur ces indices.

Ces indices peuvent être assouplis afin de réagir, par exemple, lorsque la plupart des symptômes (et pas forcément tous) sont présents. Il est également possible de donner plus ou moins d'importance à certains symptômes, selon les pannes.

Une représentation plus raffinée des connaissances (causales) permet de distinguer entre les valeurs (plus ou moins) certainement non impossibles et un sous-ensemble de valeurs garanties possibles, lorsque la panne est présente. La représentation s'appuie alors sur une double distribution de possibilités.

Enfin, certaines pannes peuvent ne pas avoir d'effet sur des attributs, ou maintenir certains attributs dans des plages de valeurs normales. Un système de diagnostic cherche à trouver une explication aux attributs qui prennent des valeurs anormales. Il est donc intéressant de disposer également d'infor-

mations sur les valeurs des attributs en absence de pannes (également sous forme de distributions de possibilités).

Ces informations peuvent alors aider à détecter de nouvelles pannes ou à rechercher des pannes multiples. Un traitement pratique pour le cas des pannes multiples, ne nécessitant pas d'informations supplémentaires, fait l'hypothèse que les effets de même nature se cumulent et que tout est possible en cas d'effets contraires, sauf information disponible différente.

La plupart de ces raffinements ont été exposés dans [De Mouzon et al. 2001].

Application à la détection de pannes sur bancs d'essais de moteurs automobiles :

L'approche a été mise en œuvre sur une application en ligne pour la détection des pannes lors de la mise au point de calculateurs automobiles sur bancs d'essais moteurs, à Siemens VDO Automotive S.A.S. Le prototype développé se place dans l'hypothèse de la panne unique.

L'automobile est aujourd'hui équipée de calculateurs, afin de contrôler au mieux les paramètres du moteur. Ces calculateurs sont indispensables pour respecter les normes anti-pollution actuelles. Ces normes ont tendance à se renforcer et on essaie aussi de donner un plus grand confort de conduite (démarrage à froid, maintien de la puissance lorsque la climatisation est mise en marche ou arrêtée...) à travers le calculateur. Les calculateurs ont donc des stratégies de plus en plus complexes qui

nécessitent un nombre toujours plus grand de capteurs et de tests pour les mettre au point (pour un type de moteur donné).

Dans ces conditions, il devient important de pouvoir détecter les erreurs de mesures et pannes qui surviennent durant la phase de mise au point, afin de réduire les temps et les coûts de développement, tout en garantissant une meilleure qualité de la mise au point.

La figure 1 décrit le processus de mise au point du calculateur (ECU : Electronic Control Unit). Celui-ci est paramétré par un "outil de calibration". Il possède des capteurs propres sur le moteur. Ces informations sont disponibles dans l'outil d'acquisition des mesures. Des capteurs spécifiques au banc sont également placés sur le moteur pour les besoins de la mise au point.

L'absence de connaissances déjà constituées en ce qui concerne le bon fonctionnement ou les pannes nous a amené à nous intéresser à la seule source de connaissances disponible : celle des techniciens et ingénieurs des bancs moteurs. Afin d'acquérir et de formaliser leurs connaissances nous avons développé un système informatique convivial. La réussite d'un système de diagnostic fondé sur des connaissances humaines dépend en effet pour partie de l'attrait de cet outil et donc de la facilité à l'utiliser. L'outil de formalisation permet ainsi de définir des distributions de possibilités de façon purement qualitative, à l'aide d'une IHM fonctionnelle et agréable.

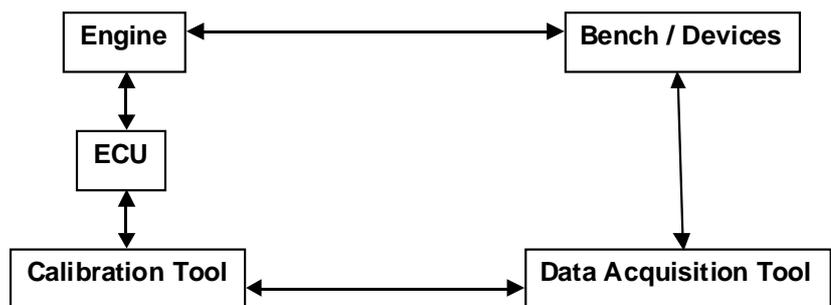


Fig. 1 : schéma de la mise au point d'un calculateur sur banc moteur automobile.

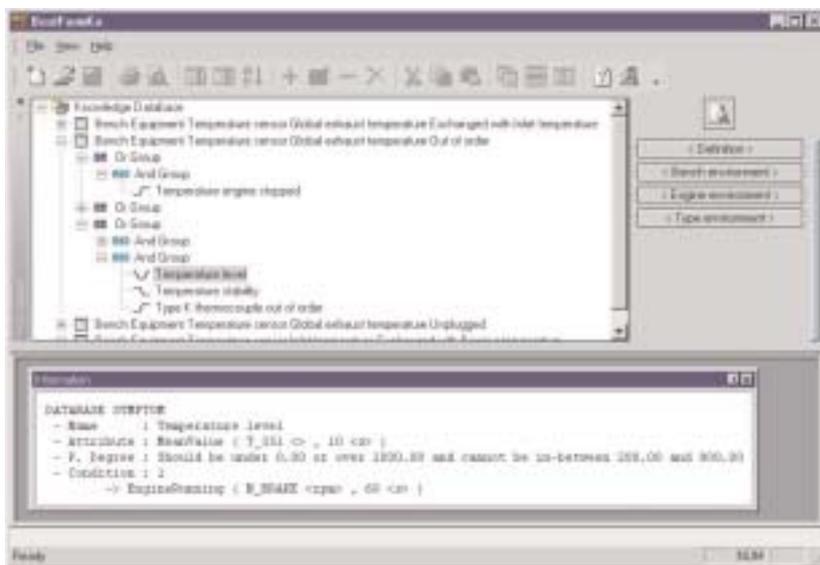


Fig. 2 : écran principal pour l'acquisition des connaissances.

La figure 2 donne un aperçu de cet outil. Les différentes pannes sont listées dans l'ordre alphabétique. On peut avoir accès à leurs symptômes, selon une arborescence à développer (ou refermer). Les symptômes sont de plus organisés en structure et/ou. Une petite icône indique la forme de la distribution de possibilité associée. Enfin, une fenêtre donne les détails pour le symptôme sur lequel on se trouve.

Il est possible d'ajouter ou de modifier des symptômes et des pannes de façon très intuitive.

Cet outil est utilisé hors ligne. Ensuite, le système de diagnostic en ligne peut prendre en compte tout ou partie des connaissances formalisées. Un prototype de ce système de diagnostic a été développé pour expérimentations, tests et validation. Une étude sur un ensemble représentatif de 35 pannes dans 20 différents environnements a mis en évidence l'efficacité du système : les pannes sont détectées en ligne (en moins de 2 minutes, ce qui est largement suffisant pour le problème considéré) lorsqu'elles sont présentes. Très peu de

fausses alarmes ont été observées, qui d'ailleurs s'expliquaient par le manque d'informations disponibles dans certains environnements.

Un exemple simplifié du processus de diagnostic est donné figure 3 (chaque ligne correspond à une panne ; on ne considère ici que la situation par rapport à deux pannes possibles ; les barres verticales indiquent l'évolution de la valeur des indices CONS et PER au cours du temps). Il s'agit de la détection d'une fuite à l'entrée d'air, dans le moteur. La fuite est bien détectée ; les autres pannes sont absentes.

Une synthèse de cette application se trouve dans [Boverie et al. 2002].

Actuellement, l'ensemble des connaissances sur les pannes est en phase de formalisation. Une extension du prototype aux pannes multiples et, éventuellement, en cascade, est également souhaitée.

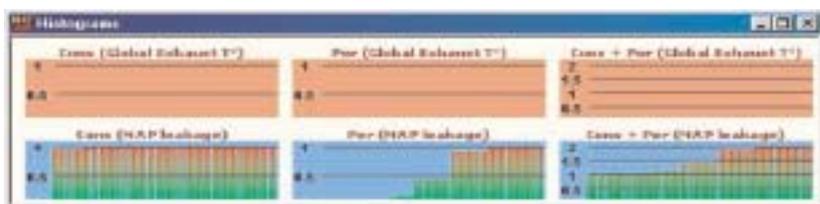


Fig. 3 : détection d'une fuite d'air à l'entrée du moteur.

Bibliographie :

L'ensemble des références (et quelques résumés des articles) peuvent se trouver sur http://dbweb.irit.fr:8902/appli/publi/publis_recc.menu.

[Boverie et al. 2002] Serge Boverie, Didier Dubois, Xavier Guérandel, Olivier De Mouzon, Henri Prade. *Online diagnosis of engine Dyno test benches: a possibilistic approach*. Dans: *Proc. of the 15th European Conference on Artificial Intelligence, ECAI 2002, Lyon - France*. IOS Press, Amsterdam - The Netherlands, p. 658-662, 21-26 juillet 2002.

[De Mouzon à p.] Olivier De Mouzon. *Une approche du diagnostic basée sur la logique floue - Application aux bancs d'essais de moteurs automobiles*. Thèse de doctorat. Université Paul Sabatier, Toulouse, à paraître.

[De Mouzon et al. 2001] Olivier De Mouzon, Didier Dubois, Henri Prade. *Twofold fuzzy sets in single and multiple fault diagnosis, using information about normal values*. Dans: *Proc. of the 10th IEEE International Conference on Fuzzy Systems, Melbourne - Australia*. FUZZ-IEEE'01 Programm Committee, Melbourne, Vol.3, 2-5 décembre 2001.

[Dubois et al. 2001] Didier Dubois, Michel Grabisch, Olivier De Mouzon, Henri Prade. *A possibilistic framework for single-fault causal diagnosis under uncertainty*. *International Journal of General Systems*. V. 30 N. 2, p. 167-192, 2001.

LAAS-CNRS / Groupe DISCO Diagnostic, Supervision et Conduite Qualitatifs

7 Avenue du Colonel Roche
31077 Toulouse cedex 4
Tel. 05 61 33 62 00
<http://www.laas.fr>

Responsable :

José Aguilar-Martin (DR)

Chercheurs permanents concernés par le thème « Diagnostic et IA » :

Louise Travé-Massuyès (DR),
louise@laas.fr

José Aguilar-Martin (DR),
aguilar@laas.fr

Marie-Véronique Lelann (Prof),
mvlelann@laas.fr

Andrei Doncescu (MC),
adoncesc@laas.fr

Thème général :

La thématique du groupe DISCO se situe dans le cadre de la sûreté de fonctionnement et de la maîtrise des systèmes complexes. Ses activités couvrent les domaines du diagnostic, de la supervision et de la conduite qui sont abordés par trois types d'approches : à base de modèles, à base de connaissances et par apprentissage et reconnaissance des formes. Les techniques utilisées s'inscrivent dans les domaines de l'Automatique et de l'Intelligence Artificielle (IA).

Travaux liés au thème du dossier :

Diagnostic et reconfiguration de systèmes concurrents hybrides

Les systèmes hybrides se définissent par une double dynamique, due à la coexistence des phénomènes continus et à événements discrets. Ce thème s'intéresse au diagnostic et à la reconfiguration de tels systèmes, dans une approche à base de modèles, avec une application aux satellites autonomes. Ce travail s'effectue dans le cadre d'un projet en collaboration avec le LIPN ainsi que le CNES et ASTRIUM.

Le formalisme KOALA (fig. 1), intègre des automates concurrents représentant les différents modes opératoires des composants (e.g. vanne ouverte ou fermée) avec des modèles continus causaux incertains [Benazera et al. 2001]. Une interface continu/discret permet de gérer les changements de configuration

du système. Pour cela, les différentes régions de l'espace d'état ont une représentation logique, ce qui permet une gestion par des techniques de satisfaction de contraintes booléennes. La détection de fautes s'effectue au niveau continu et l'identification de l'état du système est gérée par le niveau discret (logique) par une recherche de la (les) configuration(s) cohérente(s) avec les observations courantes. Suivant l'état identifié et les fonctions à restaurer, un état cible est déterminé et un module de planification à base de modèles produit un plan de manière réactive.

Détermination des états fonctionnels d'un système dynamique

Ce thème s'intéresse à l'abstraction du comportement d'un système dynamique continu, donné par les différents signaux disponibles, en un nombre fini d'états fonctionnels à l'aide d'un classificateur (LAMDA). A partir d'informations quantitatives et/ou qualitatives, ce classificateur permet, de façon récurrente et en temps réel, de déterminer dans quel état fonctionnel se situe le système. L'ensemble fini des états, qui est confirmé par un expert de manière interactive, constitue un automate. Ce modèle sert de base à la surveillance du système (Fig. 2) [Aguilar Martin et al. 2001]. Des applications de cette approche se sont faites sur des machines outil et des processus biotechnologiques. Elle est actuellement appliquée à un procédé pétro-chimique dans le cadre du projet Européen CHEM, coordonné par l'IFP.

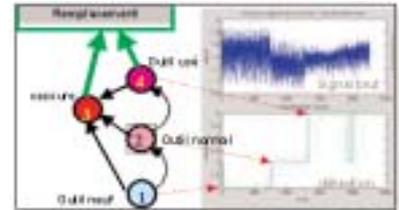


Figure 2 : Exemple de surveillance d'une machine outil

Modèles causaux incertains pour la détection de fautes et le diagnostic

Une ligne importante de notre activité porte sur la détection de fautes et le diagnostic fondé sur des modèles causaux incertains. Ces travaux sont à la base de notre système Ca-En qui fut appliqué dans les projets Européens TIGER et TIGER SHEBA dans le domaine des turbines à gaz. Ils font l'objet actuellement d'une collaboration avec le LAG et l'IFP. Les prédictions effectuées grâce à ces modèles se présentent sous forme d'enveloppes (Fig 3), synthétisant tout un ensemble de trajectoires, et fournissent des seuils adaptatifs au-delà desquels une alarme est déclenchée [Travé-Massuyès et al. 2001a]. Pour de meilleures performances, on peut utiliser des *enveloppes à erreur bornée*, constituées d'une paire d'enveloppes, l'une surdimensionnée et l'autre sous-dimensionnée (collaboration avec l'Université de Gironne) [Armengol et al. 2001b].

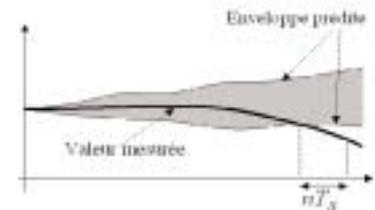


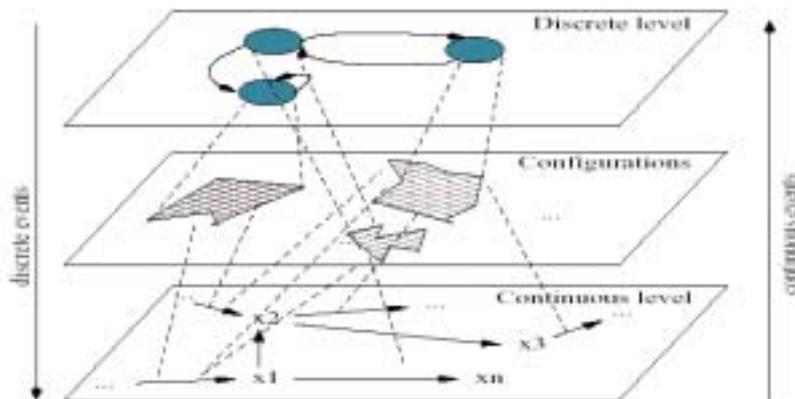
Figure 3 : Exemple de prédiction

La localisation des fautes s'inscrit dans l'approche de Reiter et exploite la nature causale des modèles qui lui permet, à partir des variables en alarme, de retrouver les chaînes de dépendances correspondant à des composants en conflit [Travé-Massuyès et al., 2001a].

Diagnosticabilité et placement de capteurs

Ce thème s'intéresse à l'analyse de la

Figure 1 : Echanges entre niveau discret et niveau continu par une interface



diagnosabilité d'un système et à la détermination de l'instrumentation nécessaire pour atteindre un certain degré de diagnosabilité. Notre approche [Travé-Massuyès et al. 2001] se base sur un modèle du comportement nominal du système dont l'interprétation causale permet de déduire les redondances analytiques correspondant à différentes hypothèses de disponibilité de capteurs. Nous travaillons actuellement à une généralisation de la méthode exploitant une recherche exhaustive des différentes interprétations causales des relations du modèle. Ce travail réalisé en collaboration avec l'UPC (Espagne) a été appliqué à plusieurs sous-systèmes d'une turbine à gaz Frame 6 General Electric de National Power (UK) dans le cadre du projet Européen TIGER SHEBA.

Génération d'arbres de diagnostic à base de modèles incertains

Dans le domaine du diagnostic hors ligne, comme celui effectué dans les garages dans le domaine de l'automobile, le problème essentiel est celui de la détermination de la séquence de tests et mesures aux points de contrôle (cf. Figure 4), qui permette de localiser la faute le plus rapidement possible et avec un coût minimal.

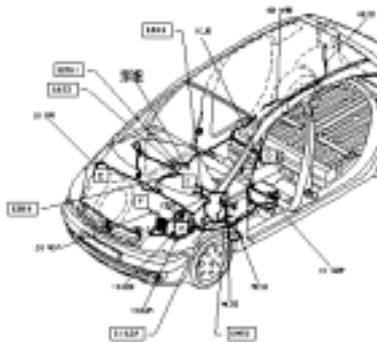


Figure 4 : Points de contrôle

L'outil utilisé est l'arbre de diagnostic et le problème posé est donc celui de la génération automatique d'arbres de diagnostic optimaux. Ce travail, en collaboration avec la société ACTIA, a abouti à la réalisation du prototype AGENDA [Faure et al. 2001].

La procédure d'AGENDA est en 2 étapes :

- 1) Prédiction des valeurs des tests dans chaque cas de faute anticipé, à partir

des modèles incertains correspondants par une technique d'optimisation globale.

- 2) Génération de l'arbre de diagnostic par un algorithme de recherche heuristique du type meilleur d'abord AO*.

Formalismes de raisonnement qualitatif

Les techniques de raisonnement qualitatif constituent une ligne de recherche de longue date, avec des contributions reconnues concernant les algèbres qualitatives et les modèles aux ordres de grandeur, la simulation causale qualitative et les techniques d'ordonnement causal [Travé-Massuyès et al. 1997].

Plus récemment, des travaux théoriques visant à fournir une passerelle entre les modèles aux ordres de grandeur absolus d'une part et relatifs d'autre part ont été finalisés. Ils établissent les conditions sous lesquelles un modèle aux ordres de grandeur absolus est cohérent avec un modèle aux ordres de grandeur relatifs [Travé-Massuyès et al., 2002].

Par ailleurs, un travail de synthèse sur la causalité dans le domaine du raisonnement qualitatif a été réalisé dernièrement en collaboration avec P. Dague du LIPN, Paris Nord [Dague & Travé-Massuyès, 2002].

Intégration de Méthodes ALiant Automatique et IA (IMALAIA)

Dans le cadre du groupe IMALAIA, une de nos activités a consisté à faire l'inventaire des modèles et méthodes pour la diagnostic à base de modèles proposés en Automatique et en IA et leurs motivations, ainsi qu'à effectuer des travaux plus techniques, tels qu'une analyse comparative de l'approche dite des résidus structurés et de la théorie logique du diagnostic, qui a donné lieu à trois publications : à l'ECAI'00 [Cordier et al. 2000], à Safeprocess'00 et à DX'00. Cette action envisage actuellement d'étendre le cadre commun développé en se focalisant sur les aspects temporels.

Activités d'animation dans le domaine

Louise Travé-Massuyès co-anime le Groupe IMALAIA rattaché aux GdRs

I3 et Automatique et labelisé « groupe AFIA » (<http://www.afia.polytechnique.fr/node.php?lang=fr&node=7>). Ce groupe constitue un projet du groupement S3 « Sureté, Surveillance, Supervision » (<http://www.univ-lille1.fr/s3/fr/fenetre.htm>).

Références

L'ensemble des articles et références peut se trouver sur <http://www.laas.fr/laasvif/index.htm>.

- [Aguilar Martin et al. 2001] J. Aguilar Martin, L. Haenlein, R. Sarrate-Estruch, J. Waissman-Vilanova, Knowledge-based signal analysis and case-based condition monitoring of a machine tool, *9th IFSA World Congress and 20th NAFIPS International Conference*, Vancouver (Canada), 25-28 Juillet 2001, pp.286-291.
- [Armengol et al. 2001] J. Armengol, J. Vehi, L.Travé-Massuyès, M.A. Sainz, Application of modal intervals to the generation of error-bounded envelopes, *Reliable Computing*, Vol.7, N° 2, pp.171-185, 2001.
- [Benazera et al. 2001] E. Benazera, L. Travé-Massuyès, P. Dague, State Tracking of Uncertain Hybrid Concurrent Systems, *DX Workshop on « Principles of Diagnosis»*, Semmering, Austria, 2-4 May 2002, pp. 106-114.
- [Cordier et al. 2000b] Cordier M.-O., Dague P., Dumas M., Levy F., Montmain J., Staroswiecki M., Travé-Massuyès L., A comparative analysis of AI and control theory approaches to model-based diagnosis, *ECAI '00*, Berlin, Germany, August 20-25, 2000.
- [Faure et al. 2001] P.P. Faure, X. Olive, L. Travé-Massuyès, H. Poulard, AGENDA: Automatic GENERation of DiAgnosis trees, *Journées Doctorales d'Automatique (JDA'2001)*, Toulouse (France), 25-27 Septembre 2001, pp.165-170.
- [Travé-Massuyès et al. 1997] L. Travé-Massuyès, P. Dague, F. Guerrin (Eds), *Le Raisonnement Qualitatif pour les Sciences de l'Ingénieur*, Editions Hermès, Paris, 1997.
- [Travé-Massuyès et al. 2001b] L. Travé-Massuyès, T. Escobet, R. Milne, Model-based diagnosability and sensor placement. Application to a frame 6 Gas turbine subsystem, *7th International Joint Conference on Artificial Intelligence (IJCAI'01)*, Seattle (USA), 4-10 Août 2001, pp.551-556.
- [Travé-Massuyès et al. 2001a] L. Travé-Massuyès, T. Escobet, R. Pons, S. Tornil, The Ca-En diagnosis system and its automatic modelling method, *Computación i Sistemas Journal*, Vol 5, n°2, pp. 128-143, 2001.
- [Travé-Massuyès et al. 2002] Travé-Massuyès L., F. Prats, M. Sanchez, N. Agell, Consistent relative and absolute order-of-

magnitude models, In Working Notes of the 16th Int. Workshop on Qualitative Reasoning QR'02, Sitges, Spain, 10-12 June 2002, pp. 185-192.

[Dague & Travé-Massuyès 2002] Dague P., Travé-Massuyès L., Raisonement causal en physique qualitative (in french), submitted to the Special Issue on Causality of *Intellectica*.

**Laboratoire d'Automatique
Mécanique et d'Informatique
industrielles et Humaines.
Université de Valenciennes**

Correspondant :
sylvain.piechowiak@univ-valen-
ciennes.fr

M.A. Maalej (doctorant) V. Delcroix et
S. Piechowiak (LAMIH – CNRS UMR
8530)

J. Rodriguez (INRETS-ESTAS)

Diagnostic à base de modèles

Le LAMIH (Laboratoire d'Automatique, de Mécanique et d'Informatique industrielles et Humaines) étudie le problème du diagnostic par des approches issues de l'IA depuis de nombreuses années.

Initialement, en 1986, l'approche « système expert » a été suivie et a débouché sur SEDIAG, un générateur de systèmes experts. SEDIAG était original du fait que différents types de connais-

sances pouvaient être exploitées (connaissances de surface et profondes). A la suite de ces premiers travaux, des études liées à « l'extraction des connaissances ont été menées. De même, les problèmes liés au diagnostic « homme machine » ont également été entrepris [2].

En 1992, c'est l'approche à base de connaissances profondes (dans la lignée des approches « model-based ») qui a été étudiée. L'intégration d'une certaine forme de prise en compte du temps permet de traiter des circuits séquentiels évoluant dans le temps [6].

Depuis quelques années, un rapprochement est fait avec la communauté scientifique qui s'intéresse aux « contraintes ». En effet, la représentation des modèles par des réseaux de contraintes permet de traiter le diagnostic comme un véritable problème de maxCSP. En parallèle, une réflexion est menée concernant l'élaboration des modèles ; en particulier, on essaie de définir des méthodes de recherche des erreurs de modélisations [4]. Un premier travail a été fait en appliquant les approches à base de modèles sur des dispositifs complexes réels comme, par exemple, le système de gestion du trafic ferroviaire [3].

Depuis environ 3 ans, un élargissement des investigations sur le problème du diagnostic est réalisé. Nous nous intéressons aux approches bayésiennes

pour lesquelles nous envisageons des approches hiérarchisées [1]. L'objectif visé est de hiérarchiser les modèles selon différents niveaux de finesse puis de rechercher les diagnostic de manière « focalisante ».

Quelques références bibliographiques

- [1] V. Delcroix, S. Piechowiak, J. Rodriguez, Computing diagnoses with higher posterior probability using bayesian networks, soumis à *International Conference on Information Processing and Management of Uncertainty in Knowledge-based systems (IMPU'2002)*, Annecy, july 1-5, 2002
- [2] D. Jouglet, *Coopération homme-machine pour le diagnostic technique : Application aux dérangements téléphoniques*, Thèse de Doctorat, UVHC, 26 janvier 2000.
- [3] F. Lefebvre, *Contribution à la modélisation pour le diagnostic des systèmes complexes: Application à la signalisation des lignes à grande vitesse*, Thèse de doctorat, UVHC, 17 juin 2000.
- [4] S. Piechowiak, *Raisonnement à base de contraintes: Application au diagnostic*, HDR, UVHC, 19 Décembre 2001.
- [5] S. Piechowiak, J. Rodriguez, A constraint network manager for diagnosing, *Proceedings seventh International Conference on Industrial & Applications of Artificial Intelligence & Expert Systems*, Austin, Texas, U.S.A., Juin 1994.
- [6] J. Rodriguez, S. Piechowiak, Expressing the behaviour of time-varying devices for the model-based diagnosis, *Proceedings of the 14th International Workshop on Expert Systems and their Applications*, Paris, France, 1994.

Note de Jean-Baptiste Berthelin
sur le *Gédéon* de Jean-Gabriel Ganascia (été 2002).

« Une synecdoque est-elle vraiment une anecdote ? », se demandait gravement l'un des boulets que l'illustre savant allait précipiter de la Tour de Pise. Et si une pomme tombait de la Lune ?

Les expériences de pensée remontent à la plus haute Antiquité. En Chine par exemple, un armurier du pays de Chu vendait à la fois des lances dont la pointe pouvait tout transpercer, et des boucliers capables de tout arrêter. Un client l'interrogea respectueusement : « Si je menace un adversaire avec une de tes lances et qu'il se protège avec un de tes boucliers, qui va gagner ? »

Mais c'est là une petite expérience, dont l'enjeu n'est que l'impudence de l'armurier. Il en est de plus grandioses, conçues par les physiciens, mathématiciens et informaticiens, qui concernent la Matière, le Calcul et la Pensée. Jean-Gabriel Ganascia, que nous connaissions comme homme de science et homme de savoir, devient ici homme de lettres, et lance un candide narrateur dans une aventure délirante ; son guide, le fameux Gédéon, ne cesse de mettre la réalité à l'épreuve de sa perverse curiosité.

Le narrateur se laisse conduire dans cet univers improbable. Il pense en profiter pour retrouver des diamants verts que des Chinois ont perdus, sans parler d'une pauvre princesse qui se trouve (ou ne se trouve pas) dans

une chambre qui communique avec l'extérieur par le canal d'idéogrammes tracés sur divers supports. Il sauve un chat de la mort par le cyanure, il visite des planètes presque jumelles de la nôtre, il apprend le « vleur » et le « blert ».

Toutes ces péripéties le laissent, à la fin, dans un état de confusion mentale certaine. Le lecteur, heureusement, peut avancer un peu plus loin, dans la partie pédagogique de l'ouvrage, où nous retrouvons notre Jean-Gabriel Ganascia professeur des Universités. Patiemment, il nous redit le rôle crucial de l'imaginaire dans la découverte scientifique, aussi bien que dans d'autres aventures de l'esprit.

Après plusieurs traités de facture plus académique, cet ouvrage apporte une note de fantaisie dans un domaine qui, à mon avis, ne peut qu'en devenir plus attractif. Si de très jeunes lecteurs tombent sous son charme, cela contribuera à en orienter certains vers les métiers de la recherche, ce qui est on ne peut plus souhaitable.

Gédéon, ou les aventures extravagantes
d'un expérimentateur en chambre, par
Jean-Gabriel Ganascia, éditions Le Pommier,
juin 2002, ISBN 2-476-50034-5

Compte rendu de la conférence

CAP 2002

Conférence Francophone d'Apprentissage

Orléans les 17 -19 Juin 2002

<http://www.univ-orleans.fr/CONFERENCES/CAP2002/>

*Présidente du comité scientifique : **Christel Vrain** LIFO, Université d'Orléans*

*Président du comité d'organisation : **Matthieu Exbrayat** LIFO, Université d'Orléans*

La Conférence d'Apprentissage (CAp) est le 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 dans le domaine de l'apprentissage automatique qui 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 connais-

sances à partir de données) et de la fouille de textes.

L'édition 2002 a été organisée à Orléans, du 17 au 19 juin 2002, par le LIFO (Laboratoire d'Informatique Fondamentale d'Orléans).

Une vingtaine de conférences, une session de démonstrations et posters, une session "industrielle" et trois conférences invitées ont été accueillies cette année.

Les transparents de la plupart des conférences invitées et des sessions industrielles sont désormais téléchargeables à l'adresse <http://www.univ-orleans.fr/CONFERENCES/CAP2002/>

Calendrier des conférences AFIA

<i>Octobre 2002</i>			
21 - 22	LFA	Montpellier	Rencontres Francophones sur la Logique Floue et ses Applications http://www.lirmm.fr/lfa02
28 - 31	JFIADSMA	Lille	Journées Francophones d'Intelligence Artificielle Distribuée et Systèmes Multi-Agents http://www.lifl.fr/jfiadsma2002
<i>Janvier 2003</i>			
22 - 24	EGC	Lyon	Journées francophones d'Extraction et de gestion des Connaissances http://www.univ-lyon3.fr/EGC2003
<i>Juin 2003</i>			
	Plate-forme AFIA		Plate-Forme AFIA

Programme de la conférence

LFA02

Rencontres Francophones sur la Logique Floue et ses Applications

Orléans les 21-22 Octobre 2002

<http://www.lirmm.fr/lfa02/>

Présidente du comité scientifique : **Sylvie Galichet** LAMII, Annecy
Président du comité d'organisation : **O. Strauss** LIRMM, Montpellier

Les rencontres sur la logique floue et ses applications sont l'occasion pour les chercheurs de la communauté francophone de se réunir afin de faire le point sur les développements actuels de la discipline. Depuis plusieurs années, la manifestation s'est élargie, au delà des ensembles flous et de la théorie des possibilités, à d'autres théories de l'incertain. L'édition 2002 de la manifestation s'inscrit dans cette volonté d'évolution et d'ouverture.

Sous l'appellation de « théorie de l'incertain », sont ainsi considérées la théorie des possibilités et les ensembles flous, la théorie des fonctions de croyance, les probabilités imprécises, les ensembles approximatifs ("rough sets"), le raisonnement qualitatif...

Un nombre croissant de domaines est concerné par l'exploitation de ces différentes théories. Si d'un point de vue historique la commande floue a joué un rôle précurseur, elle est aujourd'hui rejointe par bon nombre de disciplines telles que la recherche opérationnelle et l'aide à

la décision, le mesurage et l'évaluation subjective, l'analyse de données, la classification et la reconnaissance de formes, le traitement et l'interprétation des images, les bases de données et les systèmes d'informations, l'intelligence artificielle, la fusion de d'informations, la robotique, mais aussi l'économie, la gestion, la psychologie, l'ergonomie...

La conférence aura lieu cette année au palais des congrès de Montpellier et se déroulera avec 2 sessions en parallèle. Une trentaine de présentations sont attendues ainsi que quelques présentations de logiciels. Deux invités sont également attendus : Thomas A. Sudkamp qui donnera une conférence intitulée « Similarity and the measurement of possibility » et Ludmilla I. KUNCHEVA qui interviendra sur « 'Fuzzy' vs 'Non-fuzzy' in combining classifiers: an experimental study ».

Le programme est consultable à l'adresse <http://www.lirmm.fr/lfa02/>.

Programme de la conférence

JF IAD SMA 2002

Conférence Francophone d'Intelligence Artificielle Distribuée et Systèmes Multi-Agents

Lille les 28 - 30 octobre 2002

<http://www.lifl.fr/jfiadsma2002>

Président du comité scientifique : **Jean-Pierre Müller**, CIRAD, Montpellier
Président du comité d'organisation : **Philippe Mathieu**, LIFL, Lille

La dixième édition des Journées Francophones en Intelligence Artificielle Distribuée et Systèmes Multi-Agents (JFIADSMA) est organisée par le LIFL et aura lieu à Lille. Ces journées constituent un forum annuel d'échange privilégié entre chercheurs provenant de plusieurs communautés scientifiques telles que l'intelligence artificielle, la vie artificielle, les sciences humaines, les systèmes distribués ou encore le génie logiciel.

Le thème principal de cette année est « systèmes multi-agents et systèmes complexes ». En effet, la recherche dans le domaine des systèmes multi-agents (SMA) revêt de plus en plus d'importance pour leur capacité à aborder les systèmes complexes, c'est-à-dire les systèmes constitués de nombreux composants en interaction dynamique entre eux et avec le monde extérieur. Cette problématique touche aussi bien l'ingénierie logi-

cielle (programmation orientée agent/interaction/organisation), la résolution de problèmes (gestion de réseaux, pilotage d'entreprises virtuelles, marchés électroniques) ou la simulation (entreprises, écosystèmes, sociosystèmes).

Plus d'une vingtaine de conférences sont attendues. Associées à ces présentations scientifiques et techniques, des démonstrations d'outils logiciels compléteront le programme (consultable à l'adresse <http://www.lifl.fr/jfiad-sma2002>).

David Hill (ISIMA / LIMOS, UMR CNRS 6158) sera un des deux conférenciers invités et interviendra sur le thème des « Nouveaux défis pour les simulations et les systèmes multi-agents ». Jean Bézivin (CRGNA, Université de Nantes) sera le deuxième invité et donnera une conférence intitulée: « Nouveaux défis des systèmes complexes et la réponse MDA de l'OMG ».

Présentation de la conférence

EGC 2003

Journées francophones d'Extraction et de Gestion des Connaissances

Lyon les 22 - 23 - 24 Janvier 2003

<http://www.univ-lyon3.fr/EGC2003>

Co-présidents du comité de programme (Académique)

Mohand-Saïd Hacid, LISI, Université Lyon 1

Yves Kodratoff, LRI, Université Paris Sud

Présidente du comité de programme (Entreprise)

Eunika Mercier-Laurent, EML Conseil, Montpellier

Présidente du comité d'organisation

Danielle Boulanger, MODEME, IAE, Université Jean Moulin, Lyon

Dans le prolongement des deux manifestations précédentes, EGC 2003 ambitionne de regrouper des chercheurs, industriels et utilisateurs francophones issus des communautés Bases de Données, Apprentissage, Représentation des Connaissances et Statistiques.

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 de données, systèmes d'information et bases de données, ingénierie des connaissances, etc.), et d'autre part les spécialistes d'entreprises qui déploient des méthodes d'EGC, afin de contribuer à la formation d'une communauté scienti-

fique dans le monde francophone autour de cette double thématique de l'extraction et de la gestion de connaissances.

Cette manifestation est ainsi un lieu privilégié d'échanges entre chercheurs et spécialistes industriels, afin de présenter l'état de l'art de cette thématique, d'en faire émerger les besoins et les enjeux, ainsi que d'en stimuler le développement.

Principaux thèmes (liste non limitative) :

- Modèles de représentation des données et des connaissances.
- Méthodes de prétraitement. Traitement de données volumineuses.
- Méthodes et algorithmes d'extraction de connaissances à partir de données hétérogènes (bases de données, textes, images, vidéos, audio, données spatiales, données temporelles, données semi-structurées, Web).
- Algorithmes d'apprentissage et analyse de données, réseaux neuromimétiques, réseaux bayésiens, programmation logique inductive, ensembles approximatifs, découverte de règles...
- Mesure de la qualité des connaissances.
- Visualisation des connaissances.

- Entrepôts de données.
- Web Sémantique.
- Business Intelligence.
- Mémoires d'entreprises.
- Synthèse de connaissances. Ontologies et méta-informations.
- Méthodes de recueil, d'acquisition des connaissances et de gestion du cycle de vie des connaissances.
- Déploiement des connaissances (travail coopératif, workflow), réseaux de diffusion et de partage.
- Systèmes d'information et de gestion des connaissances.
- Applications : tout domaine d'application quel qu'en

soit le secteur ou l'activité professionnelle est éligible : industrie, administration, défense, biomédical, télécommunications, gestion de la clientèle, gestion des ressources humaines, production et maintenance...

À noter enfin, les dates limites de l'appel à communication :

- 1 octobre 2002: pour la réception des résumés des communications académiques
- 14 octobre 2002 pour la réception des articles pour les communications académiques et d'entreprises, les tutoriels et les ateliers

Présentation de la Plate-Forme AFIA

Plate-Forme AFIA 2003

Laval - juin/juillet 2003 (dates à préciser)

Présidente du comité d'organisation : Anne-Marie Kempf ESIEA Recherche, Laval
Responsable AFIA : Gilles Bisson laboratoire Leibniz-IMAG, Grenoble

Organisée tous les deux ans, en alternance avec la conférence RFIA, la Plate-forme de l'AFIA rassemble sur un même site plusieurs conférences spécifiques du domaine de l'Intelligence Artificielle. Cette manifestation est l'occasion de réunir des chercheurs ayant des thématiques voisines et complémentaires, avec des industriels.

Ouverte aux conférences émanant des groupes de travail ou des collègues de l'AFIA, la plate-forme a également pour vocation d'accueillir plus largement l'ensemble des communautés qui sont scientifiquement proches de l'IA. L'IA est en effet une discipline « ouverte » qui, à côté de communautés scientifiques clairement identifiées comme faisant partie de ce domaine de recherche, est en collaboration avec d'autres communautés pour lesquelles l'IA n'est pas le domaine scientifique principal. Une manifestation comme la plate-forme AFIA est donc une occasion de dialogues et d'échanges avec ces groupes de chercheurs situés à la périphérie de l'IA.

En juin 2001 cette manifestation (<http://www-leib->

[niz.imag.fr/AFIA/](http://www-leib-niz.imag.fr/AFIA/)) s'est déroulée sur 4 jours à Grenoble, organisée par le laboratoire Leibniz-IMAG. Elle a rassemblé 215 chercheurs qui se sont répartis en 3 conférences : (CAP : Conférence apprentissage ; IC : Ingénierie des connaissances ; A3CTE : Application, apprentissage et acquisition de connaissances à partir de textes électroniques) et en 4 ateliers : (RàPC : Raisonnement à partir de cas ; IA-KM : Intelligence artificielle et knowledge management ; SMA : Méthodologie et environnements pour les systèmes multi-agents ; EIAH : Environnement informatique d'apprentissage humain).

L'édition 2003 de cette manifestation se déroulera à Laval fin juin / début juillet (dates à préciser) et sera organisée par l'ESIEA Recherche. Les conférences CAP (Conférence Apprentissage), IC (Ingénierie des Connaissances) et RJCIA (Rencontres Jeunes Chercheurs en Intelligence Artificielle) y participeront et quelques ateliers, comme en 2001, devraient rejoindre la plate-forme très prochainement.

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://ria.e-revues.com/>

RIA V16 N°1-2 2002

- Editorial : HOMMAGE A HERBERT SIMON (1916-2001), *J.PITRAT*
- Herbert Simon, pionnier de l'Intelligence Artificielle, *J.PITRAT*
- The way it was at Carnegie: 1965-1968, *L.SIKLÓSSY*
- Travailler avec Herbert Simon, *F.GOBET*
- Sur le colloque Sciences de l'Intelligence, Sciences de l'Artificiel, Extraits des commentaires et des réponses aux questions, *H.SIMON*
- GPS peut toujours réussir... *L.SIKLÓSSY*
- Sciences de l'artificiel, modélisation et rationalité, *A.NICOLLE*
- Herbert Simon et le méta, *J.PITRAT*
- Aperçus sur la « découverte scientifique », *J.GANASCIA*
- Utilisation de dessins en résolution de problèmes, *D.PASTRE*
- Recherche et reconnaissance de patterns, *F.GOBET*
- Les composantes sémantiques dans la résolution de problèmes isomorphes, *J.RICHARD, E.CLÉMENT, C.TIJUS*
- L'apport de Herbert Simon dans le management et la décision, *J.POMEROL*
- Sur un exceptionnel manifeste épistémologique : "Symbols and Search" Merveilleux et pourtant compréhensible est le traitement heuristique des systèmes de symboles physiques, *J.LE MOIGNE*

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/inca/publications/store/5/0/5/6/0/1/>

ARTIFICIAL INTELLIGENCE V139 N°1 JULY 2002

- Yet some more complexity results for default logic, *R. BEN-ELIYAHU-ZOHARY*

- Local search with constraint propagation and conflict-based heuristics, *N. JUSSIEN, O. LHOMME*
- Clausal resolution in a logic of rational agency, *C. DIXON, M. FISHER, A. BOLOTOV*
- Summarization beyond sentence extraction: A probabilistic approach to sentence compression, *K. KNIGHT, D. MARCU*
- Accelerating filtering techniques for numeric CSPs, *Y. LEBBAH, O. LHOMME*

ARTIFICIAL INTELLIGENCE V139 N°2 AUGUST 2002

- Learning cost-sensitive active classifiers, *R. GREINER, A. J. GROVE, D. ROTH*
- Abstracting soft constraints: Framework, properties, examples, *S. BISTARELLI, P. CODOGNET, F. ROSSI*
- Narrative prose generation, *C. B. CALLAWAY, J. C. LESTER*
- Lattice-theoretic models of conjectures, hypotheses and consequences, *M. YING, H. WANG*

ARTIFICIAL INTELLIGENCE V140 N°1-2 SEPTEMBER 2002

- Binary vs. non-binary constraints, *F. BACCHUS, X. CHEN, P. VAN BEEK, T. WALSH*
- On point-duration networks for temporal reasoning, *I. NAVARRETE, A. SATTAR, R. WETPRASIT, R. MARIN*
- Nagging: A scalable fault-tolerant paradigm for distributed search, *A. M. SEGRE, S. FORMAN, G. RESTA, A. WILDENBERG*
- Solving multi-granularity temporal constraint networks, *C. BETTINI, X. S. WANG, S. JAJODIA*
- Disjunctions, independence, refinements, *M. BROXVALL, P. JONSSON, J. RENZ*
- Abduction in logic programming: A new definition and an abductive procedure based on rewriting, *F. LIN, J-H. YOU*
- Context-specific sign-propagation in qualitative probabilistic networks, *S. RENOUIJ, L. C. VAN DER GAAG, S. PARSONS*
- Tractable combinatorial auctions and b-matching, *M. TENNENHOLTZ*
- Erratum to: Drift analysis and average time complexity of evolutionary algorithms, *J. HE, X. YAO*

ARTIFICIAL INTELLIGENCE V141 N°1-2 OCTOBER 2002

- Explanations, belief revision and defeasible reasoning, *M.A. FALAPPA, G. KERN-ISBERNER, G. R. SIMARI*
- Using genetic programming to learn and improve control knowledge, *R. ALER, D. BORRAJO, P. ISASI*
- On the computational complexity of assumption-based argumentation for default reasoning, *Y. DIMOPOULOS, B. NEBEL, F. TONI*

SOMMAIRES DES REVUES

- Conditional independence in propositional logic, *J. LANG, P. LIBERATORE, P. MARQUIS*
- Unifying metric approach to the triple parity, *T. Y.T. CHAN*
- The size distribution for Markov equivalence classes of acyclic digraph models, *S. B. GILLISPIE, M. D. PERLMAN*
- The consensus operator for combining beliefs, *A. JØSANG*
- Emergence of social conventions in complex networks, *J. DELGADO*
- Coherence in finite argument systems, *P. E. DUNNE, T.J.M. BENCH-CAPON*
- On forward checking for non-binary constraint satisfaction, *C. BESSIÈRE, P. MESEGUER, E. C. FREUDER, J. LARROSA*

AI MAGAZINE

<http://www.aaai.org/Magazine/magazine.html>

AI MAGAZINE V22 N°4 2001

- Introduction to the Special Issue on Intelligent User Interfaces, *J. LESTER*
- COLLAGEN: Applying Collaborative Discourse Theory to Human-Computer Interaction, *C. RICH, C. L. SIDNER, N. LESH*
- Toward Conversational Human-Computer Interaction, *J. F. ALLEN, D. K. BYRON, M. DZIKOVSKA, G. FERGUSON, L. GALESCU, A. STENT*
- Intelligent Tutoring Systems with Conversational Dialogue, *A. C. GRAESSER, K. VANLEHN, C. P. ROSE, P. W. JORDAN, D. HARTER*
- Controlling the Behavior of Animated Presentation Agents in the Interface: Scripting versus Instructing, *E. ANDRE, T. RIST*
- Embodied Conversational Agents: Representation and Intelligence in User Interfaces, *J. CASSELL*
- Pedagogical Agent Research at CARTE, *W. L. JOHNSON*
- Interface Agents in Model World Environments, *R. ST. AMANT, R. M. YOUNG*
- Agent-Centered Search, *SVEN KOENIG*
- On the Origin of Environments by Means of Natural Selection, *M. SIPPER*
- It Does So: Review of The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology, *E. DIETRICH*
- Machine Learning and Light Relief: A Review of Truth from Trash, *J. H. MICHIE*
- Reasoning about Rational Agents: A Review, *S. ZILBERSTEIN*

AI MAGAZINE V23 N°1 2002

- Ten Years of the AAAI Mobile Robot Competition and Exhibition, *T. BALCH, H. YANCO*
- The AAAI-2001 Mobile Robot Exhibition, *B. P. ADAMS, V. VERMA*
- The Hors d'Oeuvres Event at the AAAI-2001 Mobile Robot Competition, *F. MICHAUD, D. A. GUSTAFSON*
- AAAI/RoboCup-2001 Urban Search and Rescue Events, *R. MURPHY, J. BLITCH, J. CASPER*
- AAAI/RoboCup-2001 Robot Rescue, *J. CASPER, H. YANCO*
- AAAI Hosts the National Botball Tournament!, *C. STEIN, D. SCHEIN, D. MILLER*
- RoboCup-2001: The Fifth Robotic Soccer World Championships, *M. VELOSO, T. BALCH, P. STONE, H. KITANO, F. YAMASAKI, K. ENDO, M. ASADA, M. JAMZAD, B. S. SADJAD, V. S. MIRROKNI, M. KAZEMI, H. CHITSAZ, A. HEYDARNOORI, M. T. HAJIAGHAI, E. CHINIFOROSHAN*
- Case-Based Reasoning Integrations, *C. MARLING, M. SQALLI, E. RISSLAND, H. MUNOZ-AVILA,, D. AHA*
- AI Topics, *B. G. BUCHANAN, J. GLICK*
- Reasoning with Cause and Effect, *J. PEARL*
- Abduction, Reason, and Science: A Review, *A. ALISEDA*
- The Logic of Knowledge Bases: A Review, *E. MOTTA*
- Heterogeneous Agent Systems: A Review, *P. RAVI PRAKASH*
- Dynamic Logic: A Review, *A. SILVA*
- Automated Theorem Proving: A Review, *G. SUTCLIFFE*

AI MAGAZINE V23 N°2 2002

- Editorial Introduction: The Fourteenth Innovative Applications of Artificial Intelligence Conference (IAAI-2001), *S. CHIEN, H. HIRSH*
- Electric Elves: Agent Technology for Supporting Human Organizations, *H. CHALUPSKY, Y. GIL, C. A. KNOBLOCK, K. LERMAN, J. OH, D. V. PYNADATH, T. A. RUSS, M. TAMBE*
- The RADARSAT-MAMM Automated Mission Planner, *B. D. SMITH, B. E. ENGELHARDT, D. H. MUTZ*
- Interchanging Agents and Humans in Military Simulation, *C. HEINZE, S. GOSS, T. JOSEFSSON, K. BENNETT, S. WAUGH, I. LLOYD, G. MURRAY, J. OLDFIELD*
- CARMA: A Case-Based Rangeland Management Adviser, *J. HASTINGS, K. BRANTING, J. LOCKWOOD*
- Natural Language Assistant: A Dialog System for Online Product Recommendation, *J. CHAI, V. HORVATH, N. NICOLOV, M. STYS, N. KAMBHATLA, W. ZADROZNY, P. MELVILLE*

SOMMAIRES DES REVUES

- TALPS: The T-AVB Automated Load-Planning System, *P. S. CERKEZ*
- The Sixth International Conference on Intelligent User Interfaces, *K. J. HAMMOND*
- AAAI 2002 Fall Symposium Series Reports, Benjamin Bell, *L. D. CANAMERO, S. CORADESCHI, C. GOMES, A. SAFFIOTTI, C. TSATSOUKIS, T. WALSH*
- Autonomous Mental Development: Workshop on Development and Learning (WDL), *J. WENG, I. STOCKMAN*
- Human Language Technology and Knowledge Management, *M. T. MAYBURY*
- Similarity and Categorization: A Review, *B. C. LOVE*
- Evolutionary Robotics: A Review, *H. VOICU*

COGNITIVE SCIENCE

<http://www.elsevier.com/inca/publications/store/6/2/0/1/9/4/>

APPLIED ARTIFICIAL INTELLIGENCE

<http://www.tandf.co.uk/journals/tf/08839514.html>

APPLIED ARTIFICIAL INTELLIGENCE V16 N°5 MAY 2002

- Discovering causality in large databases, *S. ZHANG, C. ZHANG*
- Negotiation in a multi-agent system for construction claims negotiation, *Z. REN, C. J. ANUMBA, O. O. UGWU*
- Improving biosignal processing through modeling uncertainty: Bayes vs. Non-bayes in sleep staging, *P. SYKACEK, G. DORFFNER, P. RAPPELSBERGER, ET AL.*

APPLIED ARTIFICIAL INTELLIGENCE V16 N°6 JULY 2002

- A neural network based approach for measurement dynamics compensation, *P. GEORGIEVA, S. FEYO DE AZEVEDO*
- ABC classification with uncertain data. A fuzzy model vs. a probabilistic model, *J. PUENTE, D. DE LA FUENTE, P. PRIORE, ET AL.*
- An adaptive real-coded genetic algorithm, *L. HAY LEE, Y. FAN*

APPLIED ARTIFICIAL INTELLIGENCE V16 N°7 AUGUST 2002

- Toward merging cognition and affect in HCI, *F. DE ROSIS*
- Toward recognizing and conveying an attitude of doubt via natural language, *S. CARBERRY, L. SCHROEDER*
- Scripting affective communication with life-like characters

in web-based interaction systems, *H. PRENDINGER, S. DES-CAMPS, M. ISHIZUKA*

- Probabilistic assessment of user's emotions in educational games, *C. CONATI*
- Can a rational agent afford to be affectless? A formal approach, *C. L. LISETTI, P. GMYTRASIEWICZ*
- This time with feeling: integrated model of trait and state effects on cognition and behavior, *E. HUDLICKA*

COMPUTATIONAL INTELLIGENCE

<http://www.blackwellpublishers.co.uk/asp/journal.asp?ref=08247935&src=cts>

COMPUTATIONAL INTELLIGENCE V18 N°2 2002

- Trends in Agent Communication Language, *B. CHAIB-DRAA, F. DIGNUM*
- An ACL for a Dynamic System of Agents, *M. GASPARI*
- To Commit or Not to Commit: Modeling Agent Conversations for Action, *R. A. FLORES, R. C. KREMER*
- Toward A Formalism for Conversation Protocols Using Joint Intention Theory, *S. KUMAR, M. J. HUBER, P. R. COHEN, D. R. MCGEE*
- Negotiating the Semantics of Agent Communication Languages, *C. REED, T. J. NORMAN, N. R. JENNINGS*

COMPUTATIONAL INTELLIGENCE V18 N°3 2002

- Introduction to the Special Issue, *N. CERCONE, B. SIRINAOVAKUL, V. WUWONGSE*
- Generate and Repair Machine Translation, *K. NARUEDOMKUL, N. CERCONE*
- Handwritten Thai Character Recognition Using Fourier Descriptors and Genetic Neural Networks, *P. PHOKKHARATKUL, C. KIMPAN*
- English-Thai Structure-Based Machine Translation, *B. SIRINAOVAKUL, K. CHANCHAROEN*
- A Method for Isolated Thai Tone Recognition Using a Combination of Neural Networks, *N. THUBTHONG, B. KIJSIRIKUL, A. PUSITTRAKUL*
- Toward a Unifying Ontology Dealing with Both User-Defined Periodicity and Temporal Constraints About Repeated Events, *P. TEREZIANI*
- On the Hybrid Propositional Encodings of Planning, *A. DATTATRAYA MALI*
- Topological Spatio-Temporal Reasoning and Representation, *P. MULLER*
- A Shallow Text Processing Core Engine, Günter Neumann, *J. PISKORSKI*

SOMMAIRES DES REVUES

MINDS AND MACHINES

JOURNAL FOR ARTIFICIAL INTELLIGENCE PHILOSOPHY AND COGNITIVE SCIENCE

<http://www.wkap.nl/prod/j/0924-6495>

MINDS AND MACHINES V12 N°3 AUGUST 2002

- Games Machines Play, *W. C. STIRLING*
- Löb's Theorem as a Limitation on Mechanism, *M. DETLEFSEN*
- Selective Representing and World-Making, *P. MANDIK, A. CLARK*
- A Conflict Between Language and Atomistic Information, *C. WASHINGTON*
- Does Classicism Explain Universality?, *S. PHILLIPS*
- vDiscussion The Importance of Continuity: A Reply to Chris Eliasmith, *R. R. POZNANSKI*
- Discreteness and Relevance: A Reply to Roman Poznanski, *C. ELIASMITH*
- Dennis Ross, Andrew Brook, and David Thompson, eds., *Dennett's Philosophy: A Comprehensive Assessment*, *R. M. HEITNER*
- Jerry Fodor, Concepts: Where Cognitive Science Went Wrong, *D. COLE*
- Jean-Pierre Dupuy, The Mechanization of Mind: On the Origins of Cognitive Science, *G. PICCININI*
- Timothy R. Colburn, Philosophy and Computer Science, *B. INDURKHYA*

ARTIFICIAL INTELLIGENCE REVIEW

<http://www.kluweronline.com/issn/0269-2821/contents>

AI REVIEW V18 N°1 SEPTEMBER 2002

- An Evaluation of Evolutionary Generalisation in Genetic Programming, *I. KUSHCHU*
- EDU-EX: A Tool for Auto-Regulated Intelligent Tutoring Systems Development Based on Models, *P. DOMINGO, A. GARCÍA-CRESPO, B. RUIZ, A. IGLESIAS*
- User Modeling for Personalized City Tours, *J. FINK, A. KOBSA*

AI REVIEW V18 N°2 OCTOBER 2002

- A Perspective View and Survey of Meta-Learning, *R. VILALTA, Y. DRISSI*
- Learning Intelligent Behavior in a Non-stationary and Partially Observable Environment, *S. ENKUL, F. POLAT*
- User Interfaces and Help Systems: From Helplessness to Intelligent Assistance, *S. DELISLE, B. MOULIN*

INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES

<http://www.academicpress.com/ijhcs>

IJHCS V56 N°5 MAY 2002

- Learner outcomes in an asynchronous distance education environment, *A. D. CARSWELL, V. VENKATESH*
- KJ3-a tool assisting formal validation of knowledge-based systems, *C.-H. WU, S.-J. LEE*
- A metadata filter for intranet portal organizational memory information systems, *Y. G. JI, G. SALVENDY*
- An empirical study of on-line help design: features and principles, *H. C. PURCHASE, J. WORRILL*

IJHCS V56 N°6 JUNE 2002

- Work domain analysis and sensors I: principles and simple example, *D. V. C. REISING, P. M. SANDERSON*
- Work domain analysis and sensors II: Pasteurizer II case study, *D. V. C. REISING, P. M. SANDERSON*
- Ontological methodology, *ROBERTO POLI*
- A cooperative framework for integrating ontologies, *J. T. FERNÁNDEZ-BREIS, R. MARTINEZ-BÉJAR*

IJHCS V57 N°1 JULY 2002

- The effects of graphic organizers giving cues to the structure of a hypertext document on users' navigation strategies and performance, *R. M. NILSSON, R. E. MAYER*
- Vision-based user interfaces: methods and applications, *M. PORTA*
- A case study on integrating contextual information with analytical usability evaluation, *A. BLANDFORD, G. RUGG*

IJHCS V57 N°2 AUGUST 2002

- Loglinear and multidimensional scaling models of digital library navigation, *B. P. BUTTENFIELD, R. F. REITSMA*
- Driving usability into the public administration: the Italian experience, *T. CATARCI, G. MATARAZZO, G. RAISS*
- Footprints of information foragers: behaviour semantics of visual exploration, *C. CHEN, T. CRIBBIN, J. KULJIS, R. MACREDIE*

IJHCS V57 N°3 SEPTEMBER 2002

- Modelling multiple mind-matter interaction, *C. M. JONKER, J. TREUR*
- Understanding user acceptance of digital libraries: what are the roles of interface characteristics, organizational context, and individual differences?, *J. Y. L. THONG, W. HONG, K.-Y. TAM*

SOMMAIRES DES REVUES

COMPUTATIONAL LINGUISTICS

<http://mitpress.mit.edu/catalog/item/default.asp?sid=8563C099-9701-4DD2-85C8-8F3502E9C8AE&ttype=4&tid=10>

COMPUTATIONAL LINGUISTICS V28 N°2 JUNE 2002

- Near-Synonymy and Lexical Choice, *P. EDMONDS, G. HIRST*
- The Combinatory Morphemic Lexicon, *C. BOZSAHIN*
- Class-Based Probability Estimation Using a Semantic Hierarchy, *S. CLARK, D. WEIR*
- Incremental Construction and Maintenance of Minimal Finite-State Automata, *R. C. CARRASCO, M. L. FORCADA*

Book review

- Spotting and Discovering Terms through Natural Language Processing by C. Jacquemin, *S. ANANIADOU*
- Automatic Summarization by I. Mani, *C. D. PAICE*
- Efficient Processing with Constraint-Logic Grammars Using Grammar Compilation by G. Minnen, *S. MANANDHAR*
- The Language of Word Meaning edited by P. Bouillon and F. Busa, *M. BROMAN OLSEN*
- Empirical Methods for Exploiting Parallel Texts, *T. PEDERSEN*
- Sentence Comprehension: The Integration of Habits and Rules by D. J. Townsend and T. G. Bever, *M. W. CROCKER*

COMPUTATIONAL LINGUISTICS V28 N°3 SEPTEMBER 2002

- Automatic Labeling of Semantic Roles, *D. GILDEA, D. JURAFSKY*
- Periods, Capitalized Words, etc., *A. MIKHEEV*
- Toward an Aposynthesis of Topic Continuity and Intrasentential Anaphora, *E. MILTSAKAKI*
- The Disambiguation of Nominalizations, *M. LAPATA*
- A Note on Typing Feature Structures, *S. WINTNER, A. SARKAR*

NATURAL LANGUAGE ENGINEERING

<http://www.journals.cambridge.org/bin/bladerunner?REQUNIQ=976614197&REQSESS=5004848&116000REQVENT=&REQINT1=5&REQSTR1=NLE&REQAUTH=0>

NLE V8 N°1 MARCH 2002

- An LVQ connectionist solution to the non-determinacy problem in Arabic morphological analysis: a learning hybrid algorithm, *M. A. EL-AFFENDI*
- Word clustering and disambiguation based on co-occurrence data, *H. LI*

- SUMMAC: a text summarization evaluation, *I. MANI, G. KLEIN, D. HOUSE, L. HIRSCHMAN, T. FIRMIN, B. SUNDEHEIM*
- Contextually appropriate reference generation, *Ö. YÜKSEL, C. BOZSAHIN*

Book Review

- Christopher D. Manning and Hinrich Schütze. Foundations of Statistical Natural Language Processing. MIT Press, 2000., *A. SIDDHARTHAN*

NLE V8 N°2-3 JUNE 2002

- Robust methods in analysis of natural language data, *A. BALLIM, V. PALLOTTA*
- Parsing engineering and empirical robustness, *R. BASILI, F. M. ZANZOTTO*
- Robustness beyond shallowness: incremental deep parsing, *S. AÏT-MOKHTAR, J.-P. CHANOD, C. ROUX*
- Full parsing approximation for information extraction via finite-state cascades, *F. CIRAVEGNA, A. LAVELLI*
- Learning rules for information extraction, *J. TURMO, H. RODRIGUEZ*
- Selecting effective index terms using a decision tree, *T. TAKENOBU, K. KENJI, O. HIRONORI, T. HOZUMI*
- A bootstrapping approach for robust topic analysis, *O. FERRET, B. GRAU*
- Robust discourse parsing via discourse markers, topicality and position, *F. SCHILDER*
- Architectural elements of language engineering robustness, *D. MAYNARD, V. TABLAN, H. CUNNINGHAM, C. URSU, H. SAGGION, K. BONTICHEVA, Y. WILKS*

USER MODELING AND USER-ADAPTED INTERACTION

<http://www.wkap.nl/jrnltoec.htm/0924-1868>

COMPUTER SPEECH AND LANGUAGE

<http://www.idealibrary.com/links/toc/csla/latest>

COMPUTER SPEECH & LANGUAGE V16 N°2 APRIL 2002

- Using phonologically-constrained morphological analysis in continuous speech recognition, *M. HUCKVALE, A. C. FANG*
- Selection of the most significant parameters for duration modelling in a Spanish text-to-speech system using neural networks, *R. CÓRDOBA, J. M. MONTERO, J. M. GUTIÉRREZ, J. A. VALLEJO, E. ENRIQUEZ, J. M. PARDO*

SOMMAIRES DES REVUES

- Hidden Markov model training with contaminated speech material for distant-talking speech recognition, *M. MATASSONI, M. OMOLOGO, D. GIULIANI, P. SVAIZER*
- Transformation streams and the HMM error model, *M. J. F. GALES*
- From within-word model search to across-word model search in large vocabulary continuous speech recognition, *A. SIXTUS, H. NEY*

COMPUTER SPEECH & LANGUAGE V16 N°3-4 JULY 2002

- Spoken language generation, *M. A. WALKER, O. C. RAMBOW*
- Response planning and generation in the MERCURY flight reservation system, *S. SENEFF*
- A conversation acts model for generating spoken dialogue contributions, *A. J. STENT*
- Learning visually grounded words and syntax for a scene description task, *D. K. ROY*
- Stochastic natural language generation for spoken dialog systems, *A. H. OH, A. I. RUDNICKY*
- Training a sentence planner for spoken dialogue using boosting, *M. A. WALKER, O. C. RAMBOW, M. ROGATI*
- Trainable approaches to surface natural language generation and their application to conversational dialog systems, *A. RATNAPARKHI*
- Exploring features from natural language generation for prosody modeling, *S. PAN, K. MCKEOWN, J. HIRSCHBERG*
- Contrast in concept-to-speech generation, *M. THEUNE*
- Efficient integrated response generation from multiple targets using weighted finite state transducers, *I. BULYKO, M. OSTENDORF*

MACHINE LEARNING

<http://www.wkap.nl/jrnltoe.htm/0885-6125>

MACHINE LEARNING V49 N°1 OCTOBER 2002

- The Lagging Anchor Algorithm: Reinforcement Learning in Two-Player Zero-Sum Games with Imperfect Information, *F. A. DAHL*
- A Simple Method for Generating Additive Clustering Models with Limited Complexity, *M. D. LEE*
- Feature Generation Using General Constructor Functions, *S. MARKOVITCH, D. ROSENSTEIN*

MACHINE LEARNING V49 N°2-3 NOV-DECEMBER 2002

- Introduction, *S. SINGH*

- Reinforcement Learning for Call Admission Control and Routing under Quality of Service Constraints in Multimedia Networks, *H. TONG, T. X. BROWN*
- Building a Basic Block Instruction Scheduler with Reinforcement Learning and Rollouts, *A. MCGOVERN, E. MOSS, A. G. BARTO*
- Kernel-Based Reinforcement Learning, *D. ORMONEIT, S. SEN*
- On Average Versus Discounted Reward Temporal-Difference Learning, *J. N. TSITSIKLIS, B. VAN ROY*
- A Sparse Sampling Algorithm for Near-Optimal Planning in Large Markov Decision Processes, *M. KEARNS, Y. MANSOUR, A. Y. NG*
- Near-Optimal Reinforcement Learning in Polynomial Time, *M. KEARNS, S. SINGH*
- Technical Update: Least-Squares Temporal Difference Learning, *J. A. BOYAN*
- Continuous-Action Q-Learning, *J. DEL R. MILLÁN, D. POSENATO, E. DEDIEU*
- Risk-Sensitive Reinforcement Learning, *O. MIHATSCH, R. NEUNEIER*
- Variable Resolution Discretization in Optimal Control, *R. MUNOS, A. MOORE*
- Structure in the Space of Value Functions, *D. FOSTER, P. DAYAN*

MACHINE LEARNING V50 N°1-2 JAN-FEBRUARY 2002

- An Introduction to MCMC for Machine Learning, *C. ANDRIEU, N. DE FREITAS, A. DOUCET, M. I. JORDAN*
- EM, MCMC, and Chain Flipping for Structure from Motion with Unknown Correspondence, *F. DELLAERT, S. M. SEITZ, C. E. THORPE, S. THRUN*
- Mixtures of Factor Analysers. Bayesian Estimation and Inference by Stochastic Simulation, *E. FOKOUÉ, D.M. TITTERINGTON*
- Being Bayesian About Network Structure. A Bayesian Approach to Structure Discovery in Bayesian Networks, *N. FRIEDMAN, D. KOLLER*
- Improving Markov Chain Monte Carlo Model Search for Data Mining, *P. GIUDICI, R. CASTELO*
- Classification with Bayesian MARS, *C.C. HOLMES, D.G.T. DENISON*
- Population Markov Chain Monte Carlo, *K. B. LASKEY, J. W. MYERS*
- A Noninformative Prior for Neural Networks, *H. K.H. LEE*

NEURAL NETWORKS

<http://www.elsevier.com/inca/publications/store/8/4/1/>

NEURAL NETWORKS V15 N°4-6 JUNE-JULY 2002

2002 Special issue

- Introduction for 2002 Special Issue: Computational Models of Neuromodulation, *K. DOYA, P. DAYAN, M. E. HASSELMO*
- Cellular, synaptic and network effects of neuromodulation, *E. MARDER, V. THIRUMALAI*
- Metalearning and neuromodulation, *K. DOYA*
- Dopamine-dependent plasticity of corticostriatal synapses, *J. N.J. REYNOLDS, J. R. WICKENS*
- TD models of reward predictive responses in dopamine neurons, *R. E. SURI*
- Actor-critic models of the basal ganglia: new anatomical and computational perspectives, *D. JOEL, Y. NIV, E. RUPPIN*
- Dopamine: generalization and bonuses, *S. KAKADE, P. DAYAN*
- The computational role of dopamine D1 receptors in working memory, *D. DURSTEWITZ, J. K. SEAMANS*
- Dopamine controls fundamental cognitive operations of multi-target spatial working memory, *S. TANAKA*
- An integrative theory of the phasic and tonic modes of dopamine modulation in the prefrontal cortex, *J.-C. DREHER, Y. BURNOD*
- Opponent interactions between serotonin and dopamine, *N. D. DAW, S. KAKADE, P. DAYAN*
- Local analysis of behaviour in the adjusting-delay task for assessing choice of delayed reinforcement, *R. N. CARDINAL, N. DAW, T. W. ROBBINS, B. J. EVERITT*
- Neuromodulation of decision and response selection, *M. USHER, E. J. DAVELAAR*
- Simplified dynamics in a model of noradrenergic modulation of cognitive performance, *M. S. GILZENRAT, B. D. HOLMES, J. RAJKOWSKI, G. ASTON-JONES, J. D. COHEN*
- Control of exploitation-exploration meta-parameter in reinforcement learning, *S. ISHII, W. YOSHIDA, J. YOSHIMOTO*
- Neuromodulation, theta rhythm and rat spatial navigation, *M. E. HASSELMO, J. HAY, M. ILYN, A. GORCHETCHNIKOV*
- Cholinergic modulation of sensory representations in the olfactory bulb, *C. LINSTER AND T. A. CLELAND*
- Acetylcholine in cortical inference, *A. J. YU, P. DAYAN*
- Sensory-motor gating and cognitive control by the brainstem cholinergic system, *Y. KOBAYASHI, T. ISA*
- On-line learning in changing environments with applications in supervised and unsupervised learning, *N. MURATA, M. KAWANABE, A. ZIEHE, K.-R. MÜLLER, S.-I. AMARI*

- Neuromodulation and plasticity in an autonomous robot, *O. SPORNS, W. H. ALEXANDER*

NEURAL NETWORKS V15 N°7 SEPTEMBER 2002

- Color constancy and the functional significance of McCollough effects, *T. VLADUSICH, J. BROERSE*
- Neural network models for the gaze shift system in the superior colliculus and cerebellum, *X. WANG, J. JIN, M. JABRI*
- A complex-valued version of Nagumo-Sato model of a single neuron and its behavior, *I. NEMOTO, K. SAITO*
- Delay-dependent exponential stability analysis of delayed neural networks: an LMI approach, *X. LIAO, G. CHEN, E. N. SANCHEZ*
- Global stability of neural networks with distributed delays, *Y. CHEN*
- SpikeCell: a deterministic spiking neuron, *C. GODIN, M.B. GORDON, J.D. MULLER*
- A feed-forward network for input that is both categorical and quantitative, *R. K. BROUWER*
- A spiking neuron model: applications and learning, *C. CHRISTODOULOU, G. BUGMANN, T. G. CLARKSON*
- An intelligent sales forecasting system through integration of artificial neural networks and fuzzy neural networks with fuzzy weight elimination, *R.J. KUO, P. WU, C.P. WANG*

Book review

- Awakening a Sleeping Cat: A Review of "Information Theory and the Brain" - R. Baddeley, P. Hancock, P. Földiák (editors), Cambridge University Press, 2000, *R. P.N. RAO*
- Learning Kernel Classifiers - R. Herbrich (Ed.); MIT Press, 2002, *R. WILLIAMSON*

Erratum

Erratum to: "The double queue method: a numerical method for integrate-and-fire neuron networks", *G. LEE, N. H. FARHAT*

ARTIFICIAL INTELLIGENCE IN MEDICINE

<http://www.elsevier.com/inca/publications/store/5/0/5/6/2/7/>

AI IN MEDICINE V25 N°3 JULY 2002

- Evaluation of two different models of semi-automatic knowledge acquisition for the medical consultant system CADIAG-II/RHEUMA, *H. LEITICH, K.-P. ADLASSNIG, G. KOLARZ*
- Combining biometric and symbolic models for customised, automated prosthesis design, *S. MODGIL, T.J. HUTTON, P. HAMMOND, J.C. DAVENPORT*

SOMMAIRES DES REVUES

- NasoNet, modeling the spread of nasopharyngeal cancer with networks of probabilistic events in discrete time, *S.F. GALÁN, F. AGUADO, F.J. DÍEZ, J. MIRA*
- An evolutionary artificial neural networks approach for breast cancer diagnosis, *H. A. ABBASS*
- Objective detection of brainstem auditory evoked potentials with a priori information from higher presentation levels, *E. VANNIER, O. ADAM, J.-F. MOTSCH*

Book review

- Shape Analysis and Classification. Theory and Practice - Luciano da Fontoura Costa and Roberto Marcondes Cesar Jr.; CRC Press LLC, Boca Raton, 2001, R. C. VELTKAMP

AI IN MEDICINE V26 N°1-2 SEP-OCTOBER 2002

- Uniqueness of medical data mining, *K. J. CIOS AND G. W. MOORE*
- Confidentiality issues for medical data miners, *J. J. BERMAN*
- Data mining a diabetic data warehouse, *J. L. BREAUULT, C. R. GOODALL, P. J. FOS*
- Using dependency/association rules to find indications for computed tomography in a head trauma dataset, *S. P. IMBERMAN, B. DOMANSKI, H. W. THOMPSON*
- Analysis of respiratory pressure-volume curves in intensive care medicine using inductive machine learning, *S. GANZERT ET AL.*
- Logistic-based patient grouping for multi-disciplinary treatment, *L. MARUSTER, T. WEIJTERS, G. DE VRIES, A. VAN DEN BOSCH, W. DAELEMANS*
- Bayesian learning for cardiac SPECT image interpretation, *J. P. SACHA, L. S. GOODENDAY, K. J. CIOS*
- Syntactic reasoning and pattern recognition for analysis of coronary artery images, *M. R. OGIELA, R. TADEUSIEWICZ*

- Mining knowledge for HEp-2 cell image classification, *P. PERNER, H. PERNER, B. MÜLLER*

Book review

- Principles of Data Mining - D. Hand, H. Mannila, P. Smyth (Eds.), MIT Press, Cambridge, MA, 2001, A. SMITH

INTERNATIONAL JOURNAL OF APPROXIMATE REASONING

<http://www.elsevier.com/inca/publications/store/5/0/5/7/8/7/>

IJAR V30 N°3 SEPTEMBER 2002

- A heuristic algorithm for computing the max-min inverse fuzzy relation, *P. SAHA, A. KONAR*
- Neuro-fuzzy approach to processing inputs with missing values in pattern recognition problems, *B. GABRYS*
- vNetworks of probabilistic events in discrete time, *S. F. GALÁN, F. J. DÍEZ*
- A parametric representation of linguistic hedges in Zadeh's fuzzy logic, *V.N. HUYNH, T.B. HO, Y. NAKAMORI*

IJAR V31 N°1-2 OCTOBER 2002

- The application of the matrix calculus to belief functions, *P. SMETS*
- Belief function independence: II. The conditional case, *B. BEN YAGHLANE, P. SMETS, K. MELLOULI*
- Approximating the combination of belief functions using the fast Möbius transform in a coarsened frame, *T. DENOEU, A. BEN YAGHLANE*
- Resource bounded and anytime approximation of belief function computations, *R. HAENNI, N. LEHMANN*

Création par réutilisation de documents décrits par les intentions des auteurs

Yahya Al-Tawki

Thèse soutenue le 29 avril 2002 à l'université Toulouse 1.

Résumé :

Cette thèse est une contribution au développement des systèmes auteurs fondés sur des représentations de connaissances utilisées dans le processus de création de document par réutilisation. Pour faciliter la réutilisation de documents, considérée comme processus de base dans la création, il est nécessaire d'aider à retrouver les documents les plus pertinents au sujet développé. La description des documents en termes des concepts utilisés dans le domaine d'application et en termes d'intentions de l'auteur est étudiée d'un point de vue conceptuel. Cette thèse propose une modélisation des structures intentionnelles adéquates en utilisant la théorie des actes de langage, les connaissances du domaine sont représentées comme réseaux sémantiques fondés sur une ontologie du domaine. XML a servi pour codifier ces connaissances en tant que méta-données. Le système SABRE d'aide à la création de documents par la réutilisation, fondé sur les modèles de description (ontologie et intention) est développé. Le domaine d'application est l'enseignement de la recherche d'information.

Mots clés : système auteur, réutilisation, intention, ontologie, XML, méta-données, actes de langage.

Contact : Saïd Tazi, tazi@univ-tlse1.fr

cations scientifiques car elles sont génériques, validantes et orientées vers l'historique des données. La validation est assurée malgré l'ouverture de la base grâce à la double contrainte qu'offrent les règles d'autorisation et de production. La traçabilité – la capacité de la base à produire un historique pour chaque donnée – découle du fait que l'ensemble des contributions est strictement incrémental.

Afin de valider la méthode, nous avons développé un serveur de PCE nommé Quilombo fondé sur le modèle déductif de bases de données. Les actions sont représentées par des relations extensionnelles, et les règles d'autorisation et de production par des relations intentionnelles. Les PCE sont exprimées par des clauses du calcul des prédicats.

Quilombo a permis de mettre en oeuvre différentes bases coopératives dans des domaines aussi variés qu'un forum de discussions, qu'une base terminologique ou une aide à la nomenclature taxonomique.

Contact : Robert Bossy Laboratoire Informatique et Systématique, Université Pierre et Marie Curie, 12 rue Cuvier, 75005 Paris.

Tél. : 01 44 27 48 27

Fax. : 01 44 27 65 50

Mél : bossy@ccr.jussieu.fr

Edition coopérative de bases de données scientifiques

Robert Bossy

Thèse soutenue 25 avril 2002 à l'UTES, Université Paris 6

Résumé :

Une base de données coopérative permet aux utilisateurs non seulement de consulter le contenu mais aussi d'ajouter, corriger ou modifier des éléments. Les bases de référence biologiques ne peuvent plus être maintenues par des groupes limités d'éditeurs car la masse d'informations croît de façon exponentielle.

Nous présentons une méthode originale d'édition coopérative de bases de données scientifiques. Cette méthode s'appuie sur la déclaration d'un modèle de coopération entre différents contributeurs, appelé Politique de Contrôle d'Édition (PCE). Les informations contenues dans une base contrôlée par une PCE sont divisées en actions, en règles d'autorisation et en règles de production. Les actions sont des contributions brutes des différents éditeurs ; elles forment l'ensemble des modifications apportées aux données.

La PCE ne prévoit que l'accumulation d'actions afin de garder la trace de toute information présente dans la base. Les règles d'autorisation sont des énoncés conditionnant l'addition de nouvelles actions, elles permettent des stratégies d'autorisations plus élaborées qu'une matrice de droits d'édition. Les règles de production permettent de calculer une vision stable et valide de la base à partir des actions emmagasinées.

Les PCE sont particulièrement adaptées à diverses appli-

Démarche méthodologique de conception de systèmes d'information coopératifs inter agents pour la gestion des connaissances

Imed Boughzala

Thèse de l'Université Paris 6, soutenue le 20 décembre 2001.

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 fondées 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.

Contact : Département GSID – UTT,
12, rue de Marie Curie - BP 2060, 10010 Troyes CEDEX.
Téléphone : 03 25 71 56 88
Télécopie : 03 25 71 56 99
Mél : Imed.boughzala@utt.fr

Svetlan', un système de structuration du lexique guidé par la détermination automatique du contexte thématique Gaël de Chalendar

Thèse de l'Université Paris 11, soutenue le 21 décembre 2001 à l'université de Paris 11 (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 : Gael de Chalendar, LIMSI/CNRS, BP 133,
91403 Orsay Cedex.
Tél: 33 1 69 85 80 04
Fax: 33 1 69 85 80 88
Mél : Gael.de.Chalendar@limsi.fr
URL : <http://www.limsi.fr/Individu/gael/>

Apprentissage par renforcement utilisant des réseaux de neurones, avec des applications au contrôle moteur Rémi Coulom

Doctorat de l'Université Joseph Fourier – Grenoble
Spécialité informatique, soutenu le mercredi 19 juin 2002
à l'INP Grenoble

Résumé :

Cette thèse est une étude de l'utilisation des capacités de généralisation des réseaux de neurones artificiels dans l'apprentissage par renforcement. Ce travail a permis de mettre au point des méthodes efficaces pour

l'application de l'algorithme TD(λ) continu à des tâches de contrôle moteur en dimension élevée. En particulier, l'algorithme a été raffiné pour traiter des problèmes avec une commande et des états discontinus, et pour rendre la descente de gradient beaucoup plus rapide dans le cas de l'utilisation de réseaux de neurones feedforward. Ces améliorations ont été appliquées avec succès à des problèmes complexes originaux, comme la conduite d'une voiture de course et la natation d'un robot articulé.

Mots clés : apprentissage par renforcement, réseaux de neurones, commande optimale, contrôle moteur

Contact : Rémi Coulom – Laboratoire Leibniz
46, avenue Félix Viallet. 38041 Grenoble Cedex.
Tél: +33 4 76 57 48 08
Mél : Remi.Coulom@imag.fr
URL : <http://remi.coulom.free.fr/>
Document récupérable à :
Laboratoire Leibniz
46, avenue Félix Viallet
38031 GRENOBLE CEDEX
<http://remi.coulom.free.fr/Thesis/>

Échantillonnage, Validation et Généralisation en Extraction des Connaissances à partir des Données **Jean-Hugues CHAUCHAT**

Habilitation à diriger des recherches de l'Université Lumière, Lyon 2 soutenue le 4 janvier 2002 à Lyon.

Résumé :

En ECD, Extraction des Connaissances à partir des Données (Knowelge Discovery in Databases), la base de données disponible doit être considérée comme un échantillon d'un univers-cible auquel les résultats devront être généralisés. Les procédures de validation (estimation du taux d'erreur en généralisation) doivent être adaptées pour correspondre au procédé de constitution de la base de données ; on détaille les cas de bases « en grappes » (tous les malades d'un échantillon de services, tous les clients d'un échantillon d'agences...) ou « en strates » (un échantillon de malades et un échantillon de témoins, ou un échantillon d'acheteurs et un de non-acheteurs...). Devant une très grande base de données, on peut accélérer l'apprentissage, sans perdre significativement en précision, en apprenant sur un échantillon ; l'échantillon équilibré est plus puissant. Enfin, les « graphiques de Bertin » permettent de visualiser les données et les connaissances.

Mots clés : data mining, validation, généralisation, échantillon, visualisation

Contact : Jean-Hugues Chauchat, Laboratoire ERIC - Université Lumière, Lyon-2,

5 avenue Pierre Mendès-France 69676 Bron Cedex.

Téléphone : 33 (0)4 78 77 23 79

Télécopie : 33 (0)4 78 77 23 75

Mél : chauchat@univ-lyon2.fr

URL : <http://eric.univ-lyon2.fr/~chauchat/>

Modélisation de la description d'images : Application au domaine médical **Richard Chbeir**

Thèse de l'Institut national des sciences appliquées de Lyon, soutenue le 14 décembre 2001 à Lyon.

Résumé :

De nombreux axes de recherche se sont développés autour de la gestion des images. Dans ce travail, nous nous sommes intéressés à la problématique de la recherche d'images dans une base de données médicales. Cette problématique est liée principalement à la complexité de la description d'image.

Nous proposons, dans cette thèse, un modèle original qui permet de décrire toutes les caractéristiques de l'image. Il est structuré selon deux espaces : 1- L'espace externe contenant des informations externes liées à l'image telles que le nom du patient, la date d'acquisition, le type d'imagerie, etc., 2- L'espace interne de l'image considérant son contenu physique (couleur, texture, etc.), spatial (forme, position), et sémantique (scène, interprétations, etc.). Le modèle a été élaboré avec plusieurs niveaux de granularité permettant de considérer les

caractéristiques de l'image entière et celles de ses objets. Nous avons défini un module de références et un module de règles assurant la cohérence entre les espaces de description.

Nous proposons également un méta-modèle de relation qui a pour but de fournir, de façon précise, plusieurs types de relations entre deux objets en se fondant sur leurs caractéristiques communes (forme, couleur, position, etc.). Ce méta-modèle contribue à définir un mécanisme d'indexation performant.

Afin de valider notre approche, nous avons développé un prototype nommé MIMS (Medical Image Management System) utilisant des interfaces conviviales d'interrogation et de stockage d'images à base d'icônes et d'hypermédias.

Mots clés : modèle de description, méta-modèle de relation, indexation, multimédia.

Contact : Richard Chbeir, LISI - INSA de Lyon, bâtiment Blaise Pascal,

20, Avenue Albert Einstein,
69621 VILLEURBANNE Cedex.

Téléphone : 04 72 43 88 99

Mél : richard.chbeir@insa-lyon.fr

URL : <http://mims.myip.org>

Codage et traitement phylogénétique de caractères structuraux de génomes entiers **Cyril Gallut**

Thèse de l'Université Pierre et Marie Curie Paris 6, soutenue le 18 décembre 2001 Paris.

Résumé :

Une approche originale de reconstruction phylogénétique 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.
Téléphone : 01 44 27 65 21
Télécopie : 01 44 27 65 60
Mél : gallut@ccr.jussieu.fr
URL : <http://lis.snv.jussieu.fr/~gallut>

Modélisation et traitement de formes spatiales et dynamiques. Reconstruction 3d, mouvement et Geste, Architecture du Génome. Rachid Gherbi

Habilitation à diriger des recherches de l'Université de Paris-Sud 11, soutenue le 5 décembre 2001 à l'université de Paris-Sud (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 bioinformatique. 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'objets à la synthèse d'images.

Dans ce contexte de recherche fondamentale menée 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 et 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 d'avantage 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, Groupe Geste & Image, LIMSI-CNRS & Université Paris-Sud, Orsay.

Tel : +33/0 1 69 85 81 64

Fax : +33/0 1 69 85 80 88

URL : <http://www.limsi.fr/Individu/gherbi>

Adresse du document : <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 Ambroise Ingold

Thèse de l'université d'Aix-Marseille 3, soutenue le 4 janvier 2002 à Marseille.

Résumé :

La thèse propose des solutions pour mettre automatiquement en relation des informations bibliographiques avec des informations factuelles. Les bases de données bibliographiques donnent 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 électroniques 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.
Tél. : 06 61 66 70 90
Mél : ingold@free.fr.

recherche (systèmes trop interactifs, critères visuels, problèmes d'accessibilité, etc.).

Mots clés : recherche documentaire, réutilisation d'expérience, raisonnement à partir de cas, reformulation de requête, profil utilisateur, modélisation, déficients visuels.

Contact : LISI, Bat Blaise Pascal, INSA de Lyon,
69621 Villeurbanne Cedex
Mél : lobna.jeribi@lisi.insa-lyon.fr

Aide à la recherche documentaire adaptée à l'utilisateur : approche par réutilisation d'expériences. Cas d'application : accès au contenu des documents scientifiques par les déficients visuels

Lobna Jeribi

Thèse de l'INSA de Lyon, soutenue le 7 décembre 2001 à Lyon.

Résumé :

Dans le but de concevoir un système de recherche documentaire permettant d'obtenir des réponses adaptées et pertinentes pour l'utilisateur, on propose d'intégrer le Raisonnement à Partir de Cas (RàPC) dans le processus de recherche. En effet, le RàPC reflète une méthode de l'esprit humain qui consiste à résoudre un problème donné en se remémorant une expérience similaire passée.

Ainsi, les résultats d'instances antérieures de recherche sont exploités afin d'aider l'utilisateur à mieux formuler sa requête, moyennant une mémoire collective d'instances. Dans ce but, nous avons proposé une modélisation et formalisation d'une instance de recherche, caractérisées par un modèle de l'utilisateur et un modèle de la recherche documentaire et défini les fonctions de similarités correspondantes. La réutilisation a pour but d'enrichir la requête initiale de l'utilisateur par les termes de documents contenus dans les instances similaires retrouvées. Contrairement à la méthode classique Rocchio, ces documents sont ceux préalablement validés par des utilisateurs, ayant des profils similaires à l'utilisateur courant et se trouvant dans un contexte de recherche similaire.

L'originalité de notre approche consiste alors à doter le système de recherche par des connaissances a priori, grâce à la mémoire collective d'instance, lui permettant de « contextualiser » et de personnaliser les réponses documentaires, tout en réduisant les interactions avec l'utilisateur.

Ces propositions ont été implantées dans notre prototype COSYDOR (système coopératif de recherche documentaire), développé à partir d'Intermedia (Oracle 8i). Les évaluations ont été effectuées à partir du corpus de tests de TREC. Les premiers résultats montrent une amélioration nette de performance par rapport à Intermedia.

Ce travail est effectué dans le cadre d'un projet de la région Rhône-Alpes, dans lequel les déficients visuels constituent notre cas d'application. Cette recherche représente un apport considérable pour ces utilisateurs, au regard de leurs difficultés d'exploitation de systèmes actuels d'aide à la

Des modèles en Logiques de Description pour l'aide à la décision Martina Kullmann

Thèse de l'Université Louis Pasteur de Strasbourg, soutenue le 25 mai 2001.

Résumé :

En utilisant les Logiques de Description (LD) pour la modélisation des situations complexes et dynamiques ainsi que des réactions appropriées, nous proposons un schéma fondé 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 fondant 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 calcul. 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 : logiques de descriptions, aide à la décision, modélisation.

Contact : Martina Kullmann, Essenweinstr.
18, D-76131 Karlsruhe.
Tel.: (+49) 0721-7817344
Mél : m.kullmann@gmx.net

Extraction des connaissances basée sur les treillis de Galois: méthodes et applications. Engelbert Mephu Nguifo

Habilitation à diriger des recherches de l'Université d'Artois, soutenue le 10 décembre 2001 à Lens.

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 pré-traitement de données, et de classification par treillis pour la fouille de données.

La première partie introduit la problématique du pré-traitement de données en traitant deux types de méthodes : la sélection et la construction d'attributs. Ensuite sont abordées 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 fondées 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, bio-informatique.

Contact : Engelbert Mephu-Nguifo, CRIL - IUT de Lens - Université d'Artois, Rue de l'Université SP 16, 62307 Lens cedex.

Tél.: (+33) (0)3 21 79 32 73

Télécopie: (+33) (0)3 21 79 32 72

Mél : mephu@cril.univ-artois.fr

URL : <http://www.cril.univ-artois.fr/~mephu>

Architecture de systèmes d'accès à des bases de données hétérogènes pour l'aide à la décision Cécile Nicolle

Thèse de doctorat de l'INSA de Lyon, soutenue le 13 décembre 2001 à Lyon.

Résumé :

De tous temps, lors d'une prise de décision, le décideur a dû faire face au problème d'accès à toutes les données qui lui sont nécessaires pour prendre une décision juste. De nos jours, de nombreux systèmes proposent une aide à cette prise de décision. Mais il est encore difficile, pour le décideur, de savoir où trouver les informations voulues. De plus, il peut ne pas connaître la nature de toutes les données qui lui sont utiles dans sa prise de décision. C'est pourquoi nous proposons une architecture de système d'accès qui permet au décideur de poser une requête en langage proche du langage naturel, sans avoir besoin de préciser où rechercher les données en question. Le système sait où trouver les informations, et peut même fournir au décideur des informations liées à celles qu'il voulait, ces informations étant utiles à la prise de décision (complétant ainsi la recherche de l'utilisateur tout en fournissant des réponses pertinentes). Le système pallie ainsi d'éventuelles lacunes du décideur dans le domaine de recherche. Notre système utilise le principe des wrappers, ainsi que XML comme langage interne, langage de requêtes et de réponse. Deux prototypes ont été réalisés au cours de la thèse, un sur la recherche dans une base de textes juridiques, un autre sur l'interrogation en XML d'une base Progress avec réponse en XML.

Mots clés : bases de données hétérogènes, système d'accès, XML, wrapper, aide à la décision

Contact : Cécile Nicolle,

25bis rue Victor Hugo. 69 100 Villeurbanne.

Tél. : 04 72 34 83 53

Mél : cnicolle2001@yahoo.fr

URL : <http://www.multimania.com/cecilenicolle>

Une méthode d'indexation sémantique adaptée aux corpus multilingues Catherine Roussey

Thèse de doctorat de l'Institut National des Sciences Appliquées de Lyon, soutenue à Lyon le 10 décembre 2001.

Résumé :

Ces travaux de thèse s'inscrivent dans la problématique générale liée à l'indexation d'un corpus de texte pour la recherche d'information multilingue. Le but de l'indexation est d'identifier la connaissance contenue dans un texte et de la représenter par des mots clés appelés descripteurs. Or, la composante multilingue ajoute une complexité supplémentaire au processus d'indexation car une étape de traduction est obligatoire pour représenter document et requête par des descripteurs appartenant au même espace d'indexation. Pour traduire

correctement un terme, il est préférable de reconnaître le concept dénoté par celui-ci. Actuellement, une recherche d'information pertinente représente le contenu des documents par des concepts et non plus par des termes.

Cette thèse propose une méthode d'indexation sémantique pour les documents XML permettant de caractériser le contenu documentaire par des connaissances, non dépendante de la langue des documents. Ces connaissances sont déclinées en deux types :

- les connaissances du domaine sont utilisées pour représenter les documents et les requêtes dans ce même espace de représentation non dépendant des langues ;
- les connaissances terminologiques constituent plusieurs langages de présentation des connaissances du domaine.

Nous avons défini notre propre modèle de représentation des connaissances intitulé les graphes sémantiques. Ce modèle est un enrichissement du modèle des graphes conceptuels de Sowa. Premièrement, notre modèle différencie les connaissances du domaine, des connaissances terminologiques. Deuxièmement, une fonction de comparaison de graphes adaptée aux besoins réels de la recherche d'information est proposée.

Notre méthode d'indexation se veut générique car elle peut être utilisée aussi bien dans un système de recherche d'information multilingue que dans un hypertexte à base de connaissances. Pour valider notre proposition, un prototype, appelé SyDoM, a été implanté, répondant aux besoins d'une bibliothèque virtuelle.

Mots clés : recherche d'information multilingue, indexation sémantique, graphes conceptuels, thésaurus, XML.

Contact : Catherine Roussey, LISI Bât Blaise Pascal, INSA de Lyon,

20 av Albert Einstein, 69621 Villeurbanne Cedex.

Tél : 04 72 62 97

Fax : 04 72 43 87 13

Mél : croussey@lisi.insa-lyon.fr

URL : <http://lisi.insa-lyon.fr/~croussey>

Le modèle satisfaction-altruisme : coopération et résolution de conflits entre agents situés réactifs, application à la robotique Olivier Simonin

Thèse de l'Université de Montpellier 2, 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 satisfaction.

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 clés : systèmes multi-agents, agents situés réactifs, coopération, communication dans es SMA situés, robotique mobile distribuée

Contact : Olivier Simonin, LIRMM,

161 rue Ada, 34392 Montpellier Cedex 5.

Tél/Fax : 04.67.41.86.09 / 04.67.41.85.00

Mél : simonin@lirmm.fr

URL : <http://www.lirmm.fr/~simonin>

Automatisation du raisonnement et de la rédaction de preuves en géométrie de l'enseignement secondaire Jean-Pierre Spagnol

Thèse en intelligence artificielle de l'Université René Descartes – Paris 5, 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 de rendre opérationnels 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

RÉSUMÉS HABILITATIONS ET THÈSES

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 atteindre 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 lui-même, à par-

tir 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,

45 Rue des Saints-Pères 75270 PARIS CEDEX 06.

Mél : spaj@math-info.univ-paris5.fr

Pour l'envoi d'un résumé de thèse ou d'habilitation :

Envoyer à Marc-Philippe Huguet (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.

Dossier « L'IA dans le RNTL »

Date de parution prévue : 30 janvier 2003

Bulletin numéro 53

Coordinateurs : Bertrand Braunschweig (IFP, bertrand.braunschweig@ifp.fr) ; Jean-Luc Dormoy (EDF, jean-luc.dormoy@edf.fr) ; Patrick Gallinari (LIP6, patrick.gallinari@lip6.fr)

Objectifs

Le Réseau National des Technologies Logicielles existe depuis trois ans. L'appel à propositions 2003 sera donc le quatrième du nom.

Les technologies IA sont très présentes dans la thématique du RNTL, et de nombreux projets comportent un volet IA consistant. Le thème II, traditionnellement, est un peu le thème « IA » du RNTL. Mais des projets des autres thèmes mettent également en œuvre certains outils de l'IA, de manière explicite – dans les objectifs - ou bien cachée – dans les moyens.

Le but de ce dossier est de faire le point sur la présence des méthodes et outils de l'IA dans les projets RNTL existants. Par extension, les projets de deux autres réseaux, le RNRT (réseau national de recherche en télécommunications), et RIAM (Réseau pour la recherche et l'innovation dans l'audiovisuel et le multi-média) sont également invités à présenter leurs volets IA..

Ce dossier permettra, d'une part, aux équipes IA présentes dans les projets RNTL en cours, de faire connaître leurs contributions et de valoriser leur recherche ; d'autre part, il est également destiné aux porteurs de proposition 2003, pour leur donner des exemples de sujets considérés comme importants par le RNTL, et leur permettre de se référer à ces projets existants si nécessaire, par exemple en exploitant les résultats. Il montrera enfin l'importance de la contribution de l'IA aux programmes du RNTL

Qui peut soumettre ?

Tout laboratoire ou industriel de la communauté francophone impliqué dans un projet RNTL, RNRT ou PRIAMM 2000, 2001 ou 2002.

Comment soumettre ?

Les descriptions de projets seront organisées suivant le schéma suivant :

- identification du projet concerné (nom, titre long, année de démarrage, nature : exploratoire, pré-compétitif ou plate-forme) ;
- identification des équipes IA, adresse, téléphone, mél ; personne à contacter, contributeurs ;
- identification du reste du partenariat ;
- description générale du projet ;
- description détaillée de la partie IA du projet ;
- état d'avancement, résultat(s) obtenus ;
- courte bibliographie bien ciblée et adresse d'un ou plusieurs sites web pertinents, par exemple le site du projet.

Présentation : Les contributions feront 1000 mots maximum, et devront être envoyées pour le 20 décembre 2002 par courrier électronique au format Word ou RTF. Aucune mise en forme particulière ne doit être faite sur les textes autre que gras, italique et la taille des caractères.

Que faire ?

Si vous êtes concernés par cet appel, n'attendez pas d'être contactés. Même si les projets seront individuellement contactés, il est possible que nous ne soyons pas au courant du contenu IA d'un projet, et nous vous remercions donc de prendre l'initiative et de nous en informer.

Envoyez vos contributions à l'adresse suivante :

Bertrand Braunschweig
IFP/DISMA, 1 & 4 avenue de Bois-Préau, 92852
Rueil Malmaison Cédex
Mél : bertrand.braunschweig@ifp.fr

S'adresser pour tout renseignement, par courrier électronique ou téléphone, à un des coordinateurs du dossier.

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 (*ayer 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 :
 Nom et prénom du représentant : Fonction :
 Mél : Tél : Fax :
 Adresse postale :

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 AVEL, 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 diagnostic..... Bulletin n°52 Temps, espace et évolutif Bulletin n°51 Equipes d'IA en France Bulletin n°49/50 IA et Médecine Bulletin n°48 Fouille de données Bulletin n°46/47 IA et document..... Bulletin n°44 IA et connexionnisme..... 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</p>	<p>I.A . et biologie moléculaire..... Bulletin n°26 A. et droit Bulletin n°25 I.A. et fusion de données..... Bulletin n°24 I.A. et musique Bulletin n°23 Apprentissage..... Bulletin n°22 Les explications dans les SBC Bulletin n°20 Pétrole-Chimie Bulletin n°19 Le raisonnement à partir de cas..... Bulletin n°18 I.A. et temps-réel 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 connexionnisme..... 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>
---	--

S O M M A I R E D U B U L L E T I N N ° 5 2

L'IA au Québec	4
Présentation de laboratoires	7
Dossier « IA et Diagnostic »	20
Revue bibliographique	41
Compte rendu de conférences	42
Sommaire des revues	46
Résumés habilitations et thèses	54
Appels à dossiers	62

CALENDRIER DE PARUTION DU BULLETIN DE L'AFIA

<p><i>Hiver</i> Réception des contributions: 15 décembre Sortie le 31 janvier</p>	<p><i>Eté</i> Réception des contributions: 15 juin Sortie 31 juillet</p>
<p><i>Printemps</i> Réception des contributions: 15 mars Sortie le 30 avril</p>	<p><i>Automne</i> Réception des contributions: 30 septembre Sortie le 31 octobre</p>