BULLETIN DE L'AFIA

afia

ASSOCIATION FRANCAISE POUR L'INTELLIGENCE ARTIFICIELLE

JUILLET 1999

N°38

Association Française pour l'Intelligence Artificielle

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 reflètent le point de vue de leurs auteurs et n'engagent qu'eux-mêmes.

Pour contacter l'AFIA

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

Contributions au bulletin
Gérard SABAH
voir ci-contre

Serveur WEB
http://www.afia.polytechnique.fr/

Adhésions,
Liens avec les adhérents
René QUINIOU
INRIA/IRISA
Campus de Beaulieu
35042 Rennes CEDEX
Mél. : quiniou@irisa.fr

Parrainage de manifestations
Marc AYEL
LIA-ESIGEC- Technolac
Université de Savoie
73376 Le Bourget du Lac CEDEX
Mél. : Marc.Ayel@univ-savoie.fr

Membres d'honneur

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

Personnes morales adhérentes à l'AFIA

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

Bureau de l'AFIA

Bertrand BRAUNSCHWEIG, président
François ARLABOSSE, trésorier
Serge DUPUY, secrétaire

Marc AYEL, Jean-Paul BARTHÈS, Jean CHARLET, Patrick GALLINARI, Catherine GARBAY, Philippe LAUBLET, Eunika MERCIER-LAURENT, Christine PIERRET, René QUINIOU, Gérard SABAH, Christian de SAINTE-MARIE, Laurent SIKLÓSSY, Marc SCHOENAUER, Brigitte TROUSSE

Comité de rédaction

Gérard SABAH
Rédacteur en chef
LIMSI — CNRS
B.P. 133, 91403 ORSAY Cedex
gs@limsi.fr

Jean CHARLET
Rubrique « Conférences et comptes rendus »
INSERM U194 — DIAM
91 boulevard de l'Hôpital
75634 Paris Cedex 13
charlet@biomath.jussieu.fr

Patrick GALLINARI
Rubrique « Groupes de travail et collèges »
LIP6, Université Paris 6
8, rue du Capitaine Scott
75015 Paris
Patrick.Gallinari@lip6.fr

Sylvain GIROUX
Rubrique « Québec »
UER Science et Technologie — LICEF/Télé-université,
1001, rue Sherbrooke est, bureau 4232,
C.P. 670, Succ. C, Montréal, P.Q.,
Canada, H2L 4L5
sgiroux@teluq.uquebec.ca

Brigitte GRAU
Rubrique « Sommaires des revues »
LIMSI — CNRS
B.P. 133, 91403 ORSAY Cedex
grau@limsi.fr

Gilles KASSEL
Rubriques « Présentation de laboratoires » et « Présentation de sociétés »
Universite de Picardie Jules Verne — LARIA
5, rue du Moulin neuf
80000 — Amiens
kassel@laria.u-picardie.fr

Philippe LAUBLET
Rubrique « livres »
CAMS — Université Paris 4
96 boulevard Raspail, 75006 PARIS
Philippe.Laublet@lip6.fr

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

Serge STINCKWICH
Rubrique « IA et Internet »
GREYC — Université de Caen
Esplanade de la Paix
14032 CAEN CEDEX
Serge.Stinckwich@info.unicaen.fr

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

afia ASSOCIATION FRANCAISE POUR L'INTELLIGENCE ARTIFICIELLE

Réa. D.E.B. Copy – 01 46 31 64 53
ISSN 1273-1323
Dépot légal août 1999

EDITORIAL

2000 : Une année de transition pour le bulletin

L'assemblée générale du 16 juin a voté les tarifs de cotisation à l'association pour l'an 2000. Ce vote se place dans un contexte de déficit de l'AFIA depuis deux ans. En effet, comme je l'avais souligné lors de l'AG de décembre dernier, le tarif d'adhésion ordinaire ne couvre pas, dans la situation actuelle, les frais de composition, d'impression et de distribution des quatre numéros du bulletin[1]. Par ailleurs, la disponibilité pratiquement universelle d'Internet nous amène naturellement à reconsidérer l'ensemble des moyens de diffusion d'information de l'association, et plus particulièrement le bulletin papier, le bulletin électronique, et le site WEB.

Un comité *« bulletins »*, animé par Gérard Sabah, avait donc été chargé d'examiner les solutions possibles pour rétablir la balance cotisations/bulletins, et notamment en envisageant des modes de diffusion partielle du bulletin par le WEB.

L'AG a finalement choisi parmi trois solutions possibles :

1. maintenir les tarifs au niveau 1999, accepter un déficit prévisionnel sur l'an 2000, tout en continuant de réfléchir sur les moyens de faire baisser les coûts du bulletin, notamment au moyen du report sur le WEB d'une partie des informations contenues aujourd'hui dans le bulletin papier ;
2. augmenter le tarif d'adhésion individuel à 350 F, afin de rétablir l'équilibre, tout en maintenant le système actuel ;
3. offrir aux adhérents le choix du mode d'accès au bulletin, papier ou WEB, avec des tarifs d'adhésion différenciés.

Après une longue discussion qui a montré des avis très divers sur le sujet, et deux tours de vote, l'assemblée a adopté la troisième proposition, qui se résume comme ceci : 400 F/200 F (normal/étudiant) pour les personnes physiques payant des cotisations incluant le bulletin papier ; 200 F/100 F pour les cotisations ne donnant accès au bulletin que sur le WEB. Ceux qui garderont la formule papier paieront un peu plus cher, tandis que ceux qui prendront la formule WEB feront une économie et recevront le bulletin un peu plus tôt, non broché, il est vrai.

Je ne souhaite pas discuter longuement ici des avantages et inconvénients des solutions envisagées, et vous laisse vous faire votre propre opinion. Je veux cependant insister sur le fait qu'il s'agit d'une mesure transitoire, facile à mettre en place dès l'an prochain, permettant d'inverser rapidement la tendance. Elle contient cependant plusieurs incertitudes : le nombre d'entre vous qui choisiront l'une ou l'autre formule ; le coût d'impression pour des quantités inférieures ; l'impact des nouveaux tarifs sur les adhésions… Il est donc essentiel que Gérard Sabah et son comité continuent à réfléchir afin de proposer dès que possible, une solution qui recevrait un accueil favorable de tous les adhérents, permettant de revoir les tarifs à la baisse, et que nous nous efforcerons alors d'appliquer dans les meilleurs délais. En attendant nous considérons 2000 comme une année de test.

En début d'année prochaine, ceux qui choisiront l'abonnement sans papier recevront un *« login »* et un mot de passe sur le serveur WEB, ce qui leur donnera accès au bulletin sous forme de fichiers pdf. Il est évident qu'en échange, ils s'engageront moralement à ne pas communiquer ces codes, ni à pratiquer la reproduction en masse, même si cela peut sembler plus économique… À terme, afin d'éviter ce type de problèmes, nous proposerons sans doute des formules d'adhésion groupées, donnant plusieurs accès au site pour une somme forfaitaire.

Enfin, ceux qui, comme moi, restent particulièrement attachés au bulletin papier, doivent comprendre que nous sommes dans un processus évolutif ; il faut donc que chacun envisage cette phase transitoire sous ses aspects les plus positifs. C'est indispensable pour que l'AFIA passe une bonne fin d'année 1999 et une bonne année 2000.

J'aurais également aimé vous parler d'un autre sujet très important, l'ASTI, mais le temps et l'espace me manquent pour le faire dans ce numéro. J'essaierai de faire passer l'information par le WEB dans les semaines qui viennent, et cela sera repris dans le bulletin 39, en même temps que les comptes rendus de la plate-forme, qui se tient à l'heure où j'écris ces lignes.

Bertrand Braunschweig

1. Quatre numéros par an nous reviennent à 300 F par exemplaire et par adhérent, chiffre à comparer au tarif d'adhésion individuel ordinaire, fixé à 290 F cette année. Sachant que parmi nos 370 adhérents figurent 1/3 d'étudiants qui ne payent que 100 F, le déficit annuel dépasse les 17 kF, soit environ 50 F par adhérent…

Concours AFIA 1999 « Agent personnel Intelligent »

À l'occasion de son 10e anniversaire, l'AFIA lance un concours consistant à réaliser le ou les meilleurs agents intelligents permettant de rendre des services personnalisés (par exemple avertir des manifestations intéressentes, rappeler la date de paiement d'une cotisation, envoyer des sections choisies du bulletin, fabriquer une page web sur mesure, etc.).

Les agents des lauréats de ce concours seront rendus disponibles sur le site WEB de l'AFIA.

Les soumissions devront parvenir avant le : ***29 février 2000,*** au président du jury à l'adresse suivante :

Jean-Pierre Müller
IIUN – Université de Neuchâtel
11, rue Emile Argand
CH-2007 Neuchâtel – Suisse

Le règlement du concours qui devra être joint daté et signé lors de la soumission peut-être obtenu sur le site web : ***http://www.afia.polytechnique.fr,*** ou auprès du secrétaire de l'AFIA : Serge Dupuy, 75 Bd. Richard Lenoir, 75011 Paris, Serge.dupuy@wanadoo.fr

Règlement du concours
« Agent Personnel Intelligent »

Article 1 : Ce concours est organisé par l'Association Française pour l'Intelligence Artificielle (AFIA).

Article 2 : Ce concours est ouvert à tous à l'exclusion des membres du jury, de leur famille et de leurs collaborateurs.

Article 3 : Seuls les œuvres logicielles originales, développées pour le concours, seront admises.

Article 4 : Les participants garantissent aux organisateurs que tous les logiciels sont libres de droit et ne lèsent en aucune manière les droits de quelque personne physique et morale que ce soit.

Article 5 : Le concours porte sur la réalisation d'agents personnels intelligents sur le web pour les membres de l'association. L'idée est que chaque membre dispose d'un agent logiciel qui lui procure des services personnalisés (par exemple l'avertir des manifestations qui l'intéressent, lui rappeler qu'il doit payer sa cotisation, lui envoyer les sections du bulletin qu'il préfère, lui fabriquer une page web sur mesure). Chaque membre pourrait configurer son/ses agents en fonction de ses intérêts. Les agents sont utilisables à partir d'un client web standard tel que Netscape Communicator ou Internet Explorer.

Article 6 : la soumission des logiciels se fera par envoi au président du jury d'une disquette, ou d'un CD-ROM contenant le logiciel. Les soumissions électroniques sont possibles sous forme de mél contenant un fichier attaché, ou donnant l'URL à laquelle il est possible de se procurer l'agent. Les candidats sont priés de fournir toutes les explications techniques nécessaires à l'installation et à la configuration de leurs logiciels. En particulier toutes les spécificités d'utilisation (versions de Netscape ou d'Explorer, présence de plug-ins, système Unix, MacOS ou Windows etc.) devront être fournies. Manuel d'utilisation ou aide en ligne sont les bienvenus. Le jury appréciera également la remise d'un texte de présentation des techniques mises en œuvre. Toute soumission devra être accompagnée du présent règlement daté et signé.

Article 7 : Aucun frais ne sera remboursé pour l'envoi des logiciels.

Article 8 : L'AFIA ne pourra être rendue responsable des retards, pertes, détériorations ou destructions des envois, causés par les services postaux, ou autres.

Article 9 : La date limite d'envoi des logiciels est le 29 février 2000 à minuit, cachet de la poste (disquette ou CR-ROM) ou date du mél d'envoi, pour un envoi électronique, faisant foi.

Article 10 : Les logiciels seront soumis à un jury nommé par le bureau de l'AFIA dont les décisions seront sans appel.

Article 11 : Les logiciels seront jugés par le jury selon les critères suivants :

- *Originalité des idées ;*
- *Services rendus ;*
- *Qualité de la réalisation et de la documentation éventuelle ;*
- *Ergonomie, utilisabilité, capacité de configuration personnalisée ;*
- *Coup de cœur.*

Article 12 : Le concours est doté de 10 000 F de prix. La répartition de ces prix aux lauréats sera établie par le jury, les ex aequo étant possibles. De plus, les trois meilleurs logiciels présentés seront mis en service sur le web de l'AFIA, et feront l'objet de la publication d'un article dans le Bulletin de l'AFIA. La remise des prix se fera au cours d'une conférence organisée par l'AFIA. Les lauréats ne pourront demander de remboursement de frais de déplacement correspondant à la participation à la cérémonie.

Article 13 : Les participants autorisent les adhérents de l'AFIA à utiliser gracieusement leurs logiciels à des fins non commerciales, en particulier pour les services rendus déclarés pour le concours.

Article 14 : L'AFIA ne saurait être tenue pour responsable si, par suite de force majeure, ce concours devait être modifié, reporté ou annulé.

Article 15 : Tout renseignement concernant ce règlement pourra être demandé à : Serge Dupuy, **Secrétaire Général, Bureau de l'Association Française pour l'Intelligence Artificielle,** 75 Bd. Richard Lenoir, 75011 Paris (Serge.dupuy@wanadoo.fr)

Article 16 : La participation au concours implique l'acceptation sans réserve du présent règlement, sans aucune possibilité de réclamations.

Article 17 : Le présent règlement est déposé chez le secrétaire de l'AFIA.

Je certifie que j'ai pris connaissance du présent règlement.

Fait à……, le……

Signature

A LA MÉMOIRE DE JACQUES LEBBE

Jacques Lebbe nous a quittés prématurément le 6 juin 1999.

Il était professeur au laboratoire d'informatique et de systématique de l'université de Paris VI à Jussieu. À la fois biologiste, systématicien, - il a été le président de la société francophone de systématique - informaticien, statisticien, médecin, musicien, il aura été un génial touche-à-tout. Il avait soutenu en 1991 à l'université Pierre et Marie Curie une thèse préparée sous la direction de E. Diday et de G. Cusset, dont le titre indique un certain nombre de ses domaines de compétence : *« Représentation des concepts en biologie et médecine (Introduction à l'analyse des connaissances et à l'identification assistée par ordinateur) »*. L'analyse de connaissances préfigure ce qui allait devenir la représentation et l'extraction de connaissances dans des bases de données et l'identification assistée par ordinateur lui avait permis de se distinguer, au milieu des années quatre-vingt, dans le monde de l'intelligence artificielle et de l'analyse de données.

Par sa gentillesse et sa culture, Jacques Lebbe savait communiquer ses passions et il nous a offert un vaste terrain de réflexion et d'expérimentation sur les bases de connaissances qui, loin des canaris et des pingouins traditionnels, s'attaque à des sujets au cœur de la représentation et de l'interrogation en systématique, dans des domaines aussi variés que les phlébotomes ou les hépatites virales.

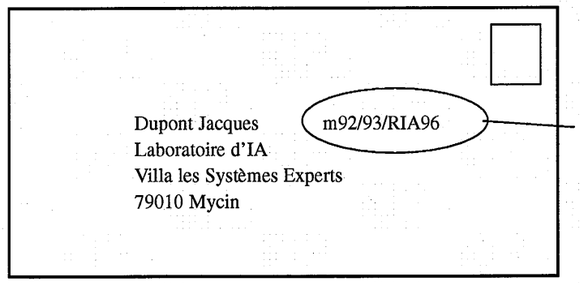
Avec sa femme et complice, Régine Vignes, qui soutenait une thèse sur le même thème le même jour, ils avaient mis au point Makey et Xper, un logiciel d'identification très en avance sur son temps. Initialement testé sur les champignons, ce logiciel accessible par Minitel a été utilisé par les pharmaciens pendant plusieurs années.

Un projet ambitieux lui tenait particulièrement à cœur : il cherchait depuis trois ans à fédérer les efforts de la recherche nationale en biosystématique et informatique, dans l'espoir de répondre au challenge international de la sauvegarde et de l'analyse massive des connaissances sur la biodiversité.

Jacques Lebbe appartenait au cercle restreint des chercheurs œcuméniques, ceux que l'on qualifiait autrefois de savants. Il s'était fait de nombreux amis dans les communautés apprentissage et classification dans lesquelles il a laissé l'image d'un chercheur passionné.

LE COIN DU TRÉSORIER

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



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

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

TIA'99
Troisièmes rencontres « Terminologies et Intelligence Artificielle »

Nantes, 10-11 mai 1999

Comme les précédentes (en 1995 et 1997), les rencontres de Nantes, organisées à l'initiative du groupe « Terminologie et intelligence Artificielle » ont connu un vif succès. Plus de quatre-vingts personnes ont participé aux deux journées dont plus du tiers d'informaticiens et près du cinquième de professionnels (responsables de service en ingénierie linguistique ou en documentation, terminologues…). La présence de nombreux étudiants est le signe d'un intérêt croissant pour des problématiques émergentes. On note cependant une contradiction entre un public majoritairement linguiste ou/et terminologue alors que les intervenants sont quasiment tous informaticiens. Plusieurs éléments peuvent expliquer ce phénomène. Tout d'abord, il se peut que les travaux de linguistes/terminologues soient moins avancés que ceux des informaticiens, sur un sujet où tout est à faire : renouveler la réflexion sur la terminologie et mettre en place une linguistique de corpus. Ensuite, le terme « intelligence artificielle » fait peur et éloigne des travaux essentiellement linguistiques qui auraient pourtant toute leur place dans ces rencontres. Enfin, le besoin d'information et de formation est patent pour les terminologues ou même les linguistes qui sont directement confrontés aux demandes des entreprises.

Sans atteindre celui des linguistes/terminologues, le nombre des chercheurs en IA a toutefois augmenté cette année, conséquence à la fois d'une prise de conscience de l'importance de la problématique pour la discipline (le nombre élevé de thèses qui devraient être soutenues prochainement est un autre indice de cette prise de conscience) et du fait que la conférence était organisée par une équipe d'informaticiens (l'équipe Langage Naturel de l'Institut de Recherche en Informatique de Nantes). Les nombreuses questions qui ont suivi chaque exposé témoignent du dynamisme de la communauté et de son désir de participer aux débats.

Ces troisièmes rencontres ont reçu le soutien de l'AFIA, de la Direction Générale à la Langue Française (DGLF), du Ministère de l'Éducation Nationale et de la Recherche et de la Technologie, et du Réseau International de Néologie et de Terminologie (Rint), qui a accepté de publier les actes dans sa revue Terminologie Nouvelles.

Les deux jours ont été organisés de la manière suivante : un tutoriel présentant la terminologie textuelle, deux conférences invitées, des présentations sélectionnées par un Comité de programme et des démonstrations d'outils (six outils étaient présentés).

Le tutoriel détaillé et très argumenté de Didier Bourigault et Monique Slodzian, par lequel les rencontres ont débuté, a permis de présenter l'évolution inéluctable de la terminologie telle qu'elle est perçue par le groupe TIA et de tracer les grandes lignes de la « terminologie textuelle » pouvant remplacer la terminologie traditionnelle,.

Asuncion Gomez Perez, de l'Universidad Politecnica de Madrid. A. Gomez Perez, première conférencière invitée, a fait le tour des méthodes et des projets concernant les ontologies. Elle a ainsi dressé un état des travaux les plus nombreux, qui, malheureusement, prennent rarement en compte la dimension textuelle des connaissances représentées.

En écho au tutoriel, la deuxième conférencière invitée, Jennifer Pearson, de l'Université de Dublin a insisté sur le rôle des corpus en terminologie et sur la nécessité de définir des critères de constructions et des méthodes d'exploration des données textuelles.

Dix communications ont été sélectionnées par le Comité de programme. Elles s'organisent en quatre rubriques.

1 - Présentation d'outils de gestion ou/et d'aide à la constitution de terminologies : trois articles peuvent être rassemblés sous cette rubrique : celui de A. Hocine et K.M. Brou (« Co-DB-Web : un système de représentation et de gestion hypertexte de concepts »), celui de C. Roche, J.-C. Marty et S. Lacroix (« Ontologie et Terminologie : le modèle OK » et celui de N. Aussenac (« Gediterm : Un logiciel pour gérer les Bases de Connaissances terminologiques »).
2 - Description de méthodes pour acquérir des données terminologiques : article de T. Hamon, D. Garcia, A. Nazarenko (« Détection de liens de synonymie : complémentarité des ressources générales et spécialisées ») et article de P. Séguéla (« Adaptation semi-automatique d'une base de marqueurs de relations sémantiques sur des corpus spécialisés »).
3 - Comptes rendus d'expériences sur la terminologie (remarquons qu'elles utilisaient toutes le logiciel Lexter, de Didier Bourigault) : article de Y.Abbas et M.L. Piccard (« Exemple de pratique terminographique en entreprise »), article de D. Bourigault, C. Chodkiewicz et John Humbley (« Construction d'un lexique bilingue des droits de l'homme à partir de l'analyse automatique d'un corpus aligné »), article de A. Chrobot (« Extraction terminologique en anglais basé sur des dictionnaires généraux et spécialisés »).
4 - Utilisation de méthodes de terminologie textuelle pour la veille terminologique : deux articles sont concernés, celui de B. Goujon (« Extractions d'informations techniques par l'exploitation de notions indépendantes d'un domaine pour la veille ») et celui de T. Poibeau (« Le repérage des entités nommées : un enjeu pour la veille technologique »).

Anne Condamines,
Présidente du Comité de programme de TIA'99

Vers une rubrique « poésie de l'intelligence artificielle » ?

CYBERCOUP DE FOUDRE

Je vais vous raconter l'histoire de Robert
Un accro d'Internet qui, été comme hiver,
Surfait sur le réseau traquant la moindre info
Qu'il repassait ensuite à deux ou trois journaux.
Car dès potron-minet, les yeux à peine ouverts
Il tournait le bouton, à peine réveillé
De son ordinateur et de la cafetière
Et il vérifiait s'il avait du courrier
Qui pendant son sommeil lui serait parvenu
De ses vieux copains ou de parfaits inconnus.
Un soir qu'il visitait un site en Australie
Il fut interrompu par une certaine Lily
Qui ne supportait plus solitude et ennui
Qui cherchait l'âme sœur au milieu de la nuit
Qui aimait les calembours, les contrepèteries
Les charades à tiroirs mais surtout les poèmes
Les vers de douze pieds étaient ses préférés
Par leur balancement et leur aspect altier
Mais elle ne crachait pas non plus
Sur les vers octosyllabiques
Tout ce que provoquait les mots et leur musique
Tout ce qui apaisait ses états bucoliques
À ces mots l'internaute ne se tient plus de joie
Diable ! une meuf poète et des propos hardis
Incroyable ! Tout simplement inespéré !
Quoi de plus exaltant, il n'aurait pas mieux dit
Depuis qu'il bourlinguait, c'est la première fois
Qu'il rencontrait quelqu'une dont les espérances
Faisaient enfin écho à ses jours de souffrance
Aux jours de vague à l'âme où il perdait la foi.
Déjà il lui répond en fiers alexandrins
Son bonheur sans égal de rencontrer enfin
Un être humain qui à l'autre bout de la terre
Communiait avec lui d'façon épistolaire.
Il lui dit ses espoirs en un monde meilleur
Où tout serait facile, même trouver le bonheur.
Ils conversaient encore quand le jour les surprit
Car les mots fleurissaient de leurs mains par milliers
Les claviers transmettaient et leurs cœurs exultaient
Nul ne pourra vous dire quand Robert et Lily
S'arrêtèrent enfin après forces promesses
De se retrouver vite pour des cybercaresses.

Alain Bonnet

Un débat c'est l'examen d'un problème entraînant une discussion animée, parfois dirigée, entre personnes d'avis différents (Larousse)...

Nouveau débat : « L'IA ne manque pas d'avenir ».

S'il y a un débat qui dépasse le cadre de notre communauté c'est bien celui qui concerne le début du prochain millénaire. Non, le 1er janvier de l'an 2000 ne marquera pas le premier jour du troisième millénaire. De fait quand on compte une dizaine, on part de 1 et l'on s'arrête à 10, la deuxième dizaine commence à... 11. Mais sans s'étendre sur cette question qui a déjà fait l'objet de nombre de rubriques dans vos quotidiens favoris (voir par exemple http://www.liberation.fr/ quotidien/debats/septembre98/980910d.html), nous aimerions lancer dans ce numéro un débat sur l'**avenir de l'IA.** Il ne s'agit pas de savoir si les programmes produits par la communauté résisteront aux bogues du millénaire — nous sommes confiants — mais plutôt de s'interroger sur l'IA dans les dix années à venir, tant sur ses thèmes de recherches que sa visibilité et ses applications industrielles. Les organisateurs de la conférence ISAI'99 soulignaient récemment lors de la plate-forme AFIA que *« l'IA est de plus en plus présente dans les applications informatiques mais trop souvent non visible »*. N'avons-nous pas — pour utiliser une formule à la mode — un problème de communication en tant que communauté ? On dit parfois que ce qui est issu de l'IA et qui est reconnu n'est plus de l'IA. Serions-nous ainsi condamnés à n'être que la jeunesse de domaines qui évolueraient ensuite par eux-mêmes ? Récemment, AI magazine titrait sur l'entrée de l'IA dans l'adolescence (AI Reaches Adolescence, Vol 19, N° 4, Winter 1998, James Allen). L'auteur y propose une définition de l'IA sujette à controverse : *« AI is the science of making machines do tasks that humans can do or try to do »*. Vu cette définition, où le mot science côtoie celui de « machines », où ladite science a pour objet de faire ce que l'être humain peut faire ou veut faire, on risque fort d'une part d'englober toutes les applications informatiques et d'autre part de perdre toute crédibilité vis-à-vis des autres sciences pour avoir confondu technique et science.

Les questions sur lesquelles nous vous proposons de réagir sont donc : Quels sont les prochains grands défis de l'IA ? quels sont les problèmes majeurs auxquels il faut nous attaquer, et plus généralement où va l'IA ?

Comme pour chaque débat, vous pouvez y participer en envoyant un petit texte à **Jean-Daniel.Zucker@lip6.fr.**

« Évaluation Scientifique : peut-on mieux faire en IA ? ».

Voir l'annonce complète dans le bulletin n° 35.

En réponse aux questions de Jean-Daniel Zucker, et suite à la réaction de Jérôme Euzenat, Gérard Sabah reprend ici quelques éléments pertinents précédemment publiés dans TSI (Sabah Gérard 1996, *À propos des limites des mesures objectives...,* Technique et science informatique, 15, 2, p. 233-236.)

S'il est traditionnel de considérer les sciences exactes comme complètement objectives, il convient de souligner le paradoxe d'une telle attitude : quel que soit le comportement étudié, l'interprétation de toute observation demande obligatoirement l'intervention subjective d'un observateur introspectif ; ainsi, tous les résultats de la science, aussi précis et froids soient-ils, ne sont que les résultats *conscients* de telles *interprétations* : toute mesure, tout calcul produit par les sciences exactes ne nous est accessible que par l'action (et donc transformation) subjective de notre compréhension. (Sur ce point, on pourrait aisément développer des arguments plus techniques, fondés sur le principe d'incertitude d'Heisenberg qui montre que les observations affectent l'observé jusqu'à supprimer toute possibilité de prédiction...)

Ainsi, les principes fondamentaux des sciences dures ont été établis par des expériences précises, quantifiables et reproductibles : leurs outils de mesure tentent d'évacuer une trop grande complexité, en étudiant isolément les variables qui décrivent le domaine étudié ; elles ne mettent en évidence que des corrélations simples, manquant ainsi totalement tout ce qui est fondé sur les interactions complexes.

Par ailleurs, voulant dépasser une simple technologie informatique, on peut fixer comme but à l'intelligence artificielle de formaliser et construire des modèles de l'esprit humain. Vu le grand nombre de variables impliquées et l'impossibilité de reproduire d'une façon homogène les résultats d'expériences portant sur des individus particuliers, il est alors clair qu'on ne dispose pas d'une méthodologie directe pour extraire les invariants de l'esprit humain.

Vouloir mesurer les résultats d'une telle approche implique qu'on ne retient qu'une infime partie de la réalité observée : on tente de modéliser l'esprit comme un programme informatique, en le réduisant à sa capacité rationnelle ; on suppose l'autonomie de la pensée rationnelle, modélisée comme un calcul logique, niveau fonctionnel de l'esprit jugé suffisant pour rendre compte de l'intelligence.

Beaucoup de compétences cognitives humaines, comme le langage en particulier, sont trop complexes pour être décrites au moyen d'une programmation directe. Si, alors, on s'intéresse, comme moi, aux mécanismes subliminaires liés à la compréhension du langage, la plausibilité psychologique des composants d'un système complexe peut sans doute être étayée par des expériences scientifiques, mais leur plausibilité opérationnelle collective est très difficile à prévoir analytiquement : on ne peut que tenter de simuler le comportement du système complet et porter un jugement subjectif sur le comportement observé, en suivant une démarche typique des sciences de l'ingénieur.

Mais, comment exiger d'une telle simulation qu'elle fournisse une prédiction exacte d'un phénomène cognitif naturel, même précisément déterminé, alors que nous sommes incapables de le faire pour nous-mêmes ?...

Ainsi, les différences entre « sciences dures » et « sciences molles » résident donc principalement dans la différence entre mesures objectives et mesures subjectives. Il n'est certes rien de mauvais avec les sciences dures (elles font étonnamment bien ce qu'elles font), le problème porte sur ce qu'elles abandonnent (ce qu'elles ne peuvent ou ne savent pas mesurer !) et que les sciences « molles » considèrent comme pouvant être essentiel.

Débat : « IA et Médecine ».

Le nouveau débat proposé dans le numéro précédent concernait le lien entre l'IA et la médecine. Vincent Corruble nous donne ici son point de vue.

L'IA et son application à la Médecine : histoire et perspectives

La médecine joue depuis les années soixante-dix un rôle crucial pour l'Intelligence Artificielle (IA). D'ailleurs, exagérerait-on beaucoup en imaginant qu'un martien observant l'IA depuis sa lointaine planète pourrait classer ses applications en deux catégories : les applications jouets et les applications médicales ? Cette situation a bien entendu des causes historiques que nous rappelons brièvement. Une question peut-être plus importante que nous abordons aussi est de savoir si cette situation est justifiée actuellement en regard des objectifs des deux disciplines.

Dans les années soixante-dix, l'IA est en quête de reconnaissance en dehors de sa propre communauté universitaire. En particulier dans son incarnation californienne, à l'Université Stanford, on pense que les

méthodes heuristiques en général, les systèmes à base de connaissances en particulier, sont aptes à résoudre des problèmes du monde réel de façon égale ou complémentaire à l'homme. Comme la plupart des applications de l'informatique, cette idée peut être validée si l'on trouve des domaines où règne une certaine complexité : l'effet de levier de la puissance de calcul y est alors le plus intéressant. Pour que l'IA, en contraste avec d'autres méthodes informatiques plus traditionnelles, y soit justifiée, il faut que le problème à résoudre soit en plus difficile à formaliser, et qu'une solution exacte (non-heuristique…) ne vienne pas naturellement. Ces critères de taille et de complexité sont de notre point de vue les justifications historiques principales de l'intérêt porté à la médecine par l'IA. Les connaissances du domaine médical sont étendues, difficiles à modéliser parce que toujours partielles et susceptibles de remise en question. Ces propriétés du domaine médical expliquent la longueur et la complexité d'organisation des études médicales.

L'intérêt originel pour les SBC médicaux s'est depuis élargi à d'autres techniques. De nouvelles disciplines se sont même développées, en particulier aux États-Unis, qui se situent à la frontière de l'informatique et de la médecine *(Medical Informatics),* avec des références à l'IA qui restent souvent vagues. Des techniques récentes, le data-mining, ou fouille de données, a trouvé dans la médecine un domaine d'application autre que l'industrie, même si l'ampleur des résultats pratiques y reste encore modeste. Un autre développement récent est l'intérêt pour le raisonnement temporel, aussi bien pour le suivi de patient, le diagnostic, ou même la découverte de connaissances. Aussi, la modélisation et la simulation, techniques qui ne se trouvent qu'en partie dans l'IA, ont trouvé en médecine de nouvelles applications, comme en témoignent les thèmes du *World Congress on Computational Medicine.* En médecine comme dans les autres applications, l'IA souffre du fait que les techniques qui obtiennent le plus de succès viennent souvent des sous-communautés qui oublient leurs origines.

Il est aussi intéressant de regarder les évolutions récentes de la médecine. Les dernières années ont vu se développer, surtout dans les pays anglo-saxons, la mode des protocoles qui détaillent la conduite à tenir pour traiter un patient en fonction d'un diagnostic initial et de son évolution. Il est intéressant de remarquer que cette mode traduit deux soucis : d'une part de permettre à un plus grand nombre de médecins (et donc de patients) d'utiliser les résultats des recherches les plus récentes ; d'autre part, de répondre au souci de rationaliser davantage la médecine, ce qui permet de contrôler les coûts.

Pour l'IA, cette évolution n'est peut-être pas une si bonne nouvelle. L'existence de protocole (dont beaucoup sont déjà informatisés) facilitera sans aucun doute le développement de l'informatique médicale. Mais en figeant ainsi la connaissance (qu'en est-il, par exemple, des diagnostics multiples et des patients atypiques ?), ils présagent d'une médecine uniformisée qui ne répondrait plus vraiment aux critères de taille/complexité qui en faisait jusqu'ici un domaine si intéressant pour l'IA. Ce risque est apparu dans les discussions de la dernière conférence AIMDM (Artificial Intelligence in Medicine and Medical Decision Making, Aalborg, juin 1999) où une tension s'est fait jour entre les deux communautés IA et Medical Decision Making : la première s'intéresse généralement au développement ou à l'application de nouvelles méthodes pour une hypothétique médecine du futur, alors que la deuxième tend à ne voir qu'un intérêt à court terme, mesuré par le degré d'utilisation dans l'hôpital et les perspectives offertes pour une baisse des coûts, souvent indépendamment de la santé du malade.

En somme, il nous semble que l'IA s'est à l'origine penchée sur la médecine en espérant donner aux machines cette étendue et cette flexibilité des connaissances qui caractérisent l'expertise médicale, tandis que l'évolution récente tend à faire évoluer la médecine vers une science plus rationnelle, plus proche de la machine. En allant ainsi l'une vers l'autre, il est bien possible qu'IA et médecine finissent par se rencontrer. À moins, et ce serait ironique, qu'elles ne se croisent sans même se reconnaître…

XML et la recherche d'informations

Serge Stinckwich
(Serge. Stinckwich@info.unicaen.fr)

1. Introduction

Il est clair, que le WEB a été conçu à l'origine essentiellement comme un hypertexte dédié à un parcours par des individus. Sa taille toujours plus grande ne permet plus de s'y retrouver. Sa nature distribuée et faiblement structurée est néanmoins un frein important à l'automatisation d'activités de recherche et de collecte d'informations.

Les nombreuses limitations du WEB sont d'abord dues à l'inadéquation du langage de marquage HTML (Hyper Text Markup Language), y compris dans ses versions les plus récentes :

- les balises HTML sont prédéfinies. L'utilisateur ne peut pas définir son propre jeu de balises spécifiques à son champ disciplinaire ou ses applications ou encore modifier la sémantique des balises d'HTML,
- il n'existe pas de distinction claire entre la structure physique ou de présentation du document et la structure logique. HTML 4 accompagné des feuilles de styles CSS (Cascaded Style Sheet) sont une tentative encore timide dans cette direction,
- il n'y a pas de système de nommage logique. Les URIs pallient cette insuffisance,
- enfin la structure et la sémantique des liens sont très pauvres.

Ces limites ont engendré de nombreuses activités de recherche notamment en intelligence artificielle, mais celles-ci restent vaines si les différents outils ne peuvent pas interopérer de manière distribuée et souple. XML (eXtensible Markup Language) est la tentative la plus intéressante dans ce sens. Il s'agit d'un langage de définition de balises autorisant la définition de documents structurés. Il s'affranchit de plusieurs des limites de HTML. Lancé en 1998 par le consortium World Wide WEB, il s'agit d'une adaption de SGML (Standard Generic Markup Language) aux contraintes d'Internet.

XML distingue deux types de document : les documents *bien formés,* respectant les règles syntaxiques de la spécification XML, et les documents *valides,* documents bien formés, qui respectent une DTD (Définition de Type de Document). Une DTD définit une classe de documents spécifiques au moyen d'un ensemble de balises, appelés éléments, ainsi que la structure d'enchâssement correspondante. On consultera avec profit les sites indiqués ci-dessous afin de trouver une liste d'outils (analyseurs, éditeurs et bases de données XML).

De très nombreuses DTD spécifiques sont en cours de développement. Parmi celles-ci, nous pouvons citer : les feuilles de styles XSL (eXtensible StyleSheets Language), les documents multimédias SMIL (Synchronized Multimedia Integration Language), les documents mathématiques MathML (Mathematical Markup Language), les documents pédadogiques EDUML, AstroML…

Un document XML devient vraiment multi-usage et multi-forme, puisqu'il est non seulement une ressource textuelle, mais également une méta-donnée, c'est-à-dire qu'il contient des informations à propos du document lui-même. Les avantages d'XML sont nombreux, non seulement dans le cadre de l'informatique documentaire, mais plus généralement, pour l'ensemble des applications manipulant des données sur le WEB : les balises deviennent extensibles, les documents sont modulaires et réutilisables dans d'autres contextes, XML peut s'interfacer avec des sources externes hétérogènes (base de données, autres services réseaux…), HTML et XML peuvent coexister dans un premier temps.

2. La recherche d'informations dans le WEB

On est, à présent, face à une alternative quand on cherche à améliorer les possibilités de recherche d'information sur le WEB. La première possibilité consiste à utiliser des outils de recherche documentaires ou des outils linguistiques. La faible structuration des documents HTML ne permet pas d'avoir des résultats satisfaisants avec cette approche. La seconde cherche à associer des méta-informations à chacun des documents.

RDF (Ressource Description Framework) est une DTD XML qui permet de représenter des connaissances et des schémas (au sens des bases de données). Ces schémas sont des triplets : ressource, propriété, valeur. RDF peut être utilisé pour améliorer les performances des moteurs de recherche, pour décrire des relations entre documents ou à l'intérieur de documents, il peut être utilisé par des agents de recherche d'information pour faciliter le partage et l'échange de connaissances, pour exprimer les préférences d'un utilisateur. RDF est maintenant utilisé dans la redéfinition d'autres projets du consortium W3C comme PICS (Platform for Internet Content Selection) et P3P (Platform for Privacy Preferences).

RDF s'inspire à la fois des modèles de représentation de connaissances (systèmes de frame, réseaux sémantiques) et du modèle entité-relation des bases de données. Il est néanmoins plus simple que le standard de description de connaissances KIF (Knowledge Interchange Format ; http://logic.stan-

ford.edu/kif/kif.html) et propose divers outils :

- SIRPAC (SImple RDF Parser And Compiler) est un outil permet de déterminer les triplets (ressource, propriété, valeur) à partir d'une spécification RDF. Une version online sous la forme d'une appliquette Java permet de dessiner le graphe correspondant.
- SILRI (SImple Logic-based RDF Interpreter) est une base de données déductive écrite en Java et utilisant RDF. Cet outil a été développé dans le cadre du projet On2Broker.
- RDF est également utilisé pour la constitution de thésaurus (voir CERES : http://ceres.ca.gov/thesaurus/).

Une des sources d'inspiration de RDF a été Dublin Core, un modèle de description de ressources électroniques proposé par les spécialistes des bibliothèques numériques. Il se veut un modèle simple de remplacement des formats classiques de catalogage utilisés principalement dans les bibliothèques et les centres de documentation (MARC, UNIMARC). Néanmoins, il ne propose pas de modèle d'implémentation, mais il est tout à fait possible pour cela d'utiliser RDF. DC comprend notamment 15 méta-informations associées à un document : Titre, Auteur, Sujet, Date…

XQL (XML Query Language) est un langage de requêtes dédié aux documents XML, en cours de développement depuis l'atelier de travail QL'98. Plusieurs acteurs du marché proposent déjà leur propre version d'XQL intégré à leurs outils, mais l'absence de standard bien défini ne permet pas de coupler ces différents outils.

Des DTD spécifiques concernant l'intelligence artificielle mais non la recherche d'informations sont en cours de développement : pour l'échange de messages KQML entre agents : KQML-XML, pour la spécification d'agents, XML-A, pour la constitution d'ontologies, OML (Ontology Markup Language) et CKML (Conceptual Knowledge Markup Language).

La nouvelle génération de navigateur (WEB Mozilla 5.0 et IE5) gère les documents XML.

3. Références

Portail XML.COM http://www.xml. com. De nombreuses listes d'outils, de spécifications de DTDs, des articles…

La référence en français : Alain Michard, « XML, Langage et applications », Eyrolles, 1999, http://www.eyrolles.com/livres/michard/.

RDF-1, Resource Description Framework (RDF) Model and Syntax Specification, W3C Recommandation, http://www.w3.org/TR/REC-rdf-syntax

RDF-2, Resource Description Framework (RDF) Schema Specification, W3C W3C Proposed Recommandation, http://www. w3.org/TR/PR-rdf-schema/

La spécification XML annotée par un de ses concepteurs Tim Bray : http://www.xml.com/xml/pub/axml/axmlintro.html. Essentiel à la compréhension d'XML.

Exemple de base RDF de 35 Mo utilisé dans un site d'indexation de sites : http://www.dmoz.org/rdf.html

Foire aux questions pour XQL http://metalab.unc.edu/xql/.

Mozilla, le navigateur de netscape qui gère XML (http://www.mozilla.org).

SHOE (Simple HTML Ontology Extensions ; http://www.cs.umd.edu/projects/plus/SHOE/). Extension d'HTML permettant de modéliser des ontologies.

Grosof B., Labrou Y., « An Approach to using XML and a Rule-based Content Language with an Agent Communication Language », Proceedings of the IJCAI-99 Workshop on Agent Communication Languages (ACL-99), Stockholm, Sweden, Aug. 1, 1999, (http://www.research.ibm.com/people/g/grosof/paps/rc21491.ps).

HTML : http://www.w3.org/HTML

URI : http://purl.org
http://www.w3.org/XML
http://www.w3.org/AudioVideo/

PICS : http://www.w3.org/PICS/

P3P : http://www.w3.org/tr/wd-p3p10-syntax

RDF : http://www.w3.org/RDF/

Dublin Core : http://purl.org/DC

The Query Languages Workshop : http://www.w3.org/QL/QL98

XML-Agents : http://www.cybercom.net/~laforge/xmlagent.html

Ontology Markup Language : http://wave.eecs.wsu.edu/CKRMI/OML. html

Conceptual Knowledge Markup Language : http://wave.eecs.wsu.edu/CKRMI/ CKML. html

PRÉSENTATION DU PÔLE INTELLIGENCE ARTIFICIELLE DU LABORATOIRE D'INFORMATIQUE DE PARIS 6 (LIP6)

http://www.lip6.fr/

Responsable P. Gallinari

(N.D.L.R. : nous donnons ici la suite de la description du pôle IA du LIP6 parue dans le précédent numéro du bulletin où les thèmes apprentissage et acquisition des connaissances -APA- (responsable B. Bouchon-Meunier) et objets et agents pour les systèmes d'information et la simulation -OASIS- *(responsable J.-P. Briot) étaient présentés. Le thème* systèmes d'aide à la décision et à la formation -SYSDEF- *(responsable J.-Y. Jaffray) est présenté ici, ainsi que les projets transversaux.)*

Thème de recherche Systèmes d'Aide à la Décision et à la Formation (SYSDEF)

Présentation du thème

Le thème rassemble des chercheurs issus de différentes disciplines et dispose d'un éventail de compétences le mettant en mesure de traiter, de la recherche théorique aux applications industrielles, des problèmes complexes de décision et de formation. L'orientation principale de ses travaux est la conception de systèmes intelligents hautement interactifs.

Trois actions de recherche structurent les activités du thème : optimisation et heuristiques, aide à la décision, systèmes d'aide à la formation.

Groupe Optimisation et heuristiques (responsable : Philippe Chrétienne)

Dans ce groupe se retrouvent des chercheurs issus les uns de l'IA, les autres de la RO, avec l'objectif commun de fournir des méthodes de résolution approchées à certains problèmes « difficiles » en optimisation combinatoire : problèmes d'ordonnancement, d'affectation, de satisfiabilité ou de satisfaction de contraintes, dont beaucoup proviennent directement du monde industriel (CNET, SNECMA). La théorie de la complexité fournit des outils puissants permettant de classer les problèmes selon leur difficulté et les algorithmes selon leur efficacité. Les méthodes de l'IA conduisent à la découverte d'heuristiques fournissant des solutions approchées. Enfin, une approche probabiliste peut permettre de prédire l'existence de solutions et d'identifier les cas favorables à une résolution effective servant ainsi de guide dans la recherche d'algorithmes efficaces. Toutes ces techniques sont mises en œuvre dans le cadre de l'action et appliquées à des problèmes combinatoires. Les principaux axes sont présentés dans les trois paragraphes ci-dessous.

Les problèmes d'ordonnancement concernent les ordonnancements avec communications (où, les tâches étant exécutées par un réseau de processeurs, aux contraintes classiques viennent s'ajouter les délais de communication) et les ordonnancements cycliques (rencontrés aussi bien en architectures parallèles qu'en ateliers flexibles ; ils sont caractérisés par l'existence de tâches génériques exécutées une infinité de fois et où l'objectif est de maximiser la fréquence d'exécution des tâches). Dans les deux cas, des résultats de complexité ont été démontrés et des algorithmes approchés ont été construits avec garantie de performance. D'autres problèmes ont été étudiés : le problème du *Bin-Packing* en dimension 2, le classique problème de Steiner ou encore le problème de l'ordonnancement sous la forme du *job shop* généralisé.

Les algorithmes intégrant IA et RO, à base d'heuristiques et de méta-heuristiques « dynamiques » sont appliqués à l'ordonnancement mais aussi aux problèmes d'affectation. L'algorithme général HOLSA testé sur le problème du sac à dos multidimensionnel a donné des résultats très probants (en temps et en valeur trouvée) par rapport à des algorithmes de recherche locale comme TABOU ou le Recuit Simulé. Nous travaillons également sur les problèmes de taquin à grande dimension.

L'étude du problème SAT par des méthodes exactes efficaces et des méthodes probabilistes qui permettent de réaliser des sauts de complexité est une approche à la fois expérimentale et théorique. Les chercheurs concernés essayant de se servir de l'une pour étayer l'autre. La méthode développée a actuellement les meilleures performances et a montré son efficacité sur d'autres problèmes tels que la reconstruction d'image à partir de projections, la construction de polyminos. Un modèle générique, Hyper-Implicant, a permis de proposer une méthode de compilation de bases de connaissances donnant un gain de temps exponentiel dans l'interrogation de ces bases.

Groupe Aide à la décision (responsable : Jean-Yves Jaffray)

Les recherches menées au sein de l'action *aide à la décision* procèdent du souci de représenter correctement en machine des situations décisionnelles et d'y apporter des solutions satisfaisantes. Les travaux partent du principe que la représentation des préférences et des croyances comme les processus de décision humains doivent effectuer un compromis entre idéaux normatifs et réalisme cognitif. Le but est donc la réalisation d'outils aussi bien théoriques (modèle de la coopération, mécanismes contextuels) que logiciels (configurateur de connaissances, systèmes d'aide à la décision). De même, les études théoriques en décision multi-critère ou dans l'incertain sont à la base de réalisations, en cours, de logiciels (réseaux probabilistes, diagrammes d'influence) et d'applications. Ces recherches sont largement ouvertes sur l'extérieur par le biais

de collaborations, nationales avec l'INRIA ou au sein de GDR-PRC du CNRS, internationales et dans l'industrie (RATP, SNCF). Les axes développés sont présentés dans les deux paragraphes suivants.

L'interaction entre l'utilisateur et la machine au cours de la résolution d'une tâche.

L'élément central d'un système d'assistance intelligent est la coopération homme-machine à travers l'intégration des explications, du dialogue, de l'acquisition incrémentale de connaissances et de la prise en compte explicite du contexte de la coopération. Nous travaillons sur la notion de contexte pour les Systèmes d'Assistance Intelligents fondés sur le Contexte (SAIC), sa modélisation et sa représentation en machine, et sur l'architecture générale de ces systèmes. Ces recherches font l'objet de plusieurs applications en cours pour la RATP et en œnologie. Par ailleurs, analysant les obstacles à l'utilisation des systèmes d'aide à la décision et des systèmes coopératifs, ils confirment la nécessité de maîtriser le contexte et d'étudier les flux réels de décision dans les organisations. Une autre direction de recherche concerne l'intégration, dans un modèle de dialogue, des concepts de l'explication de la négociation et de l'argumentation. L'objectif est de transposer les principales fonctions du dialogue en langue naturelle dans une interface graphique interactive. Une première réalisation informatique permet d'évaluer l'ergonomie d'une telle interface.

La modélisation des connaissances, des préférences et des croyances, dans un but d'aide à la décision.

En décision dans l'incertain, nous nous sommes intéressés aux difficultés posées par l'évaluation des probabilités subjectives et étudions les possibilités offertes par la représentation moins exigeante des croyances sur les événements par des intervalles de probabilité. Leur approche, axiomatique, conduit à la justification de critères adaptés. Nous avons également entrepris des travaux sur la pratique de la décision à l'aide de scénarios et en particulier sur l'utilité du contexte pour la maîtrise de l'explosion combinatoire. Dans le cadre de l'analyse multi-critère, nous élaborons de nouveaux modèles permettant d'appréhender plus finement les préférences, de concevoir de nouvelles procédures permettant d'exploiter l'information préférentielle, enfin de réaliser des systèmes de sélection, de classement, de tri ou d'affectation. Des études plus théoriques traitent des problèmes soulevés par la décision et l'agrégation de préférences : justification axiomatique d'opérateurs d'agrégation, de procédures de sélection ou de rangement, du codage numérique de l'information. Le traitement probabiliste de l'incertitude a été récemment bouleversé par l'apparition de techniques de calcul s'appuyant sur une représentation graphique de l'interdépendance entre variables aléatoires : la théorie des réseaux probabilistes ou réseaux bayésiens. Ces techniques sont susceptibles d'améliorations et d'extensions, soit à un cadre non-probabiliste, soit à un contexte décisionnel (diagrammes d'influence). Nous poursuivons des recherches théoriques dans ce domaine, qui ont déjà permis de proposer de nouveaux algorithmes de propagation de l'information. De plus, une librairie de programmes, MoByDiCk, outil indispensable aux recherches futures, est en cours de réalisation.

Groupe Systèmes d'aide à la formation (responsable : Michelle Joab)

Un effort important a été fait lors de la création du LIP6 pour regrouper les chercheurs de Paris 6 qui travaillent sur les systèmes d'aide à la formation, afin d'obtenir une meilleure visibilité de ces systèmes et de favoriser les coopérations, l'expérience ayant montré la difficulté à développer des systèmes complets sans équipe de taille suffisante. Ce regroupement permet d'ores et déjà de mettre en évidence des réalisations significatives pour la formation professionnelle et la formation initiale. Ces recherches sont ouvertes sur l'extérieur par le biais de collaborations au sein du GDR I3 du CNRS, de l'ATIEF (association fondatrice de l'ASTI) et de collaborations industrielles. Les systèmes d'aide à la formation sont des systèmes interactifs destinés à favoriser des apprentissages chez leurs utilisateurs. Les « apprenants » sont placés généralement en situation active de résolution de problèmes, tout en étant guidés et contrôlés à des degrés divers par le système. L'ingénierie des connaissances propre à ces systèmes comporte la modélisation du domaine et de la tâche cible, de l'apprenant, de l'interaction et de l'explication ainsi que l'intégration de ces différents modèles. Les difficultés tiennent en grande partie à la quantité des connaissances et des métaconnaissances mises en jeu et à leur diversité, ainsi qu'à la qualité nécessaire de l'interaction homme-machine et des interfaces. La validation des modèles et des systèmes réalisés est un point délicat qui nécessite une coopération pluridisciplinaire notamment avec la psychologie et l'ergonomie cognitives, les didactiques et les sciences de l'éducation. Dans cette action, les travaux suivent deux axes principaux :

- Les systèmes de formation fondés sur la simulation pour les applications industrielles.
- Les EIAO (environnements interactifs d'apprentissage avec ordinateur) pour l'enseignement.

Dans le premier axe, le projet DIAPASON (en collaboration avec EDF) a permis de développer un environnement complet de formation fondé sur la simulation pour les chargés de conduite du réseau électrique moyenne tension. L'explication est construite en fonction des actions du stagiaire et de l'évolution du dispositif industriel simulé. Cette recherche aboutit à une méthode de conception d'explications dynamiques. L'évaluation porte non seulement sur les performances mais surtout analyse de manière critique la démarche du stagiaire ; pour cela elle s'appuie sur la comparaison avec une résolution de référence construite par un résolveur dynamique et paramétrable, couplé à un simulateur du dispositif.

Un projet en collaboration avec Thomson Training & Simulation a débuté fin 1997. Il a pour but de développer un système d'aide à l'instructeur pour un simulateur d'entraînement destiné à l'équipage du char Leclerc (Figure 8). Il s'agit de réaliser essentiellement deux fonctions d'aide à l'instructeur : à court terme un générateur d'exercices et à moyen terme une évaluation du stagiaire.

PRÉSENTATION DE LABORATOIRES

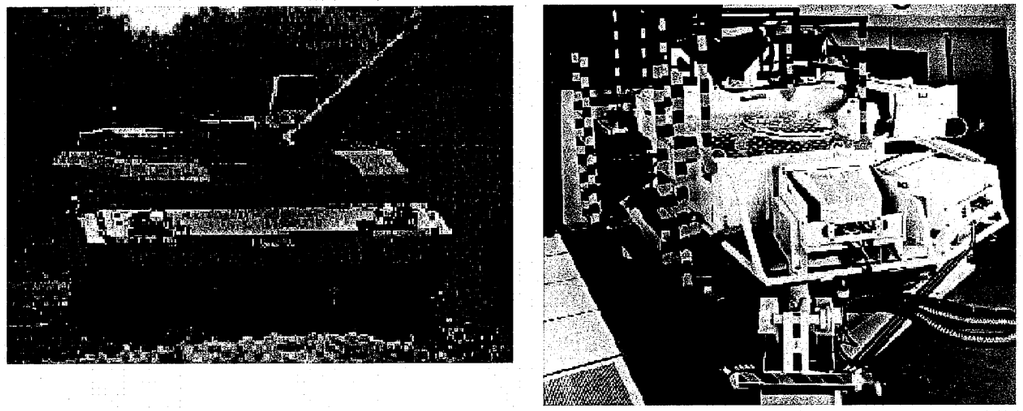


Figure 8 : Un simulateur « pleine échelle » pour le char Leclerc

Nous avons développé une série de systèmes pour la prise de décision en situation de crise : SAGESSE, destiné aux officiers de la sécurité civile, FORMENTOR, pour la formation à la conduite et à la prise de décision en cas d'incidents, de processus industriels complexes tels que le remplissage en combustible d'un satellite lors de son lancement, et INTEMPERIES pour l'analyse des catastrophes naturelles.

Les hypermédias et le multimédia grâce à leurs facilités d'accès aux connaissances et à leur souplesse de présentation contribuent à une meilleure adéquation de l'interface à l'apprenant. Nous concevons et réalisons des outils d'aide à la navigation pour les spécialistes d'un domaine s'initiant à un nouveau système « complexe » ou pour des utilisateurs « naïfs » (Projet PROFIL).

Dans le second axe, le projet « COMBIEN ? » est un EIAO dans le domaine des dénombrements. Le tutoriel est fondé sur la méthode constructive de dénombrement qui consiste à concevoir une définition de l'ensemble à dénombrer qui puisse être interprétée comme un algorithme de génération des éléments de cet ensemble, puis à raisonner sur cet algorithme au niveau « méta », sans l'exécuter. Le système réalisé permet l'introduction d'énoncés formalisés des exercices, de définitions constructives, et d'autre part la mise en œuvre de ces constructions pour trouver la solution et vérifier la validité de ces constructions.

Sur le thème de la modélisation d'apprenant et du diagnostic cognitif, nous étudions plusieurs approches de représentation et de traitement de « connaissances imparfaites » (logiques de connaissance et de croyance, logiques non monotones, formalisation de raisonnements révisables, systèmes de maintien de cohérence, révision). Les modèles de croyances ont été appliqués dans le cadre d'un environnement d'apprentissage collaboratif, où deux apprenants communiquent à distance, dans le cadre du projet CHENE.

Au sein du groupe de travail interdisciplinaire PROVERB de l'ARC (Association pour la Recherche Cognitive), nous travaillons sur la modélisation cognitive d'un humain engagé dans une tâche de résolution de problèmes. Le groupe développe plusieurs modèles de résolution, simulant à différents niveaux le comportement du sujet. Les modèles sont construits à partir de l'analyse psycholinguistique de protocoles verbaux.

Les projets transversaux

Le LIP6 développe également une politique de *projets transversaux* aux thèmes de recherche. Ces projets ont pour vocation d'aider des recherches émergentes et jugées importantes pour le laboratoire, des collaborations entre les thèmes ou d'aider à finaliser et à promouvoir des recherches bien établies au sein du laboratoire. Ils bénéficient d'un financement interne au laboratoire. Les projets actuels auxquels participent actuellement les chercheurs en IA concernent l'indexation multimédia, les mondes virtuels, les assemblées de robots, la gestion de systèmes de production, les entrepôts de données, la recherche d'information et l'apprentissage en analyse d'informations textuelles.

La formation de troisième cycle

Le LIP6 est impliqué dans l'école doctorale « EDITE » (directeur B. Robinet (ENST) - sous directeur C. Girault (LIP6)) qui outre les laboratoires de Paris 6, regroupe également des laboratoires du CNAM, de l'ENST, de l'ESPCI et de l'École Polytechnique.

Les enseignements en IA sont dispensés au sein de deux DEA et de deux DESS :

- le DEA IARFA (Intelligence Artificielle et Reconnaissance des Formes), responsable J.G. Ganascia, *http://dea-si.lip6.fr/EcoleDoc/DEA_IARFA. html*
- le DEA IRO (Informatique et recherche Opérationnelle), responsable M. Minoux *http://dea-si.lip6.fr/EcoleDoc/DEA_IRO. html*
- le DESS IA (Intelligence Artificielle), responsable A. Doucet,
- le DESS GLA (Génie des Logiciels Applicatifs), responsable J.-F. Perrot *http://gla.ecoledoc.lip6.fr*

Les relations industrielles

Nous donnons ci dessous une brève description de projets réalisés dans le cadre de relations contractuelles avec des industriels ou avec des grands organismes.

Accès intranet à un datawarehouse :
AIDA (ministère de l'industrie)
Prise en compte de la modélisation des utilisateurs pour l'accès à des ressources intranet.
Partenaires : ATOS, Renault (1997-1999)

Aide au diagnostic flou à partir d'images médicales : GEMS-E
Outils issus de la théorie des sous ensembles flous pour l'analyse d'images médicales issues de mammographies, aide au diagnostic du cancer du sein. (1997-2000)

Conception, Analyse et Réalisation de Sous-Systèmes Électroniques Embarqués : PREDIT 1996/2000
Aide à la conception de systèmes électroniques d'automobiles selon une approche multi-agent (1999-2001)
Partenaires : LORIA, PSA

État de l'art dans les domaines de la métaconnaissance et de l'apprentissage
EDF (1998-1999)

Financial Business Object Framework
FIBOF (ESPRIT IV n° 21966)
Conception de framework pour les secteurs bancaire et financier, adaptable aux spécificités de chaque institution financière. Ce framework doit permettre d'engendrer simplement, à partir d'une spécification dans un langage compréhensible par les gestionnaires, les applicatifs nécessaires à la gestion de nouveaux produits financiers (09/96-08/98
Partenaires : El Monte (E), ECI SA, SOLERI

Générateur de cursus pour le simulateur d'entraînement d'équipage du char Leclerc
Thomson Training & Simulation
Système d'aide à l'instructeur pour les simulateurs d'entraînement du char Leclerc. (1997-1999)

Jardins de Données
CTI CNET
Conception multi-agent d'interfaces utilisateurs pour le suivi de données numériques en temps réel, partant de la métaphore d'un jardin artificiel soumis aux variations de données météorologiques (06/98-06/2000)

Linguistique, informatique, et corpus oraux dictée vocale
AUPELF-URELF
Évaluation de systèmes de dictée vocale sur des corpus de langue française. (1995/1998)

Modèles connexionnistes adaptatifs pour systèmes temps réel
Philips (LEP),
Application d'algorithmes adaptatifs à la prévision de décharge de batteries et au suivi de contours dans des images médicales (1996-1998)

Modélisation de contraintes spatio-temporelles floues dans les Gabarits
Thomson CSF
Application de la théorie des sous ensembles flous pour la prise en compte de contraintes temporelles et spatiales. (1995-1998)

Optimisation du profil d'ailes d'avion et de la réflectance radio à l'aide d'algorithmes génétiques
Dassault Aviation
Utilisation d'algorithmes génétiques pour le calcul numérique employés dans le domaine de l'avionique. (1996-1998)

Reconnaissance automatique de caractères manuscrits
La Poste
Amélioration de techniques de classification et des calculs de confiance associés aux décisions pour la reconnaissance de caractères manuscrits. (1998-2001).

Réseaux de neurones pour le diagnostic et la prévision du trafic téléphonique
CTI CNET
Diagnostic en temps réel dans les réseaux téléphoniques, à l'aide de réseaux de neurones et ed réseaux bayésiens. (1995-1998)

Système d'aide à la décision de tir par les chars en position de combat
GIAT Industries
Introduction de techniques de la logique floue pour la prise en compte de données numériques et de données subjectives, dans un système d'aide à la décision, application à la décision de tir dans les chars. durée : 1995-1998)

Réseaux d'excellence

Le LIP6 est membre de cinq réseaux d'excellence dans le domaine de l'IA :

AgentLink, *Europe's network of excellence for agent-based computing* (1998-2000)
Erudit, *Network of Excellence in Uncertainty Techniques Developments for Use in Information Technology* (1998-2000)
Ingenet, *Networked industrial design and control applications using genetics algorithms and evolution strategies* (1997-2000)
Mlnet, *Network of Excellence in Machine Learning* (1998-2001)
NEuroNet, *Neural Networks network of excellence* (1999-2001)

La diffusion de logiciels

Logiciels réalisés au sein du laboratoire et diffusés par le réseau
Actalk
framework de programmation concurrente en Smalltalk. A en particulier servi de base logicielle à de nombreux projets multi-agents dans divers laboratoires.
ftp ://ftp.lip6.fr/lip6/softs/actalk/

NéOpus
moteur d'inférence d'ordre un intégré à Smalltalk VisualWorks 2.5. Utilisé par de nombreux projets.
ftp ://ftp.lip6.fr/lip6/softs/NeOpus/SourcesNeOpus/

BackTalk
bibliothèque d'algorithmes de satisfaction de contraintes intégrée à Smalltalk.
ftp ://ftp.lip6.fr/lip6/softs/NeOpus/BackTalk/

MoByDiCk
Librairie de gestion de graphes et de réseaux probabilistes
http://www-poleia.lip6.fr/~wuille-mi/Datas/mobydick.zip

Remus
logiciel de reconnaissance d'écriture manuscrite dynamique
http://www-poleia.lip6.fr/CONNEX/HWR/

La présentation du LIP6 (bulletins 37 & 38) a été rédigée par B. Bouchon-Meunier, J.-P. Briot, P. Gallinari et J.-Y. Jaffray

Références bibliographiques choisies

Thème APA

Bothorel S., Bouchon B., Muller S., A Fuzzy Logic-based Approach for Semiological Analysis of Microcalcification in Mammographic Images, *International Journal of Intelligent Systems* (1998).

Bouchon B., Rifqi M., Bothorel S., Towards General Measures of Comparison of Objects, *Fuzzy Sets and Systems,* 84, 2 (1996) p. 143-153.

Cazenave T., Machine Introspection for Machine Learning, *Journal of Consciousness Studies* (1998).

Cazenave T., Controlled Partial Evaluation of Declarative Logic Programs, *ACM Computing Surveys* (1998).

Chopard-Guillaumot G., Hadj-Mabrouk H., Ganascia J.G., Contribution à une meilleure définition de l'analyse préliminaire de risques pour les systèmes de transports guidés, *Journal Européen des Systèmes Automatisés,* (1997).

Corruble V., Ganascia J.G., Experimentations on the dynamics of induction in medicine. *Artificial Intelligence Journal, Special issue on scientific discovery,* vol. 91 n° 2, (1997) p. 205-223.

Garcia-Salicetti S., Dorizzi B., Gallinari P., Wimmer S., Maximum mutual information training for an on-line neural predictive word recognition system, *Pattern Recognition,* à paraître.

Gérard O., Patillon J.N., d'Alché-Buc F., Discharge Prediction of Rechargeable Batteries with Neural Networks, *International Journal of Integrated Computer-aided Engineering, Special Issue on Neural Network technique for Industrial Applications* (1997).

Leray P., Gallinari P., Feature Selection with Neural Networks, Behavimetrika, *Special Issue on Analysis of Knowledge Representations in Neural Network Models* (1998).

Pitrat J., Implementation of a Reflective System, *Future Generation Computer Systems,* 12 (1996) p. 235-242.

Ramalho G., Rolland P.Y., Ganascia J.G., An Artificially Intelligent Jazz Performer, Journal of *New Music Research* (accepté).

Rolland P.Y., Ganascia J.G., Musical Pattern Extraction and Similarity Assessment, *Contemporary Music Review,* Special Issue on Music and Artificial Intelligence (accepté).

Toudeft A., Gallinari P., Neural and adaptive controllers for a non-minimum phase varying time-delay system, Artificial Intelligence in Engineering, special issue on *Neural Networks in Process Engineering* (1997).

Ying M., Bouchon B., Quantifiers, Modifiers and Qualifiers in Fuzzy Logic, *Journal of Applied Non-classical Logics* 7, 3, (1997) p. 335-342.

Zaragoza H., Gallinari P., Multiple Multivariate Nonlinear Regression : A Large Scale Thermodynamical Application, *Signal Processing,* vol. 64, N°. 3 (1997) p. 371-382.

Zucker J.D., Ganascia J.G., Bournaud I., Relational Discovery in a Chinese Characters Data-Base, *Applied Artificial Intelligence Journal,* (accepté).

Thème OASIS

Cardon A., Les Systèmes d'Information et de Communication de gestion de crise : une modélisation par agents, *Ingénierie des Systèmes d'Information,* Vol. 5, No 2, p. 167-193, Ed. Hermes, Juin (1997).

Collinot A., Drogoul A., Approche orientée agent pour la conception d'organisations : application à la robotique collective, *Revue d'Intelligence Artificielle,* 12 (1), Ed. Hermès (1998), pp.125-147.

Doucet A., Gancarski S., Jomier G., Monties S. Maintien de la cohérence dans une base de données multiversion, *In Ingénierie des Systèmes d'information,* 5 (1), (1997).

Drogoul A., Fukuda T., Tambe M. (Eds), *Collective Robotics,* Lecture Notes in Artificial Intelligence, Springer-Verlag, (1998).

Guessoum Z., Briot J.P., From active objects to autonomous agents, *IEEE Concurrency Special Series on Actors & Agents,* à paraître, (1999).

Meyer J.A., Husbands P., et Harvey I., Evolutionary Robotics : a Survey of Applications and Problems. In Husbands, P. and Meyer, J.A. (Eds.). *Proceedings of The First European Workshop on Evolutionary Robotics — EvoRobot'98.* Springer-Verlag (1998).

Roy P., Liret A., Pachet F., « The Framework Approach for Constraint Satisfaction », *ACM Computing Surveys on Frameworks,* à paraître (1998).

Sahraoui H., Revault N., Blain G., Perrot J.F., Un outil pour la conception de bases de données à objets, *Technique et Science Informatiques (TSI),* (1997).

Thème SYSDEF

Auzende O., Aide à l'opérateur dans un système de formation à la conduite de processus. *Revue d'Intelligence Artificielle,* 12 (4), 467-496. 1998.

Brézillon P., Modeling and using context : Report on the first international and interdisciplinary conference CONTEXT-97, *The Knowledge Engineer Review* (1997). (à paraitre)

Bruillard E., de La Passardière B., Fonctionnalités hypertextuelles dans les environnements d'apprentissage, Les Hypermédias, Tricot A., Rouet JF, Eds, pp 95-122, Hermès 1998.

Caetano H. et Mathieu J., Des systèmes de formation à la gestion de catastrophes fondés sur la coopération entre institutions intervenantes, Première Conférence Francophone de Modélisation et de Simulation, Systèmes de Production et de

Logistique (MOSIM'97), Rouen (1997).
Chrétienne P., List Schedules for Cyclic Scheduling, *Discrete Applied Mathematics* (1997). (à paraitre)
Dubois O. et Boufkhad Y., A general upper bound of the satisfiability threshold for random r-SAT formulæ, *Journal of Algorithms,* vol. 24 (1997), p. 395-420.
Gonzales C. et Wuillemin P.H., Réseaux Bayésiens en Modélisation Utilisateur, *Sciences et Techniques Educatives* (1998).
Jaffray J.-Y. et Philippe F., On the Existence of Subjective upper and lower Probabilities, *Mathematics of Operations Research,* vol. 22, n° 1 (1997), p. 165-185.
Labat J.M. et Mynard L., Oscillation, heuristic ordering and pruning in neighborhood search, Third International Conference on Principles and Practice of Constraint Programming (CP'97), Schloss Hagenberg, Autriche (1997), p. 506-518.
Lund K., Baker M. et Baron M., Modelling Dialogue and Beliefs as a basis for generating guidance in a CSCL environment, Third International Conference on Intelligent Tutoring Systems, Montréal, Canada (1996), vol. LNCS n° 1086, Springer Verlag, p. 206-214.
Moinard C., Joab M., Acquisition de connaissances techniques pour le système DIAPASON. *Sciences et Techniques Éducatives,* 1 (4), Hermès, Paris, 457-481. 1994
Munier A. et Hanen C., Using duplication for scheduling unitary tasks on m processors with communication delays, *Theoretical Computer Sciences* (1997), p. 119-127.
Perny P., Multicriteria Filtering methodes based on concordance and non-discordance principles, *Annals of Operations Research,* n° Special issue on « Preference Modelling » (1997).
Pomerol J.C., Artificial Intelligence and human Decision making *European Journal of Operational Research,* vol. 99 (1997) pp 59-72.
Tisseau G., Giroire H., Le Calvez F., Urtasun M., Duma J., « Design of a learning Environment in combinatorics : Nondeterministic machines to improve modelling skills », in proceedings of AI-ED 99, Le Mans, july 19-23, 1999.

Présentations de laboratoires dans le bulletin de l'AFIA

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

IA et WEB : premiers rapports.

Dossier AFIA, coordonné par
Philippe Laublet, Amedeo Napoli, Serge Stinckwich

Introduction

L'idée de ce dossier est née d'une double interrogation : d'une part, quel est l'impact du développement d'Internet (existence de sites WEB de plus en plus nombreux) sur les travaux et les recherches en intelligence artificielle. D'autre part, de manière corrélée, quels sont les nouveaux besoins émergents et les attentes provoquées par ces développements sociologiques et technologiques auxquels les recherches et techniques de l'IA peuvent contribuer à répondre ?

Bien sûr ce dossier ne prétend pas répondre de manière synthétique à ces questions Il est simplement une photographie à un moment donné de travaux pour la plupart relativement récents même s'ils s'appuient sur des recherches souvent anciennes. Ce sera au lecteur de se faire une première opinion. Il nous a semblé néanmoins intéressant de mettre rapidement à disposition des exemples de ces travaux particulièrement pour de jeunes chercheurs, commençant par exemple une thèse.

Pour certains, ces nouvelles technologies ne changent rien fondamentalement aux objectifs majeurs de l'IA. Pour d'autres, l'impact est important, ne serait-ce que par le changement d'échelle et un rapport différent aux utilisateurs des systèmes réalisés ainsi que la disponibilité de techniques de plus en plus standard. De nouveaux défis seraient alors posés à l'IA. Le plus visible est bien sûr de contribuer à construire des assistants ou agents pour l'aide à la collecte d'information dans le nouveau contexte du WEB, mais ceci n'est qu'une des facettes de ce paysage mouvant.

Le WEB est plus qu'un serveur de documents et les transformations induites par le WEB vont toucher beaucoup d'autres activités. Dans un livre récent « Internet Dreams », sous-titré « Archetypes, Myths and Metaphors », Mark Stefik a collecté un ensemble d'articles historiques ou récents pour nous aider à resituer, à comprendre et à explorer les possibilités offertes par les infrastructures d'information émergentes. Il identifie quatre grands domaines actuels d'utilisation des réseaux à partir de quatre « archétypes » anciens qui ont influencé la pensée humaine depuis des millénaires, le gardien de la connaissance, le communicateur, le commerçant et l'aventurier. Ces domaines sont les bibliothèques électroniques, le courrier électronique, le commerce électronique et les mondes virtuels.

De manière plus générale, ce sont de nombreux modes d'échanges professionnels ou autres, d'interactivité ou de travail coopératif qui sont ou seront concernés. L'Internet est avant tout une plate-forme pour chercher, se former, travailler… de manière collaborative. Notons au passage que la dimension mondiale d'Internet ne doit pas cacher, plus important peut-être encore, son utilisation et ses potentialités pour des communautés réduites dispersées, partageant certains intérêts.

Différentes équipes de recherche sont aujourd'hui engagées dans la conception et la réalisation de logiciels apportant une aide accrue aux utilisateurs du WEB investis dans différentes tâches. Sans souci d'exhaustivité, on peut regrouper leurs travaux selon les thèmes suivants : assistants à la « navigation », agents de recherche, collecte et filtrage d'informations, personnalisation de services, construction, mise à la disposition et utilisation de bases de connaissances, partage et réutilisation de connaissances via le WEB, représentation de textes et raisonnements, fouille de textes et de documents hétérogènes, enrichissement des outils de commerce électronique, outils pour le travail coopératif, pour la réalité virtuelle en réseau ou l'éducation à distance…

Il nous semble que la réflexion sur ces thèmes et sur la conception d'outils prenant mieux en compte les différentes dimensions de l'homme dans ces activités ne peuvent que bénéficier de la collaboration et des apports des praticiens de l'intelligence artificielle.

Les thèmes de la liste ci-dessus ne sont pas représentés au même niveau dans les travaux des différentes équipes présentées ci-dessous. Une majorité de ceux-ci peut être classée sous la problématique de la recherche et de la collecte d'information, même si, comme nos lecteurs le constateront, d'autres problématiques sont déjà présentes. Ce dossier semble refléter à peu de choses près la diversité des travaux effectués en France à l'heure actuelle. Quoi qu'il en soit, on peut parier sur un élargissement rapide des problématiques considérées dans l'avenir. Dans le domaine les choses vont vite et il est vrai que ce dossier aurait été certainement plus complet et plus étoffé s'il avait été programmé un ou deux ans plus tard.

Collecte d'information sur le WEB

Le WEB porte bien son nom : c'est une gigantesque toile d'araignée avec d'innombrables nœuds et mailles. Les nœuds matérialisent des documents électroniques de toutes natures, les mailles les chemins pour les atteindre. Le WEB est devenu un outil commun à beaucoup, informaticiens ou non ; les enjeux scientifiques et économiques sont donc considérables.

La problématique de la collecte d'information est ainsi sans aucun doute la meilleure illustration du renouvellement de vieilles questions dans le contexte du WEB :

- La taille du WEB, les estimations, il y a deux ans, donnaient entre 30 millions et 150 millions de pages WEB indexées automatiquement et plus de 10 000 groupes de discussions. (une « page » WEB correspond en fait à une adresse et peut être de longueur variable).
- Il n'y a pas de catalogues centralisés de toutes les ressources.
- Chaque utilisateur peut être à la fois consommateur et fournisseur d'informations.

- Il n'y a généralement pas de contrôle ou d'évaluation sur les informations mises à disposition d'où le problème de la validité des informations.
- Il existe une grande diversité
  - des genres de ressources : textes, images, logiciels, sons, etc.
  - des textes eux-mêmes (rapports techniques, articles de journaux, informations ponctuelles, discussions dans les news, etc.)
  - des formats informatiques, plus ou moins utilisables par les outils de recherche.
- Les informations gratuites sont en grande quantité. La diffusion de l'information publique est une des principales motivations d'Internet, cette information est généralement non structurée au sens des bases de données.
- Mais le WEB donne aussi accès à des informations payantes (abonnement, au coup par coup...). Les banques de données commerciales sont de plus en plus souvent accessibles avec les mêmes interfaces. On disposera dans l'avenir d'un espace unique d'accès aux données : futures versions de Windows, outils d'indexation en local, couche de présentation des bases de données, etc.
- Il y a une répartition inégale des documents et informations suivant les domaines, les zones géographiques, les différentes langues, etc.
- La croissance exponentielle, les fréquences variables de renouvellement et de mise à jour et les changements de localisation compliquent la tache des usagers et... des outils de recherche.
- Mais il y a une contrepartie : « sur le WEB, où il y a un besoin, il y a un programmeur ». La multiplicité des initiatives diverses, au départ souvent individuelles, a permis de créer des outils pour aider à s'y retrouver.

L'utilisateur explore en général le WEB pour y rechercher des informations avec essentiellement deux attitudes possibles : (1) une recherche avec un but défini et éventuellement une adresse précise ou bien (2) une recherche sans but défini, en forme de demande générale d'information.

Par exemple, les questions suivantes peuvent être posées dans le cadre de la première recherche : recherche d'un article portant le titre X, recherche de l'article de l'auteur X publié à la conférence Y, tandis que dans la seconde, elles sont plutôt du type : qu'est-ce qui existe sur le sujet X ? recherche d'un article de X qui parle du sujet Y.

Dans les deux cas, l'attitude de l'utilisateur est différente et n'implique pas le même temps passé à la recherche. Force est de constater que, face à la toile qu'est le WEB, l'utilisateur ressemble à la mouche prise aux mailles du filet de l'araignée, et qui essaie désespérément de s'en sortir : comment finaliser une recherche, comment traiter cet immense volume d'informations, comment s'y retrouver, comment y trouver un bon chemin et exploiter au mieux les immenses ressources de cette toile ?

La recherche d'une information fait intervenir les désormais fameux moteurs et leur recherche sur des pages indexées. La recherche d'un document se fait « par le contenu », c'est-à-dire sur la base d'une comparaison entre la question de l'utilisateur et le contenu de la base d'informations. Il faut aussi souligner que, sur le WEB, la recherche d'informations ou de documents est fondamentalement différente de la recherche d'informations structurées (i.e., les bases de données) : la première est plutôt une recherche par similarité alors que la seconde s'appuie en général sur un appariement exact puisque la structure des documents est connue.

Plusieurs constats peuvent être faits à l'issue d'une recherche d'information :

- le rappel et la précision sont plutôt médiocres. Le rappel mesure le « silence » ou le taux de documents pertinents qui n'apparaissent pas dans le résultat d'une recherche ; la précision mesure le « bruit » ou le taux de documents retrouvés mais non pertinents. Une bonne recherche d'information nécessite de minimiser bruit et silence.
- pour explorer le WEB correctement, il faut « savoir » poser les bonnes questions ; en général il faut en poser plusieurs et s'appuyer sur les premières recherches pour affiner les suivantes. Le processus de recherche est approximatif et itératif (essai-erreur).

De plus, l'utilisateur ne sait en général rien sur la façon dont le WEB est alimenté et sur le vocabulaire utilisé, ce qui pose des problèmes lorsqu'il s'agit d'interroger la base (formulation d'une requête en fonction des éléments en présence).

Questionner intelligemment le WEB est un problème d'importance majeure à l'heure actuelle, et il est devenu naturellement un thème de recherches pour de nombreuses équipes d'IA qui s'intéressent au WEB, et ce dossier reflète bien cette constatation. En réalité le problème se formule plutôt comme suit : comment fournir une assistance à l'utilisateur qui explore le WEB pour rechercher une certaine information, précise ou non.

Il s'agit d'assistance à la navigation ou « navigation intelligente » (qui peut faire intervenir des agents intelligents ou encore des guides qui accompagnent l'utilisateur pendant sa promenade électronique par exemple).

(IA et WEB) ou (WEB et IA)

Comme souligné ci dessus, les travaux autour du WEB qui concernent l'intelligence artificielle ne se limitent pas aux seules recherches d'information et de documents. Sans chercher à en faire une liste exhaustive, notons parmi les tendances principales :

(1) l'aide à la navigation,
(2) l'extraction de connaissances à partir des données du WEB et principalement la fouille de textes,
(3) l'intégration de documents hétérogènes et de données en provenance de bases de données hétérogènes,
(4) la technologie du WEB appliquée à la technologie de l'intelligence artificielle dans le cadre par exemple de la construction de systèmes à bases de connaissances,
(5) la technologie du WEB et ses rapports avec la linguistique et le multimédia, avec par exemple le traitement des problèmes liés au multilinguisme,
(6) le développement des outils de formation, de loisir, et de travail et d'échange coopératif à l'aide du WEB.

Tous ces points entrent dans la problématique de la conception des systèmes intelligents de l'avenir, qui s'appuieront sur des systèmes à bases de connaissances couplés avec des bases de données (et leurs gestionnaires), avec un accès au WEB, à partir duquel il sera possible d'exploiter les systèmes à bases de connaissances et les bases de données.

Les points (1), (2) et (3) concernent plutôt la technologie de l'IA appliquée au WEB, qui recouvre essentiellement les problèmes de recherche intelligente d'informations et de documents, la recherche selon un point de vue ou sur la base de contraintes, le raisonnement sur l'accès au WEB et au document (classification de requêtes, détection de requêtes plus générales ou de requêtes analogues, historiques des requêtes, gestion des retours arrières, sauvegarde de requêtes, etc.), l'intégration d'informations de données en provenance de sources hétérogènes en vue de fusionner ou faire collaborer plusieurs bases de données ou plusieurs documents provenant de sources hétérogènes, la prise en compte du profil de l'utilisateur et de son comportement pour mettre en œuvre un apprentissage en fonction des explorations déjà pratiquées par l'utilisateur (raisonnement à partir de cas).

Toutes ces recherches impliquent de s'intéresser à la combinaison de techniques de gestion de bases de données et de bases de connaissances, à la gestion de très grandes bases de données et de très grandes bases de connaissances, à l'aide à la navigation en s'aidant de connaissances sur le domaine exploré, à la mise en œuvre de systèmes de traitement de l'information intelligents et intégrés.

Le point (4) concerne la technologie du WEB appliquée à l'IA et recouvre principalement la place que peuvent occuper les navigateurs - comme Netscape par exemple - dans la gestion des accès à un système à base de connaissances ou à une base de données ; le navigateur est alors considéré comme un « frontal » permettant d'utiliser - depuis n'importe quel endroit dans le monde - un système à base de connaissances. Par ailleurs, les navigateurs peuvent être aussi utilisés avec profit dans le cadre des tuteurs intelligents (outils de recherche et de visualisation, associés à une base de documents).

Le point (5) concerne essentiellement les problèmes liés à la traduction de documents électroniques (dans le cadre des besoins de la communauté européenne par exemple), mais aussi tout ce qui est en rapport avec la recherche et l'indexation de documents multimédia, et en particulier la gestion de documents non textuels comme les séquences vidéo, les bandes sonores, les collections d'images et de photographies, etc.

Soulignons pour terminer (point 6) que les applications des systèmes multi-agents et de travail coopératif sont peu présents dans ce dossier. Pour cette raison, ces points sont pratiquement absents de cette brève introduction, bien que, IA distribuée oblige, ils constituent un des champs principaux du possible avenir commun entre Internet et IA. Cette introduction a comme seule prétention d'ouvrir une discussion, à poursuivre donc éventuellement dans la rubrique « débats » du bulletin.

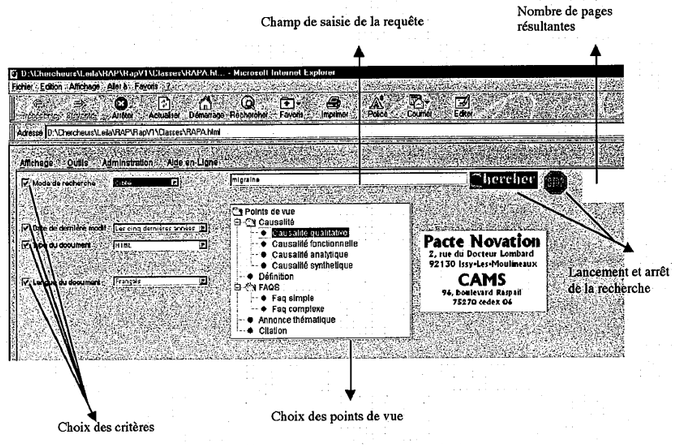
Philippe Laublet, Amedeo Napoli

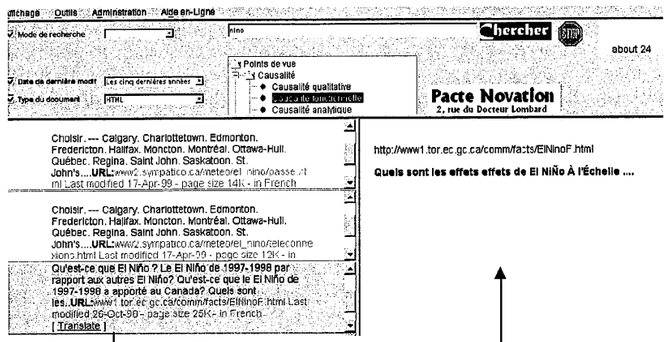
CAMS, ÉQUIPE LALIC (LANGAGE, LOGIQUE, INFORMATIQUE ET COGNITION)

CAMS (Centre d'Analyse et de Mathématiques Sociales)
UMR 17, CNRS, EHESS, Paris-Sorbonne
96 boulevard Raspail 75006 PARIS
Contact : Philippe. Laublet@msh-paris.fr

Les travaux sur l'extraction de concepts, de relations et de connaissances opératoires à partir de textes prennent une place croissante dans le contexte de l'informatique linguistique en France. La nécessité largement reconnue de meilleurs outils d'assistance à la collecte d'informations sur le WEB pose de nouveaux défis en vraie grandeur auxquels il semblait prometteur et intéressant de s'attaquer à l'aide des connaissances linguistiques largement étudiées dans l'équipe dirigée par J.P. Desclés. Ces travaux sont menés, en collaboration avec la société Pacte Novation et avec l'URFIST de Paris, dans le cadre du projet RAP « Recherche, Analyse, Propose », projet financé par le Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche depuis fin 1997. Pour la collecte d'information sur le WEB, nous combinons en fait deux méthodes qui s'appuient toutes deux sur l'utilisation de marqueurs linguistiques : la reformulation de requêtes et l'analyse sémantique des textes par exploration contextuelle [Desclés 91, 97] afin de présenter des extraits significatifs de ces textes à l'utilisateur de RAP.

La démarche classique dans une recherche d'informations consiste à construire des requêtes à partir de termes ou d'expressions nominales devant donner accès à des documents pertinents pour cette recherche, par exemple : (i) « migraine », (ii) « migraine » et « traitements de fond » ou « traitements de la crise ». Cette démarche est





Panneau d'affichage des pages résultantes

Panneau d'affichage des extraits correspondants aux pages sélectionnées dans le panneau de gauche

d'autant plus aisée que l'utilisateur connaît bien le domaine d'investigation, la terminologie qui en relève, et que l'information qu'il recherche a des identifiants incontestables. L'utilisateur peut alors affiner aisément sa requête si celle-ci ramène un nombre trop important de réponses. Dans le cas où la recherche d'informations concerne un domaine non familier ou émergeant, la formulation des requêtes devient plus délicate. Par exemple, aborder le phénomène météorologique El Niño à partir d'une requête simple « El Niño » revient à s'exposer à un flot considérable d'informations difficilement appréhendable sans recours à d'autres critères de sélection [Naït-Baha *et ali.* 98].

De plus, l'utilisation de requêtes booléennes, même enrichies d'opérateurs variés comme la troncature ou l'opérateur de proximité, nécessite une mise en forme soigneuse des requêtes. Or, celle-ci est rarement effectuée et il a pu être observé que la plupart des requêtes formulées par les utilisateurs du WEB, non professionnels de la documentation, sont généralement constituées par peu de mots.

Afin d'aider l'utilisateur à mieux cibler le contenu dans les textes, nous proposons d'une part de diviser l'espace de recherche en plusieurs sous espaces de recherche présentés à l'utilisateur sous forme de points de vue, et d'autre part, d'extraire des informations textuelles des documents sélectionnés en recourant à un moteur d'exploration des textes. Cela afin de présenter des extraits permettant un choix plus informé, donc plus rapide et pertinent, des textes à consulter. Chaque point de vue capte des traces informationnelles spécifiques dans les textes.

Du point de vue de la reformulation de requêtes, l'idée générale consiste à associer aux termes de la requête de l'utilisateur des termes spécifiques (marqueurs linguistiques) correspondant à chacun des points de vue sélectionnés, et à reformuler automatiquement, en amont du processus de recherche, les requêtes selon les points de vue choisis, afin de cibler plus précisément des informations présentes dans les documents que l'on recherche. Cette reformulation est « transversale » aux domaines explorés par l'utilisateur, c'est-à-dire que l'élaboration de ces marqueurs linguistiques ne repose pas sur les connaissances relatives à des domaines particuliers.

Le point de vue qui est actuellement expérimenté dans le cadre de *RAP* — projet dont le prototype informatique a été réalisé par Brahim Djioua, puis repris par Leila Naït-Baha [Djioua *et ali.* 98] — est celui de la causalité. Les connaissances causales utilisées sont issues des travaux d'Agata Jackiewicz [Jackiewicz 98]. Notons que de nouvelles recherches sont programmées afin d'étendre progressivement les capacités du système en intégrant d'autres points de vue parmi ceux par exemple de la définition d'un terme, différentes relations entre concepts… intégration qui pourra être faite dans le système lorsque la méthodologie testée sur la causalité aura été bien éprouvée.

Nous disposons de plus de 1100 marqueurs pour les relations causales. Les marqueurs qui sont mis en œuvre dans la reformulation du point de vue de la « causalité » sont choisis actuellement selon deux critères : (i)- la fiabilité du marqueur (le marqueur le moins polysémique) ; (ii)- la valeur sémantique du marqueur. Différentes approches du point de vue de la causalité [Jackiewicz 98] à savoir la causalité qualitative, la causalité analytique, la causalité fonctionnelle et la causalité synthétique permettent d'utiliser de manière plus sélective des ensembles plus restreints de marqueurs.

Ensuite, dans une deuxième phase, à partir des documents produits par la reformulation par points de vue RAP procède à une extraction de segments de textes pertinents comme extraits significatifs des documents. Une évaluation de l'intérêt du système est en cours de réalisation.

Liste des contributions

CAMS - Paris (équipe LaLIC)
CAMS - Paris (projet Concerto))
DIAM - Paris (service d'informatique médicale AP-HP)
ENSMSE - SIMADE (département RIM)
INRIA - Rhône-Alpes (projet Sherpa)
INRIA - Sophia Antipolis (projet ACACIA)
INRIA - Sophia Antipolis (action AID)
Institut Interfacultaire d'Informatique - Neuchatel
IRIT - Toulouse (équipe SMAC)
LIUM - Le Mans (équipe EIAO)
LORIA - Nancy (équipe Langue et Dialogue))
LORIA - Nancy (équipe Orpailleur)
LRI - Paris Sud (équipe IASI)
PSI - INSA - Rouen
PSY-CO - Rouen (équipe EMC2)
Université d'Ottawa (groupe Tamale) + RES International
Université de Savoie (équipe SysCom)

Références

[Desclés 91] Desclés J.-P., Jouis C., Oh H.-G., Reppert D., « Exploration contextuelle et sémantique : Un système expert qui trouve les valeurs sémantiques des temps de l'indicatif dans un texte », in Knowledge modeling and expertise transfert, D. Hérin-Aime, R. Dieng, J.P. Regourd, J.P. Angoujard (éds), Amsterdam, Washington DC, Tokyo : IOS Press

[Desclés 97] Desclés J.-P., Cartier E., Jackiewicz A., Minel J.-L., « Textual Processing and Contextual Exploration Method », in International and Interdisciplinary

Conference on Modeling and Using Context, Context'97, Rio de Janeiro, Brésil, 1997.

[Djioua et ali. 98] Djioua B., Jackiewicz A., Laublet P., Naït-Baha L. « Recherche et filtrage d'information dur les réseaux », Rapport de Recherche du projet RAP pour le Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche. Décision 97 K6428, novembre 1998.

[Jackiewicz 98] Jackiewicz A. L'expression de la causalité dans les textes. Contribution au filtrage sémantique par une méthode informatique d'exploration contextuelle. Thèse de doctorat. Université de Paris-Sorbonne, 1998.

[Naït-Baha et ali. 98] Naït-Baha L., Jackiewicz A., Laublet P. « Reformulation de requêtes et extraction de phrases pertinentes pour la collecte d'informations sur le WEB ». In Proceedings of RIFRA'98, Sfax, Tunisie, 11-14 novembre 1998.

[Naït-Baha et ali. 99] Leila NAIT-BAHA, Agata JACKIEWICZ, Brahim DJIOUA. « Reformulation de requêtes pour la collecte d'informations sur le WEB à partir de points de vue : premiers résultats », ISKO'99, Lyon, France, octobre 1999.

ATELIER D'INGENIERIE DES DONNEES DU CAMS PROJET CONCERTO

Atelier d'ingénierie des données du CAMS (Centre d'Analyse et de Mathématique Sociales),
UMR CNRS n° 17
54, Bd Raspail
75270 Paris Cedex 06
Tél. : 01 49 54 20 33
Télécopie : 01 49 54 21 09

Gian Piero ZARRI, Sarah JACQMIN, Béatrice MARIN
{ zarri, jacqmin, marin}@msh-paris.fr

Cadre général, objectifs du projet CONCERTO

CONCERTO [1] est un projet européen (ESPRIT P29159), regroupant des industriels, des universités et des centres de recherches, qui s'achevera le 15 septembre 2000. Il rassemble cinq partenaires principaux — PIRA International (Royaume-Uni, coordinateur), ASSETT (Grèce), CNRS, IBM Italie et QUINARY (Italie) — et trois partenaires associés — le Centre de Linguistique Computationnelle de l'UMIST de Manchester (Royaume-Uni), et les Départements d'Informatique des Universités de Bari et de Milan (Italie). CONCERTO est un acronyme pour

IA et WEB : quelques informations.

Adresse utile : le consortium W3C : http://www.w3.org/

Journaux.

ACM SIGMOD Record, http://www.acm.org/sigmod/record/

Artificial intelligence techniques for emerging information systems applications

IEEE Internet Computing, http://computer.org/internet/

IEEE Intelligent Systems, http://computer.org/intelligent/ dont un numéro spécial sur

« knowledge management and knowledge distribution over the internet »

http://www-sop.inria.fr/acacia/ieee-proposal.html

IEEE Expert

Numéros spéciaux :

The AI Magazine 18 (2), 1997.

International Journal of Human Computer Studies 46 (6), 1997.

International Journal of Human Computer Studies 47 (1), 1997.

Conférences

IEEE International Conference on Intelligent Processing Systems

ACM-SIGMOD International Conference on Management of Data (SIGMOD)

http://www.research.att.com/conf/sigmod99

International World Wide WEB Conference (accessible à partir de la page du W3C).

Ateliers

Intelligent Agents — Agent Theories, Architectures, and Languages.

International Workshop ATAL

CIA — COOPERATIVE INFORMATION AGENTS — Learning, Mobility and Electronic Commerce for Information Discovery in the Internet

http://www.informatik.tu-chemnitz.de/~klusch/cia98.html

Éléments de Bibliographie

T. Catarci, D. Nardi and G. Santucci,, WAG : Web-At-A-Glance, International Journal of Intelligent and Cooperative Information, Systems, 7 (2 & 3) : 187--214, 1998.

T. Joachims, D. Freitag and T. Mitchell,, WebWatcher : A Tour Guide for the World Wide WEB,, Proceedings of IJCAI'97, 1997.

C. Leloup, Moteurs d'indexation et de recherche, Eyrolles, Paris, 1997.

H. Lieberman,, Letizia : An Agent That Assists WEB Browsing, Proceedings of IJCAI'95, pages 924-929, 1995.

D.L. McGuinness,, Ontological Issues for Knowledge-Enhanced Search,, Formal Ontology in Information Systems, N. Guarino editor,, IOS Press, Amsterdam, pages 302-316, 1998.

R.Murch, T.Johnson, Intelligent Software Agents, Prentice Hall 99.

M.Stefik, Internet Dreams, Archetypes, Myths and Metaphores, The MIT Press 1996.

D. Wood,, Standard Generalized Markup Language : Mathematical and Philosophical, Issues, Computer Science Today — Recent Trends and Developments,, Lecture Notes in Computer Science 1000, J. van Leeuwen editor,, pages 344-365, Springer, Berlin, 1995.

C.A. Welty,, Intelligent Assistance for Navigating the WEB,, Proceedings of the 1996 Florida AI Research Symposium, 1996.

CONCEptual indexing, querying and Retrieval Of digital documents.

L'objectif de CONCERTO est le développement d'un environnement intelligent (KESE, *Knowledge Engineering Software Environment*) pour l'indexation, la recherche et l'interrogation de documents électroniques. Les documents ainsi traités peuvent être stockés sur différents supports, et en particulier sur le WEB. Plus précisément, les fonctions d'indexation et de recherche sont fondées sur la notion d'annotation conceptuelle : une annotation conceptuelle constitue une description succincte du contenu sémantique d'un document électronique donné, rédigée dans un langage formel de représentation des connaissances — NKRL dans notre cas, voir ci-dessous — que l'on associe au document d'origine au moyen d'une interface spécialisée. Dans le cas de documents textuels par exemple, une annotation conceptuelle ne se limite pas à répertorier les mots-clefs, concepts ou groupes nominaux que l'on peut retrouver dans ces documents, mais elle indique aussi les relations logico-sémantiques qui existent entre ces éléments.

La représentation des annotations conceptuelles dans le langage NKRL

CONCERTO utilise NKRL pour créer les annotations conceptuelles qui représentent le contenu sémantique des documents d'origine. NKRL (*Narrative Knowledge Representation Language*) [2] est un langage développé par G.P. Zarri qui se distingue d'autres langages de représentation des connaissances par l'utilisation parallèle d'une ontologie de concepts et d'une ontologie d'événements. On y représente sous forme de *templates* des classes d'événements du type déplacer un objet, recevoir une information, résider dans un endroit, se comporter de façon amicale ou hostile envers quelqu'un, conclure un accord, se trouver dans une situation donnée, etc.

Un événement particulier (une occurrence) est représenté ainsi par une instance d'un *template* : il comporte un prédicat, une liste de rôles et les arguments du prédicat introduits par ces rôles. Les arguments peuvent être formés d'une entité NKRL simple (concept ou instance de concept) ou complexe (associations de concepts ou d'instances). Plusieurs événements peuvent être reliés entre eux sous la forme de *binding templates*. Des modaux et des indications de temps et de lieu peuvent aussi être associés à une occurrence.

Chaque événement type (chaque *template*) de l'ontologie des événements indique quel prédicat doit lui être associé ainsi que les rôles qui doivent être remplis ou non, les contraintes particulières de temps, de modalité, d'appartenance à la hiérarchie des concepts des *fillers* de l'un ou l'autre des rôles. Une contrainte classique par exemple, est de demander que le SUJET d'un *template* donné soit une personne ou une entité sociale. Ce type de représentation du contenu des documents offre ainsi de larges possibilités d'interrogation.

L'adoption du format RDF/XML

L'un des enjeux du développement d'outils sur le WEB est celui de la portabilité. Pour la transmission des données, des connaissances, des règles d'inférence etc. entre les différents modules du KESE, CONCERTO a adopté RDF/XML, le nouveau standard pour la représentation des métadonnées développé par le consortium W3C [3]. En adoptant RDF, le projet CONCERTO adopte ainsi un format de représentation des métadonnées qui devrait pouvoir être reconnu et utilisé par d'autres systèmes. Dans le projet, les *templates*, les occurrences ainsi que les concepts et leurs instances sont représentés sous le format RDF/XML. Cela a nécessité, bien évidemment, la redéfinition complète des structures de données NKRL selon les caractéristiques de la norme RDF/XML, voir [5].

Les deux types de documents textuels considérés dans le projet et les deux prototypes

Le projet CONCERTO prévoit le traitement de deux types de documents textuels différents et la mise en place de deux applications. La première consiste dans le traitement de résumés en langage naturel extraits des bases de documents textuels de PIRA International, l'un des deux utilisateurs pilotes (et coordinateur du projet). PIRA attend des annotations conceptuelles associées aux résumés qu'elles permettent une indexation à la fois riche et souple des documents d'origine, bien adaptée aux besoins de l'utilisateur. PIRA prévoit, par exemple, de pouvoir visualiser les résumés ou des parties de ces résumés de plusieurs façons différentes en fonction de leur contenu en information (contenu sémantique).

L'autre application concerne l'automation, du moins partielle, des procédés d'exploration récursive d'Internet qui sont propres au deuxième utilisateur pilote de CONCERTO, ASSETT. Il s'agit de la recherche de documents apparaissant sur le WEB dans le domaine des secteurs émergeant de la biotechnologie, destinés à être transmis aux spécialistes et aux décideurs du domaine. Le but de cette application est l'identification des compagnies les plus importantes dans le domaine des biotechnologies et la détermination de leurs aires de compétence, de leurs produits, des brevets, des accords avec les autres compagnies, etc. Après leur annotation, un noyau initial de documents Internet va ainsi servir de guide pour la recherche de nouvelles informations sur le secteur, au moyen d'un processus itératif et récursif.

Conclusion

CONCERTO constitue un projet d'environnement complet pour l'indexation, la recherche et l'interrogation de documents électroniques. Il intègre des outils développés dans différents cadres pour ce type d'applications. En adoptant le nouveau format XML/RDF pour la représentation de ressources Internet, il répond aux nouvelles normes pour la représentation des métadonnées sur le WEB.

Références

[1] Zarri, G.P., Bertino, E., Black, B., Brasher, A., Catania, B., Deavin, D., Di Pace, L., Esposito, F., Leo, P., McNaught, J., Persidis, A., Rinaldi, F. and Semeraro, G. (1999) « CONCERTO, An Environment for the 'Intelligent Indexing, Querying and Retrieval of Digital Documents », in Foundations of Intelligent Systems — Proceedings of 11th International Symposium on Méthodologies for Intelligent Systems, ISMIS'99. Berlin : Springer-Verlag, *à paraître*.

[2] Zarri, G.P. (1997) « NKRL, a Knowledge Representation Tool for Encoding the 'Meaning'of Complex Narrative Texts », Natural Language Engineering 3, 231-253.

[3] Lassila, O., and Swick, R.R., eds. (1999) Resource Description Framework (RDF) Model and Syntax Specification — W3C Recommendation (http://www.w3.org/TR/REC-rdf-syntax/).

[4] www.alphaworks.ibm.com/formula nsf/toolpreview

[5] Jacqmin, S., and Zarri, G.P. (1999) Representation of the NKRL Data Structures According to the XML/RDF Norm (Concerto Technical report NRC-TR-5). Paris : CAMS.

DIAM SERVICE D'INFORMATIQUE MÉDICALE (AP-HP)

DIAM – Service d'Informatique Médicale DSI/AP-HP
CHU Pitié-Salpêtrière,
91, boulevard de l'Hôpital,
75634 PARIS Cedex 13.
Tél : 01-45-83-67-28,
Courriel : diam@biomath.jussieu.fr,
Page d'accueil :
http://www.biomath.jussieu.fr/

Jacques Bouaud, Vincent Brunie, Jean Charlet, Sawsan El Kassar, Brigitte Séroussi, Pierre Zweigenbaum.

Objectifs, cadre général

Au sein de ses activités, l'équipe DIAM du SIM étudie comment l'ingénierie des connaissances (IC) et les techniques de traitement automatique des langues qui sont liées contribuent à la définition et à la mise en œuvre du système d'information hospitalier. Dans ce contexte, nous développons un certain nombre de travaux et réflexions sur l'utilisation de techniques d'Intelligence Artificielle pour accéder à l'information médicale, que ce soit au niveau de la modélisation de « guides de bonnes pratiques » sous forme hypertextuelle (projet ONCODOC) ou pour la recherche d'informations dans un dossier médical (Intranet – projet HOSPITEXTE). Dans ces approches, le *WEB* est le paradigme d'accès à l'information choisi, en particulier pour son adéquation aux approches documentaires et ses capacités à instrumenter une navigation hypertextuelle.

Projets

OncoDoc

Les variations de pratiques thérapeutiques observées au sein des acteurs du monde de la santé témoignent aujourd'hui de la disparité de formation des médecins aux nouveaux traitements. Dans un souci d'optimisation de la qualité des soins, la synthèse de « guides de bonnes pratiques » (GBP) fondés sur la médecine factuelle ou *evidence-based medicine* vise à harmoniser les pratiques. Il s'agit donc de mettre à la disposition des cliniciens un savoir, l'état de l'art, qui puisse être mis en application dans leur pratique quotidienne. Si la diffusion (au sens strict) de ces GBPs a d'abord été l'enjeu majeur que la numérisation et la télématique ont en partie résolu, leur appropriation et leur respect par les médecins restent des problèmes en suspens.

Actuellement élaborés sur le mode strictement textuel, et diffusés sur les réseaux télématiques (internet-intranet), ces GBP ne répondent pas aux besoins des utilisateurs comme le confirment les mauvais taux d'observance relevés. Pour pallier ce fait, des travaux visent à implémenter les GBPs sous la forme de systèmes à base de connaissances. Toutefois, leurs implémentations sont confrontées à des problèmes de formalisation. Les difficultés proviennent des connaissances des GBPs qui s'avèrent souvent incomplètes, de l'utilisation de notions médicales qualifiées d'imprécises et par conséquent de l'influence du contexte sur leur interprétation. Aussi, les modélisations proposées comme outils d'aide à la décision automatiques ne sont-elles pas validées par la pratique.

À mi-chemin entre texte et formalisation, tirant partie des technologies du *WEB* pour la diffusion et des enseignements et de l'intelligence artificielle pour la représentation des connaissances, OncoDoc est une base de connaissances développée dans une logique documentaire. Les connaissances du domaine sont représentées sous la forme d'un arbre de décision dont les branches sont sémantiquement exclusives et exhaustives. Mais l'automatisation du raisonnement médical n'est pas recherchée. Au contraire, les connaissances génériques sont « opérationnalisées » en situation clinique par le médecin. Sa lecture de la base de connaissances en une navigation hypertextuelle à travers la réplique HTML de l'arbre de décision, implémenté en SGML, lui permet d'interpréter en contexte les informations qui lui sont présentées. Il construit ainsi de façon interactive la meilleure représentation de son patient et aboutit ainsi à la proposition de recommandations thérapeutiques centrées-patient.

Appliqué au traitement du cancer du sein, le système a été évalué par une expérimentation en vraie grandeur sur site pilote, le Service d'Oncologie Médicale Pitié-Salpêtrière, démontrant un taux d'observance significativement supérieur aux meilleurs chiffres publiés.

Hospitexte

À la suite du projet DOME, qui avait permis de valider auprès des médecins une conception hypertextuelle de l'accès à l'information, le projet HOSPITEXTE a développé l'informatisation du dossier médical dans une approche d'« ingénierie documentaire ». En effet, puisque les informations médicales sont indissociables de la structure documentaire qui les contient, il importe de disposer, d'une part, du moyen de représenter la structure pour la manipuler et, d'autre part, du moyen de visualiser cette structure et ses contenus. SGML a été retenu comme norme de structuration des documents (chaque type de document respecte une structure d'organisation ou Définition de Type de Documents, la DTD), et HTML comme norme de visualisation ou de formatage. Un prototype a été réalisé et il est testé par des médecins dans le cadre d'une étude de cas sur 10 dossiers médicaux réels. Deux thèmes de recherche du projet sont mis en avant ici de par leur intérêt dans le cadre de ce dossier. Ils font chacun l'objet d'une thèse qui sera soutenue en 1999 :

- Utilisation du raisonnement à partir de cas (RàPC) afin de proposer des stratégies de parcours de l'hypertexte à des praticiens inexpérimentés. Ici,

on va considérer qu'un praticien expérimenté a les moyens de se servir efficacement de l'hypertexte et de mettre en œuvre ses stratégies de parcours. Il s'agit alors de mémoriser ces « bonnes » stratégies de parcours dans des « cas » spécifiques. Il s'agit de trouver la structure de cas et les critères de généralisation ou d'appariement adéquats qui permettront de rendre les services prévus en fonction de critères médicaux (buts du praticien, information recherchée, etc.) et hypertextuels (taxinomie des documents, sessions de « parcours » hypertextuels, etc.).

- Réflexions sur la sémantique hypertextuelle et son instrumentation (recherche encadrée par B. Bachimont – INA). Le dossier informatisé est un hypertexte, c'est-à-dire un ensemble de nœuds (ici, des éléments documentaires ou textuels) reliés par des liens. Le principe de la sémantique hypertextuelle est que la position d'un nœud d'information dans le réseau constitué des nœuds et des liens conditionne et détermine la signification du contenu du nœud. La sémantique hypertextuelle doit alors expliciter des règles interprétatives prescrivant quels parcours à travers l'hypertexte peuvent constituer des lectures, c'est-à-dire faire sens. Ces parcours sont alors instrumentés en documents d'aide à la navigation comme une table des matières.

Références

BOUAUD J., SÉROUSSI B. & ANTOINE E.-C. (1999). OncoDoc : modélisation et « opérationnalisation » d'une expertise thérapeutique au niveau des connaissances. In Actes d'IC'99. À paraître.

BOUAUD J., SÉROUSSI B., ANTOINE E.-C., GOZY M., KHAYAT D. & BOISVIEUX J.-F. (1998). Hypertextual navigation operationalizing generic clinical practice guidelines for patient-specific therapeutic decisions. Journal of the American Medical Informatics Association, 5 (suppl), 488-492.

BRUNIE V., BACHIMONT B. & MORIZET-MAHOUDEAUX P., (1999). Modélisation des connaissances structurelles documentaires pour la conception d'un dossier médical hypertextuel. In *Ingénierie des Connaissances, évolutions récentes et nouveaux défis*. J. CHARLET, M. ZACKLAD, G. KASSEL et D. BOURIGAULT, éds., Eyrolles, 1999.

CHARLET J., BACHIMONT B., BRUNIE V., ZWEIGENBAUM P. & BOISVIEUX J.-F. (1998). Hospitexte : towards a document-based hypertextual electronic medical record. *Journal of the American Medical Informatics Association*, 5 (suppl).

EL KASSAR S. & CHARLET J. (1997). Représentation de connaissances et aide à la navigation hypertextuelle à partir de cas. *In Actes des 1ères Journées Ingénierie des Connaissances*, pages 387 – 401, Roscoff, France, 20 – 22 mai, 1997.

SÉROUSSI B., BOUAUD J. & ANTOINE E.-C. (1999a). Enhancing clinical practice guideline compliance by involving physicians in the decision process. In Proc of AIMDM'99, Aalborg, DK. À paraître.

SÉROUSSI B., BOUAUD J., ANTOINE E.-C., GOZY M., FERRARI C., KHAYAT D. & BOISVIEUX J.-F. (1998). Hypertextual navigation operationalizing decision-oriented guidelines to manage breast cancer patient. In Proc. 34th ASCO Annual Meeting.

SÉROUSSI B., BOUAUD J., ANTOINE E.-C., KHAYAT D. & BOISVIEUX J.-F. (1999b). Aide à la prescription par la consultation télématique de « bonnes pratiques thérapeutiques » dans une démarche centrée-patient. In L'informatisation du cabinet du futur, volume 11, p. 47-58. Springer-Verlag France.

ENSMSE-CENTRE SIMADE-DÉPARTEMENT RIM

Intégration des technologies de l'hypertexte et des réseaux sémantiques

École Nationale Supérieure des Mines de St-Etienne
Département RIM (http://rim.emse.fr/)
158 Cours Fauriel
42023 Saint-Etienne Cedex 2
FRANCE

Composition de l'équipe (e-mail : NOM@emse.fr) :
Membres permanents (6) : Jean-Jacques GIRARDOT, Annie CORBEL, Michel BEIGBEDER, Philippe JAILLON, Roland JEGOU, Jean-François CHAMBON

Doctorants : Bich-Liên DOAN, Yamina ZELLOUF, Fernando AGUIAR, Jérôme HABRANT

Cadre général :

« Information » est un des mots clefs du prochain millénaire. L'information est la matière première de l'avenir, ce que tout un chacun, décideur, ingénieur ou chercheur devra savoir gérer pour simplement survivre dans son domaine. Fournir, obtenir, maintenir, créer et vendre de l'information — en particulier, sous forme électronique, opérations préfigurées par ce qui se passe aujourd'hui avec l'Internet sont des problématiques importantes des métiers de l'information de demain.

Objectifs :

Notre équipe s'intéresse à l'intégration des technologies de l'hypertexte et des réseaux sémantiques dans une vision pédagogique. Notre but est de concevoir et de développer des outils qui puissent répondre de façon plus satisfaisante aux besoins des apprenants. Il est indispensable de tenir compte des besoins des utilisateurs lors de la phase de conception et de développement d'un outil informatique. Actuellement, les serveurs WEB n'offrent que peu de possibilités pour associer aux documents bruts (textes, images, sons) des informations sémantiques de qualité suffisante. Il en est de même pour les outils de recherche qui utilisent des mécanismes trop primaires pour prendre en compte toute la subtilité des documents et des relations existant entre eux. Nous nous intéressons plus particulièrement à l'aspect navigation au sein de la base de documents.

Le premier but est de fournir un outil répondant mieux que les serveurs WEB actuels aux besoins évoqués ci-dessus. Nous nous intéressons à deux aspects en particulier :

- La recherche de documents. L'utilisateur, qu'il soit enseignant, élève, chercheur ou simple quidam, doit pouvoir demander et obtenir aisément des documents compréhensibles pour lui dans le domaine qui l'intéresse. Ceci implique que l'outil de recherche tienne compte de certains aspects sémantiques des documents dont il dispose, ainsi que de certaines caractéristiques de l'utilisateur (le profil).
- La création de documents. Les groupes pédagogiques, actifs, doivent pouvoir créer et rendre continuellement accessibles aux autres utilisateurs de nouveaux documents. Ceux-ci doivent pouvoir s'intégrer continuellement au sein des éléments préexistants. Là encore, c'est en se basant sur une description sémantique et technique précise, fournie par le créateur ou l'inventeur du document, que le serveur va pouvoir insérer le nouvel item dans l'organisation établie.

Le second but de ce travail vise à concevoir, en collaboration avec une équipe s'intéressant à la didactique des sciences, et avec plusieurs équipes d'enseignants, une application en « vraie grandeur » sous la forme d'un serveur spécialiste d'un domaine particulier, et offrant les services décrits ci-dessus dans ce domaine.

Nous nous appuyons sur les réseaux sémantiques qui permettent une représentation riche et économique à partir de 3 notions très simples : notion de graphe où les concepts sont modélisés par des nœuds et les relations entre concepts modélisés par des arcs. Ces trois notions nous semblent parfaitement adaptées au développement de nouveaux outils de recherche et de navigation à la fois puissants et conviviaux.

Publications :

BL. Doan, M. Beigbeder, Virtual WEB Documents : a concept to explicit the structure of WEB sites, in proceedings of the 21st Annual Colloquium on IR Research, 19-20 April 1999, Glasgow, Scotland.

J.Habrant, A.Corbel, JJ.Girardot, « Utilisation des réseaux sémantiques pour la navigation dans l'hypertexte », Colloque Multimédia et Construction des Savoirs, Besançon, France, mai 99.

BL. Doan, M. Beigbeder, JJ. Girardot, Ph. Jaillon, A model of a self-configuring organization to improve information retrieval on the WEB, KRIMS, II (Second International Workshop on Knowledge Representation for Interactive Multimedia Systems), Trento, Italy, juin 1998

BL. Doan, M. Beigbeder, JJ. Girardot, Ph. Jaillon, Using metadata to improve organization and information retrieval on the WEB, WebNet98, Orlando, USA, Novembre 1998.

INRIA RHÔNE-ALPES PROJET SHERPA

655 avenue de l'Europe,
38330 Montbonnot Saint Martin, France
http://exmo.inrialpes.fr
contact : Jérôme Euzenat
(Jerome. Euzenat@inrialpes.fr)

Échanges de connaissances structurées médiatisés par ordinateur

Le projet Sherpa a conçu et développé des modèles de représentation de connaissances à objets qui ont été utilisés pour construire plusieurs bases de connaissances, en particulier dans des domaines scientifiques comme la biologie moléculaire [3] ou les mémoires techniques (voir Bulletin de l'AFIA n° 36).

L'un des principes guidant notre travail est que l'intérêt d'une représentation formelle et explicite des connaissances est non seulement de résoudre des problèmes mais surtout d'élaborer ces connaissances en les formalisant, en les manipulant et en leur appliquant des techniques liées à sa structure formelle (par exemple, la classification ou la catégorisation).

Au sein de la composante consacrée aux « Échanges de connaissances structurées médiatisés par ordinateur (EXMO) », trois types d'activité liés à Internet sont développés :

- la mise en œuvre et l'édition de serveurs de connaissances,
- le développement d'un protocole de construction collaborative de bases de connaissances,
- l'intégration de connaissances formelles dans des documents XML.

Depuis 1994, notre logiciel de représentation de connaissances par objets TROEPS est utilisable en tant que serveur de connaissances à l'aide d'un client HTTP [2, 5]. Cela signifie que n'importe quelle base développée en TROEPS est accessible à travers le WEB sans développement supplémentaire. Qui plus est, les mécanismes de manipulation de connaissances (classification, filtrage, inférence de taxonomies…) sont disponibles interactivement — en fait toutes les pages HTML sont engendrées à la volée.

Les bases de connaissances peuvent être reliées directement à d'autres sites WEB, à des bases de données ou à un lexique du domaine. Au prix de quelques adaptations, il est possible d'insérer au sein des pages engendrées par TROEPS des annotations HTML, des Applets, des appels à d'autres sites fondés sur le contenu de la base ou de redéfinir complètement l'allure des pages engendrées. Par ailleurs, comme toutes les adresses utilisées par TROEPS sont normalisées (elles ne dépendent que du contenu et de la localisation de la base) il est possible de faire appel aux serveurs de connaissances depuis un site externe. Enfin, les serveurs de connaissances sont extensibles de manière simple ce qui permet d'intégrer à son serveur des fonctionnalités liées à une application particulière (par exemple, la détection de cycles dans les graphes d'interaction génique [3]).

TROEPS permet également aux personnes autorisées d'éditer directement les bases de connaissances (voir Bulletin de l'AFIA n° 34).

Le protocole de CO4 (pour Construction coopérative de bases de connaissances consensuelles) est une bibliothèque logicielle qui peut être connectée à un système de représentation d'un artefact à concevoir (à TROEPS en tant que système de représentation de connaissances). L'originalité principale du protocole est d'encourager la collaboration entre utilisateurs en imitant la soumission d'articles à une revue scientifique, c'est-à-dire en promouvant une évaluation par les pairs des modifications soumises à une base consensuelle [1]. La différence essentielle est que les soumissions sont exprimées dans un langage formel et que les bases doivent être consistantes. Afin de rendre ceci possible, les utilisateurs ne modifient pas directement la base de connaissances consensuelle mais travaillent sur un espace de travail personnel et, lorsqu'ils considèrent avoir atteint un état publiable, soumettent leurs modifications à la base consensuelle à laquelle ils adhèrent. Le protocole prend en charge cette soumission et organise un appel à commentaires entre l'ensemble des adhérents à la base consensuelle. Le résultat peut être l'un des suivants : soit tous les adhérents acceptent la modification et celle-ci est intégrée dans la base consensuelle, soit un des adhérents refuse la modification et celle-ci est rejetée, soit un des adhérents fait une contre-proposition en en acceptant une partie et en remplaçant une autre partie, auquel cas un appel à commentaire modifié est émis.

Le protocole de CO_4 a été formellement spécifié et prouvé analytiquement consistant, équitable (fair), vivace (live), terminant et surtout consensuel [4]. Cette spécification a été traduite en LOTOS et analysée par des techniques de « model-checking ». Elle a ensuite été implémentée comme une librairie [6] et connectée à TROEPS.

Afin d'organiser une base documentaire, il peut être utile de représenter les connaissances du domaine et les connaissances sur un document particulier dans un formalisme de représentation de connaissances. L'une des utilisations qui peut en être faite est la représentation du contenu des documents à des fins d'indexation et de recherche : il est alors possible de poser des requêtes sur le contenu formalisé des documents. L'avènement du langage XML permet d'intégrer dans chaque document, disponible sur un serveur, la représentation de son contenu (ce que nous appelons « inTexation Formelle »). Le langage de représentation TROEPS peut être exporté vers et importé d'XML grâce à la DTD correspondante et cette intégration peut être réalisée. Ainsi, si le contenu d'un document est représenté par une structure XML correspondant à un ou plusieurs objets TROEPS, le document pourra être classé dans une ou plusieurs taxonomies du domaine et l'on pourra retrouver des documents similaires.

Ce type d'utilisation d'une représentation de connaissances est amené à se développer, mais est assez peu étudié pour l'instant. Nous avons mis sur pied, avec les projets INRIA ACACIA et ORPAILLEUR, l'action ESCRIRE dont le but est de comparer, sur une étude de cas, les avantages et les limitations de trois types de représentation de connaissances — graphes conceptuels, représentation par objets et logiques de descriptions — dans ce cadre.

Ces travaux sont actuellement mis en œuvre dans plusieurs contextes : mémoire d'entreprise (contrat SERICS « autoroutes de l'information » SToRIA, avec STMicroelectronics), traçabilité des modèles (Génie II, avec Dassault aviation), base de connaissances sur les interactions géniques (ACC « sciences de la vie » du CNRS, avec LGPD-CNRS), gestion de documentation technique (ARC ESCRIRE, avec les projets ACACIA et ORPAILLEUR de l'INRIA).

Références

[1] Jérôme Euzenat, Building consensual knowledge bases : context and architecture, dans Nicolaas Mars (éd.), Towards very large knowledge bases, IOS press, Amsterdam, pp143-155, 1995

[2] Jérôme Euzenat, Knowledge bases as WEB page backbones, Actes 5th WEB workshop on « artificial intelligence-based tools to help W3 users », Paris, 1996

[3] Jérôme Euzenat, Christophe Chemla, Bernard Jacq, A knowledge base for D. melanogaster gene interactions involved in pattern formation, Actes 5th international conference on intelligent systems for molecular biology, Halkidiki (GR), pp108-119, 1997

[4] Jérôme Euzenat, A protocol for building consensual and consistent repositories, Rapport de recherche 3260, INRIA Rhône-Alpes, Grenoble, 1997.

[5] Projet Sherpa, Troeps 1.2 reference manual, Rapport interne, INRIA Rhône-Alpes, Grenoble, 1998. http://co4.inrialpes.fr/docs/tropes-manual.html

[6] Projet Sherpa, Co4 protocol 1.0 reference manual, Rapport interne, INRIA Rhône-Alpes, Grenoble, 1998

INRIA SOPHIA-ANTIPOLIS ACTION ACACIA

Contact : Rose Dieng,
Mél. : Rose.Dieng@sophia.inria.fr,
Tél. : 04 92 38 78 10,
Télécopie : 04 92 38 77 83

NDLR : ce projet ayant déjà été présenté dans les numéros 35 (suite du dossier « Ingénierie des connaissances ») et 36 (dossier « mémoires d'entreprises »), nous y renvoyons le lecteur…

INRIA SOPHIA-ANTIPOLIS ACTION AID

2004, route des Lucioles BP 93
06902 Sophia-Antipolis
http ://www.inria.fr/aid/
Contact : Brigitte TROUSSE
Mél : trousse@sophia.inria.fr

Présentation générale :

L'action de recherche AID de l'INRIA Sophia-Antipolis poursuit ses recherches sur les modèles de raisonnement en intelligence artificielle en vue d'aider au travail coopératif sur Internet/intranet (cf le dossier IA et Telecom du bulletin 33 (avril 1998)). Les domaines actuels d'applications de l'action AID sont d'une part la recherche collaborative d'informations sur le WEB et, d'autre part la conception de sites WEB adaptatifs et évolutifs (ou plus généralement de systèmes d'informations hypermédia adaptatifs et évolutifs).

Problématique :

Notre problématique se situe actuellement d'une part au niveau des recherches sur le raisonnement à partir de cas et son utilisation dans le cadre de systèmes de recommandations et d'autre part au niveau de la gestion des connaissances pour l'aide à la conception de systèmes complexes (sites WEB, systèmes d'informations hypermédia, mémoire collective).

Tout d'abord un système de recommandation peut être décomposé en trois entités de base : le groupe d'agents *producteurs* de recommandations, le module de *calcul de recommandations* et le groupe de *consommateurs* des recommandations. Un défi majeur dans le domaine de la conception de systèmes de recommandation est le suivant :

« Comment produire des recommandations *personnalisées* et de haute *qualité* tout en *minimisant l'effort* requis des producteurs et des consommateurs ».

Deux grandes approches complémentaires sont proposées dans la littérature : 1) l'approche fondée sur l'apprentissage automatique de profils utilisateurs et 2) l'approche fondée sur des techniques de fouille de données.

Dans l'action AID, nous travaillons à développer une approche de calcul de recommandations centrée sur la fouille de données où les comportements passés d'un groupe d'utilisateurs sont utilisés pour calculer les recommandations (cf. aussi filtrage collaboratif). La plupart des autres approches fondées sur la fouille de données sont principalement des approches statistiques où l'ordre d'occurrence d'événements dans l'historique n'est pas pris en compte lors du calcul de recommandation. Les problèmes difficiles pour la mise en œuvre de notre approche concernent les aspects suivants :

1. fournir des techniques d'identification et d'extraction de comportements pertinents (i.e., des enseignements ou des cas) à partir des données brutes des historiques,
2. définir des méthodes et des techniques de mesure de similarités entre comportements,
3. définir des techniques d'inférence de recommandations personnalisées à partir des comportements pertinents passés identifiés (ou à partir des cas remémorés).

Nous abordons les trois problèmes ci-dessus en explorant la possibilité d'application des techniques RàPC [1,2,3,5].

Deuxièmement nous étudions comment des concepteurs ayant des points de vue différents (référentiels thématiques ou isotopies i.e., leur propre ontologie locale) arrivent à partager des connaissances. Nous entendons par « ontologie » des systèmes pour classer des entités dans un univers de discours ; dans notre cas, celui-ci s'appuie sur les isotopies référencées par les agents. À partir de cela, nous étudions comment prendre en compte la notion de point de vue pour l'aide au travail coopératif sur le WEB et l'indexation d'expériences passées en conception (cf. mémoire d'entreprise).

Nos recherches actuelles visent, dans le cadre d'un contrat GIS Sciences de la Cognition en conception, à évaluer Hermès [7], notre système d'argumentation pour l'aide à la structuration d'une discussion sur le WEB pour la confrontation de points de vue en conception de sites WEB institutionnels et à y intégrer la notion de point de vue.

Principales contributions

Nos principaux résultats concernent les trois axes suivants :

- au niveau RàPC, la définition d'un modèle d'indexation par situation comportementale [5] et au niveau RàPC & modélisation à objets, la conception et la réalisation en Java d'une plate-forme à objets en RàPC appelé CBR*Tools utilisant les patrons de conception pour la conception des modèles objets [4]
- au niveau systèmes de recommandations, la proposition d'une nouvelle approche de recommandations, appelée Broadway, fondée sur nos résultats en RàPC [2] et l'application de cette approche à deux types de problèmes liés à la recherche d'informations sur le WEB [1] : la navigation sur le WEB (Cf. Broadway V1) [5] et la reformulation de requêtes (Broadway_QR et BeCBKB) [3]
- au niveau ingénierie des connaissances, appliquée à la conception de systèmes complexes, la proposition d'un modèle d'analyse de corrélations de points de vue dynamiques issue d'une approche topique (sémantico-logique et qualitative) [6] et la conception d'un système d'argumentation d'aide à la décision collective, Hermès [7].

Le lecteur intéressé pourra se reporter à nos rapports d'activité de ces trois dernières années [http://www.inria.fr/aid/ra.html]

Nos partenaires principaux sur ces sujets sont le CNET, le Centre de recherche de Rank Xerox à Meylan (France) et le CNES (Toulouse). Tous les logiciels de l'action AID sont conçus avec l'atelier Rose de Rational et implantés en Java : ils s'adressent soit à des concepteurs d'applications dites coopératives avec la spécification et la réalisation de deux frameworks orientés objet pour l'aide au travail coopératif sur le WEB (Broadway*Tools et CBR*Tools) soit à des utilisateurs finaux (cf. Broadway-V1, Broadway-QR + BeCBKB et Hermès). Voir http://www.inria.fr/aid/software.html

Applications

Broadway-V1 [5] est un assistant pour la navigation sur le WEB réutilisant les navigations passées d'un groupe d'utilisateurs. Broadway observe les navigations de différents utilisateurs, récolte les évaluations et les annotations de ces utilisateurs pour établir une liste de documents conseillés.

Broadway-QR [3] est un nouveau service issu de l'application de l'approche Broadway à la reformulation de requêtes dans un quelconque métamoteur de recherche. L'utilisation de similarités entre requêtes est l'une des approches classiques d'aide à la reformulation de requêtes. Différentes mesures de similarités ont été proposées et comparées dans la littérature. Cependant aucune de ces mesures ne prend en compte les positions relatives des requêtes comparées dans les processus de la reformulation. C'est cette nouvelle mesure de similarité que nous proposons à travers l'application de l'approche Broadway pour le problème de reformulation de requêtes. En outre notre approche Broadway-QR peut être utilisée comme un optimiseur des autres approches classiques de reformulation de requêtes. Des expérimentations de cette approche pour l'optimisation automatique des requêtes sont planifiées. Un premier prototype de l'application de l'approche Broadway-QR pour l'aide à la reformulation de requêtes est développé. Ce prototype, appelé moteur de recherche d'information CBKB est développé à XRCE. Dans ce premier prototype nous avons considéré le cas simple des requêtes constituées uniquement d'une liste de mots clés. Les résultats de cette étude sont rapportés dans [3].

Hermès [7] est un système d'aide à l'argumentation, à la structuration d'une discussion voire à la prise de décision collective sur le WEB. Hermès apporte une aide plus puissante que celle des systèmes actuels existants qui proposent uniquement des forums de discussions hiérarchiques. Pour cela, le système d'argumentation Hermès offre des mécanismes de raisonnement qui contrôlent la discussion. De plus il permet l'expression de différents arguments individuels (ou préférences) et de les noter dans le but d'aider à la sélection d'un certain choix. Hermès vise des coopérations sur le WEB entre agents travaillant ensemble dans des lieux différents et à un même instant.

Conclusions/Perspectives

L'effort dans le domaine du WEB qui fut un choix du projet depuis fin 96 devrait encore s'amplifier les années à venir : en effet notre objectif principal actuellement vise à proposer des méthodes et des outils pour aider à l'utilisation et à la conception au niveau

sémantique de sites Internet/intranet adaptatifs et évolutifs, cela, en nous appuyant à la fois sur des techniques d'intelligence artificielle et de génie logiciel. Nos travaux sur le raisonnement à partir de cas et les systèmes de recommandations constituent une première base pour l'analyse de l'usage de tels sites WEB et pour une évolution facilitée de ces sites. Nos travaux pluridiciplinaires sur l'aide à la conception de sites Internet/intranet sont en cours.

Bibliographie

(http://www.inria.fr/aid/publications.htm)

[1] B. Trousse, M. Jaczynski, R. Kanawati, Une approche fondée sur le raisonnement à partir de cas pour l'aide à la navigation dans un hypermédia, In proceedings of Hypertexte & Hypermedia : Products, Tools and Methods (H2PTM'99), Paris, 1999.

[2] B. Trousse, M. Jaczynski, R. Kanawati, Using User Behavior Similarity for Recommandation Computation : The Broadway Approach, In proceedings of 8th international conference on human computer interaction (HCI'99), Munich, August, 1999.

[3] R. Kanawati, M. Jaczynski, B. Trousse, J-M. Andreoli, Applying the Broadway Recommandation Computation Approach for Implementing a Query Refinement Service in the CBKB Meta-search Engine Conférence française sur le raisonnement à partir de cas (RàPC'99), Palaiseau, France, juin, 1999.

[4] M. Jaczynski, B. Trousse, Patrons de conception dans la modélisation d'une plate-forme à objets pour le raisonnement à partir de cas. L'Objet, Numéro spécial sur les patrons orientés objets, D. Rieu et J-P. Giraudon (Guest editors), 5 (2), 25 pages, Hermès, 1999.

[5] M. Jaczynski, « *Modèle et plate-forme à objets pour l'indexation des cas par situations comportementales : application à l'assistance à la navigation sur le WEB* », Thèse de doctorat de l'Université de Nice Sophia Antipolis, 1998.

[6] B. Trousse, « *Viewpoint Management for Cooperative Design* », In Proceedings of the IEEE Computational Engineering in Systems Applications, Nabbeul, Tunisie, Avril 1998.

[7] N. Karacapilidis, D. Papadias, Hermes97 « Hermès : supporting Argumentation Discourse in Multi-Agent Decision », in Proceedings of the AAAI-98 Conference, AAAI/MIT press, p. 827-832, 1998, http://www.inria.fr/aid/hermes/.

INSTITUT INTERFACULTAIRE D'INFORMATIQUE DE NEUCHÂTEL

Université de Neuchâtel
Pierre-à-Mazel 7
2000 Neuchâtel (Suisse)
tél : +41 36 718 1375
télécopie : +41 36 718 1231
http://iiun.unine.ch

* Prof. J. Savoy, directeur institut interfacultaire d'informatique
Jacques. Savoy©seco.unine.ch
* Dr A. Le Calvé, maîtrise Université de Rennes, Dr ès Sciences Université de Neuchâtel (1999)
Anne. LeCalvé©seco.unine.ch
* J. Picard, ingénieur, maîtrise École polytechnique de Montréal (1996)
Justin. Picard©seco.unine.ch

Cadre général et objectifs

Afin de concevoir un système de recherche d'informations adapté aux hypertextes et à Internet, six critères doivent être pris en compte : (1) le système de dépistage doit être fondé sur des sources multiples d'évidence, c'est-à-dire permettre le recours à des techniques d'indexation différentes, à des formulations de requêtes diverses et à des stratégies de repérage variées ; (2) les liens hypertexte (ainsi que d'autres liens inter- ou intra-documents) doivent être utilisés afin de retrouver davantage de documents répondant aux souhaits de l'usager et de déterminer de bons points de départ pour la navigation ; (3) le besoin d'information doit être facile à exprimer par l'usager. Notre objectif se situe dans le cadre de techniques générales (4) qui peuvent s'appliquer dans tous les domaines du savoir ; (5) qui sont indépendantes de la langue utilisée et (6) qui ne requièrent pas un important travail manuel. Ainsi, nous excluons les systèmes de recherche dédiés et adaptés à un domaine très étroit pour lesquels un important travail manuel est requis (création d'un réseau sémantique, constitution d'un thesaurus spécialisé, élaboration d'une base de règles d'un système expert).

Dans cette optique, la problématique centrale de nos recherches consiste à modéliser, à implémenter et à tester la meilleure stratégie permettant de combiner différentes sources d'évidence d'une part, et d'autre part, de tenir compte des liens hypertextes. Afin d'atteindre notre but, nous souhaitons vérifier la validité de nos approches à l'aide de collections tests comme la collection des « Communications of the ACM » (3204 documents, 50 requêtes) ainsi que d'autres corpus tirés des expériences DISPTER-DARPA TREC ou, pour des documents en français, issus des expériences Amaryllis (Action de Recherche Concertée de l'AUPELF-UREF).

Problèmes abordés et travaux réalisés

Dans le cadre général défini ci-dessus, nos travaux ont démontré l'intérêt de disposer d'un système intelligent d'apprentissage afin que le système automatique de dépistage de l'information puisse améliorer sa performance avec le temps. Spécialement adapté aux systèmes hypertexte, notre système d'apprentissage permet d'améliorer significativement la performance moyenne du dépistage [Savoy 1994].

Les liens entre documents ou entre pages WEB peuvent fournir une indication précieuse sur le contenu sémantique des textes sous-jacents. Ainsi, notre étude [Savoy 1997] indique que le repérage automatique de l'information peut bénéficier de la prise en compte des liens hypertextes au coût d'une augmentation très faible du temps de réponse.

Si la langue anglaise a été étudiée de manière intensive depuis plusieurs années, le français ne pose pas de problèmes très particuliers. Une analyse morphologique du français est certes possible pour aider la recherche d'informations [Savoy 1993], mais des procédures relativement simples permettent d'obtenir des performances intéressantes [Savoy 1999] (procédures pouvant s'intégrer aisément dans le cadre de bibliothèques virtuelles).

Cependant, la nature distribuée des réseaux locaux ou à grande distance pose quelques difficultés supplémentaires. Ainsi, il convient de choisir les n meilleures banques textuelles en fonction de la requête (problème de sélection des sources), leur soumettre la requête (par exemple, en recourant au standard

Z90.50), récupérer n listes de documents retrouvés, fusionner ces n listes pour en donner une seule à l'usager (problème de fusion de collections) et supprimer les duplicatas (ou les quasi-duplicatas comme, par exemple, le rapport technique et l'article correspondant paru dans un journal scientifique). Le Calvé & Savoy [1999] proposent de résoudre le problème de la fusion de collections au moyen d'une méthodologie fondée sur la régression logistique. La qualité des réponses ainsi obtenues s'avère significativement supérieure aux techniques habituelles.

Finalement, nos travaux portent également sur les mécanismes automatiques de raisonnement et d'inférence sous incertitude. Dans ce cas, la logique classique et la théorie des probabilités peuvent être combinées pour former des Systèmes d'Argumentation Probabiliste (SAP). À l'aide de ce formalisme permettant de traiter, de gérer et de combiner des sources incertaines sur le contenu sémantique des documents, Picard [1998] démontre comment on peut modéliser les liens hypertextes dans le cadre d'un nouveau modèle logique en recherche d'informations.

Sélection d'articles sur le sujet

• J. Savoy : Stemming and Stop-List for General French Corpora. Journal of the American Society for Information Science, 1999, to appear.

• A. Le Calvé, J. Savoy : Database Merging Strategy Based on Logistic Regression. Information Processing & Management, 1999, to appear.

• J. Picard : Modeling and Combining Evidence Provided by Document Relationships Using Probabilistic Argumentation Systems. Proceedings of the ACM-SIGIR'98 Conference, Melbourne (Australia) August 1998, p. 182-189.

• J. Savoy : Ranking Schemes in Hybrid Boolean Systems : A New Approach. Journal of the American Society for Information Science, 48 (3), 1997, 235-253.

• J. Savoy : A Learning Scheme for Information Retrieval in Hypertext. Information Processing & Management, 30 (4), 1994, 515-533.

• J. Savoy : Stemming of French Words Based on Grammatical Category. Journal of the American Society for Information Science, 44 (1), 1993, 1-9.

IRIT — TOULOUSE ÉQUIPE SMAC

Équipe SMAC
Pierre Glize
IRIT Université Paul Sabatier
118, Route de Narbonne
31062 TOULOUSE Cedex
Tel. 05 61 55 61 76
Mél : glize@irit.fr
URL : http://www.irit.fr

Pierre Glize, Marie-Pierre Gleizes, André Machonin

Cadre général, objectifs

Gérer et obtenir l'information de sources de données dispersées géographiquement est un problème difficile. C'est le cas du World Wide WEB où les sources d'information sont très dynamiques et autonomes, rendant ainsi tout contrôle global inadéquat. Nos travaux ont pour objectif d'utiliser les techniques des systèmes multi-agents auto-organisateurs coopératifs pour mettre en œuvre des outils d'aide à la recherche d'une ressource parmi un nombre de sites très important.

Motivations, problématiques considérées, travaux et réalisations informatiques

La problématique qui guide nos travaux porte sur la possibilité (ou l'impossibilité !) pour des systèmes artificiels de s'adapter à des situations véritablement imprévues. « Véritablement imprévues » signifie que le concepteur n'a en aucune manière programmé le comportement du système face à de telles situations. Pourquoi cette problématique ? Nous considérons que, même si l'homme développe des méthodes et des outils de conception, il restera toujours des bogues car les machines artificielles ont des potentialités énormes et croissantes, les réseaux multiplient leur puissance intrinsèque et les environnements de ces systèmes peuvent être tellement dynamiques que tout n'est pas concevable *a priori*. La réponse que nous apportons pour spécifier de tels systèmes est fondée sur l'auto-organisation par coopération qui garantit théoriquement la convergence vers la fonction globale désirée par le concepteur bien qu'aucune des parties qui le composent ne la connaisse.

C'est ainsi qu'Internet en tant que « système » au sens habituel, devrait fournir en permanence toutes les sources pertinentes — et uniquement celles-ci — à chaque requête utilisateur. Notre expérience quotidienne montre que nous en sommes actuellement très éloignés car même avec les moteurs de recherche par indexation ou bien par catalogue, les réponses à une recherche d'information ne couvrent pas l'ensemble des sites pertinents et comportent beaucoup de bruit. Cette incapacité à gérer Internet à un niveau global comme un système unique est tolérée car expliquée par sa taille, sa distribution, son hétérogénéité et sa dynamique. Nous pensons que cela n'est pas une fatalité, si nous l'abordons comme un système auto-organisateur à fonctionnalité émergente.

Applications à des problèmes concrets

Arcadia (ARchitecture pour la Coopération d'Agents d'Information Autonomes)

Le projet ARCADIA, qui s'intéresse à la gestion et à l'obtention d'informations de sources dispersées géographiquement, s'inscrivait dans le cadre d'un contrat entre le CNET, le CERT et l'IRIT de janvier 1995 à décembre 1997.

Chaque usager et chaque fournisseur de service sont représentés par un agent particulier appelé « agent représentant ». Son architecture est constituée de trois niveaux de systèmes multi-agents auto-organisateurs :

- Le niveau supérieur d'Arcadia est un système multi-agent de sites communiquant par envoi de messages.
- Chaque site est un système multi-agent de niveau intermédiaire contenant des « agents représentants » interagissant en son sein.
- Pour que les interactions dans chacun des deux niveaux s'effectuent de manière pertinente, les agents raisonnent en fonction de connaissances privées qu'ils possèdent les uns sur les autres : des croyances. Chaque réseau de croyances privées est un système multi-agent du niveau inférieur.

Nous avons développé un prototype dans le domaine des petites annonces de véhicules d'occasion, dont les trois niveaux de multi-agents ont été développés en C++.

Abrose (Agent Based Brokerage SErvices in electronic commerce)

Abrose est un projet européen qui a débuté en Mars 1998 et se finira en Janvier 2000. Il comporte huit autres partenaires (Deutsche Telekom Berkom, InfoMures SA, France Telecom CNET, Universidad Polytechnica de Madrid, Onyx Ltd., Dégriftour, National Technical University of Athens, Technical University of Berlin).

C'est un outil de commerce électronique de médiation de services fondée sur des agents. Une place de marché électronique est constituée d'une collection importante de fournisseurs et de clients qui ont des offres et des besoins très dynamiques. Chacun d'eux est représenté par un agent et l'approche par auto-organisation coopérative garantit une satisfaction individuelle et collective optimale alors qu'il n'existe jamais de connaissance globale de l'état du système. C'est ainsi que le système possède un certain nombre de fonctionnalités, notamment :

- La publicité ciblée : lorsqu'un fournisseur de services souhaite diffuser une offre, seuls les agents représentant des usagers dont les besoins sont conformes à celle-ci reçoivent l'information.
- La concurrence totale : un client reçoit du système la liste exhaustive des fournisseurs qui ont des produits répondant à la description de son besoin.
- L'apprentissage permanent : les agents représentants adaptent automatiquement leurs croyances tant sur les nouveaux produits des fournisseurs que sur les nouveaux besoins des usagers.

Les interactions entre agents sont conformes aux normes Fipa et le logiciel est écrit en Java. D'autres informations peuvent être trouvées sur le site (http://www.b5.berkom.de/abrose/)

Courte bibliographie

• Camps V., Gleizes M.P., Principes et évaluation d'une méthode d'auto-organisation, Troisièmes journées francophones sur l'Intelligence Artificielle Distribuée & Systèmes Multi-Agents, 15-17 mars 1995, St Baldoph, Savoie (pp. 337-348)

• Camps Valérie, Vers une théorie de l'auto-organisation dans les systèmes multi-agents basée sur la coopération : Application à la recherche d'information dans un système d'information, Thèse de l'Université Paul Sabatier N. 2890 - 1998

• Camps Valerie, Glize Pierre, Gleizes Marie-Pierre, Leger Alain, Athanassiou Eleutherios, Lakoumentas Nikos — A Framework for Agent Based Information Brokerage Services in Electronic Commerce – EMSEC Conference – 1998

• Gleizes Marie-Pierre, Leger Alain, Athanassiou Eleutherios, Glize Pierre — Abrose : Self-Organisation and Learning in Multi-Agent based Brokerage Services – IS & N Conference, 1999

• Glize Pierre, Gleizes Marie-Pierre, Léger Alain — Brokerage communication in a cooperative multi-agent based mediation service : one example in Abrose — Foundation for Intelligent Physical Agent CFP6_016 - avril 1999

LIUM — LE MANS ÉQUIPE EIAO

Équipe EIAO
LIUM — Institut Informatique
Claude Chappe
UNIVERSITÉ DU MAINE
72085 Le Mans Cedex 9
Tél. : 33 2 43 83 3858 —
Télécopie : 33 2 43 83 3868

Responsable : Martial VIVET, professeur
Enseignants — Chercheurs : Pascal Leroux, Philippe Teutsch
Doctorants : Christophe Després, Sébastien Georges, Paulo Pavel, Rousanna Bedanokova, Omar Gueye, Catherine Bonvalot

Cadre général & objectifs

Le cadre général est celui de la conception d'environnements d'apprentissage pour l'éducation et la formation en mettant en œuvre l'ingénierie de la connaissance au sens IA du terme. L'objectif essentiel est de développer ce que nous appelons les environnements d'apprentissage distribués dans le cadre de la réalisation de dispositifs de formation à distance à travers INTERNET. Nous nous intéressons fondamentalement au déplacement vers INTERNET des activités généralement identifiées dans la communauté IA & éducation (cf. AIED society)

Activités

Problématiques considérées

Nos problématiques concernent plusieurs aspects complémentaires faisant l'objet de travaux séparés.

1 - architectures de systèmes d'apprentissage à distance : nous travaillons à la conception de systèmes multi-agents distribués sur le WEB avec une famille d'agents sur la machine cliente apprenant, une famille d'agents sur la machine client du maître-tuteur, disponible sur le réseau et d'une famille d'agents sur le serveur. Ces diverses familles sont fondées sur le même modèle et les mêmes protocoles de communication et permettent la communication entre les agents humains concernés. Ce travail est à distinguer de ce qui peut être fait en matière de campus virtuels qui intègrent les dimensions gestion de groupes d'apprenants, gestion de scolarité, gestion de ressources de communication synchrone ou asynchrone au sens usuel.

2 - modélisation d'utilisateurs : nous travaillons à la modélisation d'utilisateurs au sens où on l'entend en EIAO quand on parle de modèle d'élèves. En particulier, des travaux sont en cours sur ce que N. Balacheff a appelé modèle cognitif d'apprenant, qui renvoie un modèle en terme de connaissances acquises par l'apprenant dans le domaine de connaissances concerné (par opposition au modèle comportemental qui ne renvoie que les activités de l'apprenant sur l'interface sans en faire l'analyse ou l'interprétation en termes de connaissances). Nous travaillons là en particulier sur la distribution du modèle d'utilisateur en composants répartis sur le réseau (par exemple, mémoire court terme sur le client, mémoire long terme sur le serveur).

Travaux réalisés

En mathématique

Paulo PAVEL dans sa thèse - à soutenir à l'automne 1999 - a travaillé à la réalisation d'un système (VISU@L — GD) d'apprentissage à distance de la géométrie descriptive pour des architectes brésiliens. Il a implanté un système multi-agent distribué permettant le tutorat à distance. L'élève est engagé dans une série d'activités mettant en œuvre des outils spécifiques de manipulation d'épures sous CABRI-Géomètre. Il peut communiquer en cas de besoin avec le précepteur distant qui de son côté dispose d'un véritable tableau de bord permettant le contrôle d'apprenants connectés. Les apprenants peuvent communiquer entre eux.

En technologie

Christophe DESPRES et Sébastien GEORGES travaillent à l'implantation d'un autre modèle multi-agents pour mettre en liaison des apprenants aidés d'un « assistant pédagogique logiciel local » entre eux ou avec un tuteur disponible sur le réseau. La communication intègre la prise de contrôle du poste apprenant à distance avec discussion possible à travers la visio conférence (net meeting pour l'instant). Le suivi des activités par le maître peut se faire par téléchargement d'une trace des activités faites en local sous le contrôle de l'assistant logiciel. L'analyse de cette trace par des méthodes d'IA permet de préparer un résumé destiné au maître de ce qui s'est passé en « son absence ».

En français

Catherine BONVALOT travaille à l'analyse de réponses à des questions ouvertes dans un système d'apprentissage du français langue étrangère. Il s'agit de faire un modèle cognitif d'apprenant pour assurer un diagnostic d'erreurs chez l'utilisateur concerné et à terme un guidage dans le déroulement des activités. Les apprenants seront répartis sur INTERNET et la construction de tels modèles de connaissances à partir des interactions de l'usager avec le système relève fondamentalement de l'IA.

Pour sa part, Omar GUEYE travaille sur la communication parlée dans les environnements d'apprentissage du français à distance à travers INTERNET. Il se focalise sur la compréhension et l'expression orale et au-delà des seuls aspects « apprentissage des langues », il étudie plus particulièrement l'intérêt d'une modalité orale dans la conception d'EIAO en général.

Rouzanna BEDANOKOVA travaille sur l'apprentissage du français des affaires et aborde la distribution du modèle d'utilisateur en composants répartis sur le réseau (composants sur le client, composants sur le serveur). Ces composants ont des degrés d'élaboration et une durée de vie différents.

Réalisations informatiques

- le système VISU@L GD de Paulo PAVEL est un prototype opérationnel, testé dans des classes au Brésil en novembre 1998.
- le système multi-agents développé par C. Després et S. Georges est opérationnel en laboratoire mais n'a pas encore fait l'objet de test externe.
- le système GCSE (General Certificate for Secondary Education) qui permet de tester les connaissances en français pour de jeunes anglais est opérationnel sur INTERNET (tests en cours). Le modèle cognitif développé par C. Bonvalot n'est pas encore disponible.

Applications à des problèmes concrets

Les applications qui nous concernent touchent la formation à distance dans toutes ses dimensions : formation en zones rurales -PMI/PME-, formation sur réseaux INTRANET de grandes entreprises dont les composantes/agences sont dispersées, formation dans le domaine de la francophonie (nous disposons au Mans de didacticiens spécialistes du français langue étrangère), formation initiale ou continue des enseignants dans les IUFM le plus souvent multi-sites (anciennes écoles normales), formation à l'étranger (cf. projet RESAFAD pour le LIUM en Afrique).

Bibliographie

• BEDANOKOVA R., TEUTSCH Ph. « GCSE Project for Distance Evaluation of Knowledge in French Foreign anguage », IFIP WG 3.3 & 3.6, « The virtual Campus : trends for Higher Education & Training », Madrid, November 27-29, 1997.

• BEDANOKOVA R., VIVET M. « Distributed User Models for Client/Server Architectures », Soumis à AIED '99.

• DESPRES C., LEROUX P. « Raisonner sur la trace : Analyse de sessions avec l'application ROBOTEACH », Actes des Cinquièmes journées EIAO de Cachan chez Hermès, 14-16 mai 1997, p.227-288.

• GEORGE S., LEROUX P., VIVET M. « Computer Aided learning In industrial Machine Scheduling », 4th International Conference on Computer Aided Engineering Education, CAEE'97, Krakow, Pologne, sept. 11-13, 1997, p. 70-77.

• LEROUX P., VIVET M. « Micro-robots based learning environnments for continued education in SMEs », International Journal of continuing Engineering Education and life long learning, special issue on « Intelligent Systems/tools in training and life-long learning », accepted to appear in 1999

• LEROUX P., VIVET M. « De la modélisation d'un processus de coopération à la conception de systèmes coopératifs d'apprentissage », ITS '96, In : C. Frasson, G. Gauthier and A. Lesgold eds., *Intelligent Tutoring Systems — Third International Conference*, ITS'96, Springer Verlag, Lecture Notes in Computer Science 1086, Montréal, 12-14 juin, 1996, p. 197-205.

• LEROUX P., VIVET M., BREZILLON P. « Cooperation between a Pedagogical Assistant, a Group of Learners and a Teacher », Proceedings of the European Conference on Artificial Intelligence in Education, Lisbon, sept. 30th, Oct. 2, 1996. Brna P., Paiva A., Self J. eds, Editions Colibri, Lisbon. ISBN 972-8288-37-9, p. 379-385.

• LEROUX P., VIVET M., BREZILLON P. « Cooperation between Humans and a Pedagogical Assistant in a Learning Environment », Second International Conference on the Design of Cooperative Systems (COOP'96), Juan-les-Pins, France, June 12-14, 1996, p. 391-404.

• PAVEL P., RIBEIRO POLA M.C. Spatial Visualization and Multimedia, ED-MEDIA '98, Freiburg, June 1998, p. 1791-1792.

• PAVEL P., RIBEIRO POLA M.C. « Solitary : An Educational Computer Game for the Learning of Descriptive Geometry », 8th International Conference on Engineering Computer Graphics and Descriptive Geometry -ICECGDG-, Austin, Texas, July 31, August 3, 1998. p. 309-313.

• PAVEL P., RIBEIRO POLA M.C., VIVET M. « Direct Manipulation of Working Drawings in Descriptive Geometry Learning by Computers », 8th International Conference on Engineering Computer Graphics and Descriptive Geometry -ICECGDG-, Austin, Texas, July 31, August 3, 1998. p. 608-612.

• PAVEL P., VIVET M., BALACHEFF N. « Impacto de Descoberta Dirigida Telepresencial nos Ambientes Interactivos de Aprendizagem », 8 Simposio Brasileiro de Informatica na Educaçao, Sao José dos Campos, 18 a 20 de novembro de 1997, p. 675-686.

• TEUTSCH PH. « Un modèle de situation d'évaluation pour le suivi de formation en langue étrangère ». ITS '96. In C. Frasson, G. Gauthier and A. Lesgold eds., *Intelligent Tutoring Systems — Third International Conference ITS'96*. Springer Verlag, Lecture Notes in Computer Science 1086, Montréal, 12-14 juin, 1996, p. 315-323.

• TEUTSCH Ph., CHANIER T., CHEVALIER Y., MANGENOT F., NARCY J.P. (1996). « Environnements interactifs pour l'apprentissage en langue étrangère ». Atelier au 3ème Colloque Hypermédias et Apprentissage, Chatenay-Malabry, mai 1996, p. 247-256.

LORIA — NANCY ÉQUIPE LANGUE ET DIALOGUE

LORIA, UMR 7503
Bâtiment LORIA — Campus Scientifique — B.P. 239
F-54506 Vandœuvre-lès-Nancy Cedex
http://www.loria.fr/equipes/led/

Jean-Marie Pierrel, Laurent Romary, Patrice Bonhomme, Samuel Cruzlara

Depuis plusieurs années déjà, l'équipe *Langue et Dialogue* (LED) du LORIA, utilise l'Internet comme vecteur de valorisation de ses recherches et développements fondés principalement sur le traitement automatique de la langue naturelle. LED, dont la thématique de recherche est la communication, les interfaces et le dialogue homme-machine à forte composante langagière, s'est nettement impliquée sur l'accès et la dissémination de ressources linguistiques sous format électronique. Pour ces besoins, et ceux de ses travaux, l'équipe LED a mis un fort accent sur la notion de normalisation des données et ressources linguistiques qui représentent une source générique d'information linguistique pour la spécification de ses modèles théoriques et leur validation par l'observation. Plus globalement, la réflexion menée autour de ces aspects de codage et de normalisation fondée principalement sur des normes et recommandations internationales, doit être partagée par l'ensemble de la communauté concernée. C'est à ce titre d'ailleurs, que nous avons été amenés à participer à un certain nombre de projets et d'actions nationales et internationales.

Les domaines concernés sont orientés vers la gestion de ressources linguistiques mono et multilingues et leur distribution par le biais du WEB, mais couvrent également un grand nombre de secteurs d'activité allant de la terminologie multilingue maritime et océanographique jusqu'à la télé-médecine, en passant par l'indexation de site WEB et la mise en réseau de serveurs de ressources linguistiques.

Des alignements multilingues sur le WEB

Cette compétence tire son origine des développements effectués dans le cadre d'un projet Lingua dont le but était l'élaboration et la gestion d'un corpus de textes parallèles multilingues et la mise au point d'un système d'alignement multilingue pour l'apprentissage d'une seconde langue étrangère. L'ensemble des spécifications, repose sur une utilisation importante de la norme SGML et des directives de la TEI *(Text Encoding Initiative)* pour le codage des textes et des alignements. Pour répondre à une demande importante, une démonstration du concordancier multilingue est accessible sur le WEB offrant une navigation parmi les textes d'un corpus multilingue et une interface d'interrogation pour afficher en parallèle un ensemble de phrases répondant à une requête et leur traduction.

Des ressources linguistiques avec Silfide

Silfide (Serveur Interactif pour la Langue Française, son Identité, sa Diffusion et son Étude), élaboré à partir d'un projet débuté en janvier 1996 sous l'égide de l'AUPELF-UREF et du CNRS, est une plate-forme d'échange et de diffusion de ressources linguistiques primaires (textes, lexiques essentiellement francophones) pour des utilisateurs de secteurs assez diversifiés (linguistes, enseignants, informaticiens, etc.). Silfide fournit à chaque utilisateur un espace de travail convivial qui lui permet de se constituer par requêtes successives ou navigation itérative des corpus de travail symbolisés par des paniers. Ces paniers contiennent des références aux ressources linguistiques du serveur et ne nécessitent donc aucun téléchargement des données. Le groupe de travail Silfide, complète régulièrement l'espace de travail, par des outils linguistiques de type concordancier mono et multilingue, statistique lexicale, étiqueteur morpho-syntaxique, etc.

Afin de dynamiser l'architecture de Silfide, et de lui assurer une évolution tant au niveau des contenus textuels que des outils, la totalité des développements est fondée sur une utilisation importante des nouvelles technologies de l'information et de la communication comme le langage Java, la recommandation XML et CORBA. Par exemple, Silfide utilise la norme SGML (maintenant la recommandation XML) pour le codage des ressources linguistiques mais se sert également d'XML pour le codage des messages transitant (requêtes, résultats, information utilisateur…) entre le serveur et les clients.

Vers une société de l'information multilingue

Sur la base des spécifications de Silfide et de son architecture technique, et dans le cadre des appels d'offre Européens MLIS (MultiLingual Information Society), deux propositions ont été retenues.

ELAN (European Language Network) concerne la mise en réseau de serveurs Silfide spécialisés, dans une langue ou une catégorie de ressources, et la définition d'un langage de requête commun pour l'interrogation de contenus langagiers. Un utilisateur est référencé sur un seul serveur du réseau ELAN et a accès, de manière transparente, à la totalité des ressources et outils présents sur les serveurs constituant le réseau. Par l'intermédiaire d'un navigateur WEB classique, l'utilisateur se connecte et s'identifie auprès d'un serveur local, charge un client graphique (en fait une Applet Java), choisit les serveurs qui l'intéressent et constitue ses corpus de travail. Il peut ensuite, effectuer des requêtes sur le contenu qui seront répercutées par le serveur local vers les serveurs présélectionnés.

DHYDRO (Dictionnaire Hydrographique) vise la définition, sur Internet, d'un espace terminologique multilingue (Français, Anglais et Espagnol dans un premier temps) spécialisé dans le domaine de l'hydrographie. Ce forum, offre des fonctionnalités (1) de consultation sur site des différentes versions du dictionnaire hydrographique international, (2) de mise à jour en ligne des différentes entrées du dictionnaire validées par les rédacteurs autorisés, (3) d'un espace interactif de discussion entre utilisateurs et (4) de consultation d'extraits de textes de références pour y repérer les termes utilisés dans les dictionnaires.

Indexation de site WEB et accès au contenu textuel

L'indexation contrastive consiste à connaître les termes pertinents d'un texte en le comparant à un ensemble de textes plus vaste mais couvrant un domaine d'expertise similaire. Nous avons appliqué cette méthode, provenant de la statistique lexicale et utilisée en lexicologie, à l'indexation d'un site WEB. L'architecture ICSW (Indexation Contrastive de Site WEB), soutenue par la DGLF et réalisée en partenariat avec l'INRIA, permet (1) à l'administrateur d'un site WEB de préparer une cartographie thématique de son site en utilisant le moteur d'indexation contrastive et de catégorisation et (2) à un utilisateur de naviguer parmi cette cartographie pour retrouver plus rapidement les informations qui sont pertinentes pour lui.

Surveillance de patients dialysés à domicile

La maîtrise de technologies clés telles que Java et XML, a permis à LED de participer activement à la spécification et aux développements de Diatelic, un système de surveillance de personnes dialysées à domicile. Par son orientation médicale, ce système s'adresse à des utilisateurs potentiels, malades et médecins, qui ne sont pas *a priori* des spécialistes de l'usage de l'outil informatique. Ses interfaces améliorées et conviviales, permettent d'une part une acquisition rapide et intelligente des informations, côté malade et d'autre part une présentation et une description d'alertes éventuelles, côté médecin. Les alertes, relatives à l'état de santé du malade, sont déclenchées par un système expert qui a été spécialement développé pour le projet. Pour se connecter sur le serveur WEB dédié à Diatelic, un simple ordinateur personnel relié à l'Internet et équipé d'un navigateur WEB classique suffit aux patients et au médecin pour dialoguer et échanger des informations. Outre l'aspect économie de santé publique, le système procure au patient un nouveau confort en le sécurisant davantage. Diatelic assure au patient, un suivi supplémentaire de la part de son médecin quasiment en temps réel, ce qui a pour conséquence de diminuer les temps de prise de décision en cas de problème important, nécessitant, par exemple, une hospitalisation immédiate de la personne dialysée.

Références

P. Bonhomme and L. Romary. Managing multilingual texts for educational purpose. In *Actes des Quinzièmes Journées Internationales IA 95*, Montpellier, 1995. EC2 & Cie.

P. Bonhomme, S. Cruzlara, J.-M. Pierrel, and L. Romary. Silfide : Serveur interactif pour la langue française, son identité, sa diffusion et son étude. *TALN'97*, 1997.

L. Romary, P. Bonhomme, F. Bruneseau et J.M. Pierrel (1999). Silfide : A System for Open Access and Distributed Delivery of TEI Encoded Documents. Computers and the Humanities, vol. 33, No 1 & 2, pp 31-38.

P. Bonhomme and L. Romary (1997a). A contrastive indexing method and its integration in aquarelle folders. In *3rd DELOS Workshop : Cross-Language Information Retrieval*, ERCIM Workshop Proceedings European Research Consortium for Informatics and Mathematics.

LORIA — NANCY ÉQUIPE ORPAILLEUR

UMR LORIA 7503
Bâtiment LORIA — Campus Scientifique — B.P. 239 -
F-54506 Vandœuvre-lès-Nancy Cedex
URL : http://www.loria.fr/LORIA/EXT/equipes/ORPAILLEUR

Rim Al Hulou, Fairouz Chakkour, Amedeo Napoli, Emmanuel Nauer, Yannick Toussaint.

Cadre général, objectifs

Le but de ces recherches est d'étudier la combinaison de la technologie des systèmes à bases de connaissances et de la technologie du WEB ; ceci afin de fournir une aide efficace lors de la consultation du WEB. Notre approche consiste plus précisément à utiliser des données structurées d'un domaine (références bibliographiques, thesaurus, etc.) pour construire des connaissances sur ce domaine (réseaux d'auteurs, vocabulaire employé par tel ou tel auteur, etc.). Ces connaissances sont alors exploitées pour la recherche ou le filtrage d'information sur le WEB [1]. L'accès aux données du WEB est réalisé via les moteurs de recherche classiques (AltaVista, Excite, etc.) qui sont utilisés comme des outils distants ; la connaissance construite sert à guider l'utilisateur :

- en amont des moteurs de recherche pour la détermination de requêtes, pour l'expression du besoin, etc. ;
- en aval pour l'évaluation et une présentation plus intelligible des documents.

Motivations, problématiques considérées, travaux réalisés et réalisations informatiques

Motivations. La maîtrise de l'accès à l'information dans un fonds volumineux et hétérogène tel que le WEB représente un enjeu majeur pour les consommateurs d'information (chercheurs, entreprises, etc.). Les moteurs de recherche sont débordés par l'explosion du WEB et ne satisfont plus aux tâches de recherche d'information. Depuis quelques années des agents dits *« intelligents »* ont vu le jour pour tenter de pallier les problèmes rencontrés par les moteurs. Ces agents mettent en place certaines heuristiques et combinent des techniques issues de différents domaines tels que la linguistique, les bases de données, la recherche d'information, les statistiques, etc.

Nos travaux s'attellent au même objectif : la construction de systèmes intelligents.

Problématiques considérées. L'intégration et le traitement d'informations

provenant de sources variées et hétérogènes sont des problèmes préalables à la construction de systèmes intelligents (systèmes pour la fouille de données sur le WEB en particulier). À partir de ces données hétérogènes, nous souhaitons modéliser le domaine (déterminer les concepts du domaine, les liens entre les concepts et le vocabulaire attenant aux concepts) en utilisant un formalisme de représentation capable d'effectuer des raisonnements sur ces données. De ce fait, il est indispensable de maîtriser l'hétérogénéité du vocabulaire utilisé et de sa portée sémantique [2]. La construction d'un vocabulaire de base est donc un point essentiel dans notre approche. Le fait que les données soient hétérogènes et plutôt non régulières nécessite aussi de s'intéresser au traitement des données semi-structurées et à l'intégration d'informations provenant de sources différentes.

Un autre enjeu est la production d'une information élaborée qui permette de donner une idée synthétique d'un ensemble de données et qui puisse être utilisée - à l'image des connaissances - dans un raisonnement. Concrètement, la production d'informations élaborées peut se voir comme un processus d'ECBD (Extraction de Connaissances à partir de Bases de Données) qui agit sur des informations scientifiques et techniques (IST). Nous chercherons par exemple à dégager les principaux thèmes de recherche sous-jacents à un corpus de références bibliographiques, ou encore les collaborations entre auteurs, l'émergence d'une technique bien particulière, etc. Nous touchons en cela au domaine de la bibliométrie qui fixe les bases d'exploitation de l'IST. Là aussi, une normalisation minimale des données à exploiter est indispensable pour éviter des biais statistiques [3].

Un dernier enjeu concerne le traitement distribué de l'information dans une architecture client-serveur. Les données sont naturellement distribuées - notamment sur le WEB - sur des sites distants et il est particulièrement important de pouvoir tenir compte et de maîtriser cette distribution. Les problèmes actuels de normalisation, d'échanges de formats et de gestion des accès aux données sont dans une phase préliminaire d'étude.

Travaux réalisés et réalisations informatiques.

Les travaux menés jusqu'à présent concernent l'intégration des données (locales ou distantes, multi-sources et de natures différentes) et la construction d'informations élaborées à partir de ces données. Un système d'investigation sur le WEB a déjà été mis en place. Ce système permet un accès indifférencié aux données locales ou distantes, ainsi que des croisements entre ces données [4]. Concrètement, nous avons développé un ensemble de modules permettant :

1. de réaliser des opérations fondamentales d'intégration de données telles la normalisation des données manipulées, la convergence du vocabulaire, le dédoublonnage des données, etc. ;
2. d'accéder à l'ensemble des données par une interface WEB (génération dynamique de pages HTML).

Application à des problèmes concrets

D'un point de vue pratique, une collaboration avec des chercheurs de l'INRS (Institut National de Recherche et Sécurité) a orienté notre contexte d'application vers le domaine médical. Ce domaine constitue un terrain particulièrement propice aux expérimentations, du fait qu'il s'avère riche en fonds structurés (bases de données, thesaurus, etc.) et en données en ligne. L'utilisation et le croisement de données structurées et hétérogènes pour la construction d'un système intelligent de recherche d'informations ont permis [5] :

1. la structuration du domaine pour un accès hiérarchisé à l'information : des accès thématiques sont construits automatiquement par des méthodes de classification (à partir de descripteurs de références bibliographiques par exemple).
2. la traduction de termes pour un accès multilingue : une traduction automatique du vocabulaire du domaine peut être effectuée via un thesaurus ou encore par des corrélations de descripteurs au travers de références bibliographiques multilingues ;
3. la génération d'un environnement d'investigation spécialisé (et intégré) sur WEB permettant à l'utilisateur :
   - d'être assisté dans l'étape consistant à définir le vocabulaire de la requête à soumettre à un moteur de recherche (pour une recherche d'information sur Internet) ;
   - d'obtenir des compléments d'informations (références bibliographiques locales) sur les documents WEB distants retrouvés ;
4. le filtrage d'information sur Internet : à partir des critères sélectionnés par l'utilisateur, une requête est engendrée automatiquement et est soumise aux moteurs de recherche. Cette requête est déterminée par ajout d'un contexte de recherche (vocabulaire proche), en addition aux critères sélectionnés.

Travaux en cours

Il s'agit maintenant de déterminer quelles connaissances vont permettre l'émergence de documents pertinents. Ces connaissances serviront à favoriser la recherche d'information sur le WEB (formulation automatique de requête). Elles devront également permettre d'analyser (valider, rejeter, juger, etc.) et de classer les documents, proposés en réponse par les moteurs de recherche. Dans ce but, la classification et le raisonnement ont un rôle essentiel à jouer.

Ces travaux prennent actuellement part, dans le cadre de l'action ESCRIRE, à des collaborations avec les projets ACACIA et SHERPA de l'INRIA. Le but de cette action est d'étudier différents types de représentation des connaissances pour la gestion de documents scientifiques et techniques.

Bibliographie

[1] Emmanuel Nauer, Jacques Ducloy et Jean-Charles Lamirel. Using of multiple data source for information filtering : first approaches in the MedExplore project. In *5th DELOS Workshop on Filtering and Collaborative Filtering*, Budapest, Hungary, 10-12 november 1997.

[2] Emmanuel Nauer. Les problèmes de variations terminologiques dans l'indexation de références bibliographiques. In *Journées Internationales de Linguistique Appliquée (JILA'99)*. LILLA — Université de Nice, juin 1999.

[3] Emmanuel Nauer. De l'importance de la normalisation en bibliométrie. In *Les systèmes d'information élaborée*, Ile Rousse, France, 27 septembre-1er octobre 1999, à paraître.

[5] Emmanuel Nauer et Jean-Charles Lamirel. Environnement d'investigation sur WEB : assistance à l'utilisateur par des connaissances fédérées. In Jean-Pierre Balpe, Imad Saleh and Marc Nanard, editors, *Hypertextes et Hypermédias*, pages 101-113, Paris, 1997. Hermès. 4ème Conférence Internationale H2PTM'97, 25-26 septembre 1997.

[4] Emmanuel Nauer. Croisement de multiples sources d'information : utilisation du WEB comme source et interface. In *Actes de la 5ème conférence internationale « Hypertextes et Hypermédias » réalisations, outils & méthodes »*, Paris, France, 23-24 septembre 1999, à paraître.

LRI, UNIVERSITÉ PARIS-SUD ÉQUIPE IASI

http://www.lri.fr/Français/Recherche/grille.html

Alain Bidault, François Goasdoué, Marie-Elizabeth Concalves, Chantal Reynaud, Marie-Christine Rousset, Brigite Safar, Véronique Ventos

Avec l'émergence de l'Internet et de l'Intranet, il est possible d'accéder aujourd'hui à de multiples sources d'informations réparties, hétérogènes et autonomes, qui peuvent contenir des informations pertinentes et complémentaires pour une application.

PICSEL est un système d'intégration de différentes sources d'information développé dans l'équipe IASI du L.R.I (Paris-Sud) en collaboration avec le CNET. PICSEL repose sur la construction d'un *médiateur* regroupant dans un cadre unificateur et homogène la description sémantique et abstraite du contenu des différentes sources d'information. Les données restent stockées et réparties au niveau des sources d'information sous une forme qui peut varier d'une source à l'autre. Le médiateur joue le rôle d'interface entre l'utilisateur et les sources d'information en lui donnant l'impression qu'il interroge un système d'informations centralisé et homogène. Sa tâche est de construire des plans de requêtes combinant l'appel à différentes bases de données spécialisées dans le but de fournir des réponses effectives à la requête de l'utilisateur.

La plupart des systèmes à base de médiateurs reposent sur l'utilisation d'un formalisme déclaratif qui offre un cadre unificateur et homogène de représentation aussi bien du domaine d'application (e.g., le tourisme), que du contenu des sources d'information, que des requêtes que peut poser l'utilisateur. La plupart des systèmes existants prônent le choix d'un formalisme déclaratif associé à une sémantique claire. C'est pourquoi les langages de représentation fondés sur la logique sont très utilisés. Leur pouvoir d'expression doit permettre de traduire des informations sémantiques à la fois fines et structurantes.

La représentation déclarative du contenu des sources d'information se traduit par un ensemble de *vues abstraites*. Une vue abstraite (*vue*, par la suite) d'une source d'information fait abstraction des données effectivement stockées dans la source ainsi que de son format de stockage pour n'en décrire que le contenu sémantique sous la forme des requêtes auxquelles elle permet de répondre. Dans une telle approche, déterminer comment répondre à une requête posée par l'utilisateur au médiateur consiste à reformuler (on dit aussi réécrire) la requête de l'utilisateur en termes de vues caractérisant les requêtes spécialisées pertinentes auxquelles les sources d'information disponibles peuvent répondre.

PICSEL est fondé sur le langage CARIN qui combine Datalog avec les logiques de description. Les propriétés théoriques de ce langage ont été étudiées dans [Levy-Rousset]. Son pouvoir d'expression permet de faire cohabiter des règles et des relations d'arité quelconque avec un réseau sémantique de concepts et de propriétés. Ces concepts et ces propriétés peuvent également intervenir dans les conditions des règles et ainsi participer à la définition de relations ou de requêtes. Par ailleurs, le réseau sémantique peut être automatiquement construit et mis à jour par des algorithmes de subsomption et de classification.

L'utilisation de CARIN et des algorithmes d'inférences, au cœur du moteur de requêtes de PICSEL, est décrite dans [Lattes-Rousset].

Une assistance à la conception de bases de connaissances décrivant les sources d'information dédiées au médiateur PICSEL a été proposée dans [Goasdoué-Reynaud].

L'aide à l'affinement de requêtes dans le cadre de PICSEL est un problème crucial qui a commencé à être étudié dans [Bidault].

Nous commençons également à étudier l'intégration de sources d'informations semi-structurées exprimées en XML.

Références

• [Bidault] A. Bidault, Aide à l'affinement de requêtes posées à un serveur d'informations. Rapport de stage de DEA, Université Paris-Sud, septembre 1998.

• [Goasdoué-Reynaud] F. Goasdoué, C. Reynaud, Modeling Information Sources for Information Integration. Proceedings of European Knowledge Acquisition Workshop EKAW'99. Daghstul, May 1999.

• [Levy-Rousset] A. Levy, M-C Rousset, Combining recursive Horn rules and description logics in CARIN. Artificial Intelligence, Volume 104 (1998), pages 165-209

• [Lattes-Rousset] V. Lattes, M-C Rousset, The Use of CARIN Language and Algorithms for Information Integration : the PICSEL System. Proceedings of the international workshop on Intelligent Information Integration, associated with ECAI'98, Brighton, August 1998.

PSI – INSA DE ROUEN

BP08, Place Emile Blondel
76131 Mont Saint Aignan Cedex
Tel : (33) 235 52 84 70
Télécopie : (33) 235 52 84 41
http://www.univ-rouen.fr/psi/

Contact : Jean-Pierre Pécuchet (pecuchet@insa-rouen.fr)

Participants : Nicolas Delestre, Gaetan Dussaux, Catherine Gréboval-Barry, Françoise Guégot.

Présentation de l'équipe

Le laboratoire PSI (Perception Systèmes Information) fédère sa recherche autour de trois thèmes scientifiques majeurs : modélisation des systèmes intelligents, classification et apprentissage, vision. C'est dans le thème modélisation des systèmes intelligents que trois prototypes utilisent des méthodes et outils d'intelligence artifi-

cielle pour améliorer les possibilités offertes par Internet. Deux de ces travaux ont pour objectif la distribution de cours via Internet (les projets ARCHYMEDIA et METADYNE), alors que le but du dernier projet, est l'aide à la recherche d'information sur la toile.

Projet Archymedia

ARCHYMEDIA « Architecture Hypermédia distribuée pour l'EIAO » est un environnement interactif et hypermédia d'apprentissage par ordinateur. Le but de ce système est de permettre à des enseignants de réaliser des cours hypermédias afin que des apprenants puissent les suivre et les consulter. Fondé sur une approche systémique, il contient trois sous-systèmes : un système acteur, un système d'aide à la décision et un système documentaire.

Projet Metadyne

Avec les STI (Systèmes Tutoriels Intelligents) et les EIAO (Environnements d'Apprentissage assistés par Ordinateur), les hypermédias constituent de nos jours une composante essentielle dans le domaine de la recherche sur les systèmes d'enseignement assistés par ordinateur. Malheureusement, le principal atout de ces systèmes, c'est-à-dire la non-linéarité de la progression de l'apprenant, devient vite un inconvénient majeur en perdant l'apprenant dans l'hyperespace. De ce fait, on a cherché à guider l'étudiant dans son cheminement en fonction de ses connaissances sur le domaine enseigné, en modifiant aussi bien le contenu des pages que les liens entre ces dernières : c'est ce que l'on nomme les hypermédias adaptatifs. Mais les expérimentations ont montré qu'une désorientation de l'apprenant pouvait persister. Ainsi depuis quelques années, les recherches s'orientent vers ce que l'on nomme les hypermédias adaptatifs dynamiques, c'est-à-dire des systèmes qui créent dynamiquement les pages et les liens de l'hyperespace en fonction des caractéristiques de l'utilisateur, en vue de simplifier l'espace exploré et de le réduire à la partie la plus adaptée à l'utilisateur.

L'objectif du projet METADYNE est de concevoir un hypermédia adaptatif dynamique. Bien qu'étant le sujet d'étude de plusieurs recherches, il n'y a à notre connaissance que peu de systèmes réellement multimédia dispensant des cours via le réseau Internet. METADYNE tente donc d'innover en ce sens en proposant des cours multimédias, disponibles quelle que soit la position géographique de l'apprenant. La construction de ces cours, issue d'une mise en commun des connaissances des enseignants, tiendra compte de toutes les caractéristiques de l'apprenant.

Projet d'aide à la recherche d'information

Chaque jour, la quantité d'information disponible en ligne croit inexorablement, en particulier en raison de l'essor des communications électroniques. Le public auquel de telles masses d'informations sont destinées est aussi en constante évolution : le développement rapide de l'Internet en est l'exemple le plus marquant.

Nous sommes donc entrés dans l'ère de l'information où l'utilisateur est submergé par la masse. Grâce à des outils d'aide à la navigation (AltaVista, Excite, Yahoo, Nomade…) qui reposent sur des techniques d'indexation de texte, l'utilisateur sait comment accéder à l'information qu'il cherche, mais doit encore fournir un effort considérable pour éliminer lui-même l'information non pertinente. Il apparaît donc essentiel de disposer d'outils permettant de classer, d'archiver, d'analyser et de diffuser ces éléments d'informations. Nous proposons d'améliorer ces outils en les dotant d'un mécanisme d'apprentissage consistant à indexer les documents selon l'utilisation qui en a été faite par les membres d'un groupe. La démarche consiste à observer les utilisateurs dans leur processus de recherche d'information et à s'appuyer sur ces observations pour fournir un potentiel d'indexation supplémentaire au moteur de recherche à l'aide de techniques à base de cas.

Le système met à disposition de l'utilisateur une sorte de *signet* amélioré dans lequel il pourra déposer les documents qu'il juge intéressants au cours de sa recherche, dans le cadre de ses préoccupations. Grâce à cette collection de documents pertinents, on crée des cas de navigation et on peut ainsi fournir une aide aux autres utilisateurs qui ont des préoccupations similaires. L'idée centrale est d'éliminer progressivement le bruit engendré par les moteurs de recherche reposant sur une indexation plein texte, en constituant collectivement une base de cas. La réussite d'un tel système repose sur une participation massive et rigoureuse des utilisateurs. C'est pourquoi nous allons leur fournir des outils qui vont leur apporter une réelle plus-value :

- une mémoire de navigation, avec la possibilité de classifier, d'annoter, et d'interroger cette mémoire
- une technologie Push.

Bibliographie

F.Guégot, C.Nécaille, « Proposal of an Hypermédia Distributed Architecture for an Interactive/Intelligent Computer Assisted Learning System », Actes ICCE95 pages 190-197, Singapour, Décembre 1995.

F.Guégot, C.Nécaille, « Représentation d'un Métamodèle pour un Système Hypermédia d'Enseignement-Apprentissage », NTICF'98, pp139-146, Rouen1998.

C. Nécaille, F. Guégot, J.P.Pécuchet « ArchyMeDia : un environnement multimédia distribué d'enseignement/apprentissage centré sur l'aide à l'utilisateur », NTICF'98, pp339-344, Rouen 1998.

N. Delestre, C. Gréboval, J.P. Pécuchet, METADYNE, « a Dynamic Adaptive Hypermedia System for Teaching », 3rd ERCIM Workshop User Interfaces for All, pp 143-149, 1997.

N. Delestre, J-P. Pécuchet, « L'architecture d'un hypermédia adaptatif dynamique pour l'enseignement », NTICF'98, pp383-391, Rouen, 1998.

PSYCO — ROUEN ÉQUIPE EMC2

Alain Vom Hofe, Faculté de Psychologie, Sociologie et Sciences de l'Education,
rue Lavoisier, 76821 Mont Saint-Aignan Cedex
http://hometown.aol.com/vomhofeal
VOMHOFEAL@aol.com

EMC2 (Équipe Multimédia, modélisation Cognitive & aides à la Conception) est membre du Laboratoire PSY-CO de l'Université de Rouen. Ce laboratoire a un statut d'UPRES. EMC2 est un membre actif du pôle inter-régional normand (en voie de constitution) « TIC et SHS » qui a vocation à relier des laboratoires d'Informatique et de SHS. L'équipe est composée de : Alain Vom Hofe, Professeur de psychologie cognitive (Université de Rouen) ; Valérie Saint-Dizier, Docteur en psychologie, chercheur sur contrat ; James De Almeida, Docteur en psychologie, chercheur ; Didier Viard (consultant en ergonomie), chercheur associé ; Olivier Guibernao, thésard (Université de Rouen) ; Dominique Guédon, thésard (Université de Rouen) ; Isabelle Vimare, étudiante en 3e cycle

Cadre général et Objectifs

EMC2 s'intéresse à l'acquisition de compétences sur des supports multimédias. Ses principaux objectifs concernent la formalisation de l'activité cognitive de l'utilisateur et du concepteur, l'élaboration d'outils d'assistance au concepteur (notamment aux plans du scénario d'interaction entre utilisateur et système et de la conception de l'interface), et la définition de démarches individualisées d'apprentissage (enseignement/remédiation).

EMC2 appréhende aujourd'hui, tout particulièrement, la problématique de la recherche d'informations et de la navigation dans les bases de données hypermédia « on line » et « off line ». Un développement de ses activités est actuellement en cours sur les hypermédias pour l'enseignement. EMC2 a pour particularité d'engager des interactions permanentes entre la psycho-ergonomie cognitive et l'intelligence artificielle.

Problématiques et travaux

La problématique centrale de nos travaux actuels concerne les profils des utilisateurs. Nous nous interrogeons sur la pertinence de cette notion pour rendre compte de certaines régularités observées dans l'activité des utilisateurs des grandes bases de données hypermédias et du WEB. Nous cherchons tout particulièrement à mettre en évidence le degré de généralité, à travers différentes tâches et pour une large variété de sujets, des grandes dimensions du fonctionnement cognitif lors d'une activité de Recherche d'Informations. Doit-on envisager des modèles « spécifiques » à la tâche ou bien des modèles plus « transversaux » ; doit-on envisager des modèles « généralistes » ou bien des modèles « différentialistes » ?

Un premier programme (ERGO-WEB), financé par le Ministère de l'Education Nationale, de la Recherche et de la Technologie, est actuellement en cours sur l'activité de recherche d'informations sur le réseau Internet. Dans une perspective de conception de systèmes (assistants logiciels interactifs pour la navigation, agents de recherche…) qui prennent en considération l'activité de l'utilisateur ou bien certaines de ses caractéristiques stables, ce programme est consacré à l'élaboration de modèles génériques de tâches et à l'examen de stratégies individuelles en matière d'interrogation par mots-clés et de navigation. Le repérage de la stabilité temporelle et de la (relative) cohérence à travers des tâches différentes, de ces stratégies est réalisé au moyen d'un micro-réseau interne géré par un proxy-cache.

Un second programme, dans le cadre du projet MARA financé par le Ministère de l'Industrie, aborde la question de la recherche d'informations sur un CD.ROM à structure hypertextuelle. L'objectif est de dégager des profils individuels de recherche - dans le cadre des modèles de la décision et de la résolution de problèmes - et de les mettre en relation, d'une part, avec la nature des connaissances dont disposent les utilisateurs, d'autre part, avec plusieurs dimensions de l'efficience cognitive (capacité de planification, capacité de la Mémoire de Travail, Vitesse de traitement de l'information).

Bibliographie

Vom Hofe, A., Mathieu, J. & Viard, D. (1999, à paraître). La prise en compte de l'utilisateur dans les systèmes hypermédia. Pratiques Psychologiques.

UNIVERSITÉ D'OTTAWA (GROUPE TAMALE) RES INTERNATIONAL

Contact ; Jean-François Delannoy,
e-mail : delannoy@res.ca
http://www.res.ca ; http://www.csi.uottawa.ca/tanka/ts.html ;
http://www.csi.uottawa.ca/~delannoy

RES International/Online Direct (Ottawa/Toronto) : fouille de textes appliquée aux catalogues

Patfinder est un projet de recherche de brevets avec rétroaction. L'utilisateur échantillonne des documents et leur attribue une note de un à cinq. Une courte liste des mots-clés les plus discriminants est créée dynamiquement à chaque itération, pour alimenter un réseau de neurones qui reclasse les documents par pertinence en fonction de la distribution de ces mots-clés, faisant « remonter » vers le haut de la liste la « crème » de la collection.

Le projet **Prodfind** est destiné à améliorer les connaissances d'un acheteur institutionnel ou individuel en parcourant des catalogues pour y repérer les descriptions de produits (nom, prix, code, fournisseur…) pour les afficher dans une table ou les archiver en base de données. La technique appliquée est un filtrage tenant compte à la fois du formatage (HTML, et plus tard XML) et des patrons linguistiques du texte lui-même. L'utilisateur peut définir de nouveaux filtres « par l'exemple » : en désignant à la souris certaines zones dans les pages et en précisant le genre d'information qu'elles contiennent. Le système peut insérer des marqueurs XML pour faciliter le traitement ultérieur ou permettre un reformatage sur des critères de contenu. Les applications principales sont la fouille systématique de catalogues pour la mise à jour côté acheteur, et la comparaison d'offres d'un produit donné, mais les techniques peuvent aussi être appliquées à l'acquisition automatique de connaissances en général (par exemple du domaine financier ou politique) dans les textes formatés.

Groupe TAMALE, Université d'Ottawa : résumé automatique ; question-réponse ; marquage et analyse de l'argumentation

Le groupe TAMALE (ex « MaLTe »), déjà présenté dans le bulletin de l'AFIA pour ses recherches en analyse de textes et **résumé automatique**, participe à l'effort d'évaluation en **Question-Réponse** de TREC (voir http://www.research.att.com/~singhal/qa-track.html), qui soulève de nombreux problèmes concrets en sémantique des textes et recherche d'information :

Qu'est-ce qu'une réponse adéquate à une question, dans la communication humaine ou entre humain et machine ? Comment catégoriser utilement les questions, pour fournir une réponse coopérative ? Quel est l'apport d'une analyse de Cas verbaux ? Comment évaluer la qualité de l'information obtenue, par rapport à la réponse ?

Rob Holte, Berry de Bruijn et Joel Martin travaillent à l'amélioration de la **recherche interactive** d'information, en utilisant la rétroaction, les références croisées entre les documents, la modification du texte (dont le résumé automatique), la visualisation, le raffinement de requête, et la modélisation de l'utilisateur et des ses tâches. En particulier, Berry de Bruijn étudie statistiquement les interdépendances entre diverses fonctionnalités intelligentes d'aide à la recherche et les actions de l'utilisateur (par exemple la décision de persévérer ou non dans une certaine direction après une série de résultats médiocres).

D'autre part, Jean-François Delannoy mène des recherches sur le repérage et le marquage en XML de l'argumentation, un aspect de l'analyse du discours très peu étudié par la linguistique informatique. Il s'agit d'analyser la partie centrale du « jeu de la preuve » (conclusions, faits, modèle du domaine, inférence explicite par déduction/induction/analogie, formulation et mise à l'épreuve d'hypothèses) et les effets rhétoriques (alternance d'ironie et d'indignation, etc.). L'analyse argumentaire peut rendre des services notables pour la recherche et l'acquisition d'informations en ligne, la lecture critique de textes multiples, et l'aide à la décision par la synthèse de contenu.

UNIVERSITÉ DE SAVOIE
ÉQUIPE SYSCOM

Systèmes Communicants
Université de Savoie
Domaine Universitaire Scientfique
73376 Le Bourget du Lac Cedex
Mél. : <prenom>.<nom> @univ-savoie.fr

Philippe Brunier, Laurent Denoue, Christine Ferraris, Nadia Kabachi, Christian Martel, Michel Riveill, Stéphane Talbot, Miguel Tomasena, Laurence Vignollet.

Le domaine de recherche couvert par SysCom, jeune équipe de l'Université depuis 1997, est la construction des applications coopératives. Dans ce cadre nous participons, entre autres, à deux projets sur le thème « IA et Internet » que nous décrirons brièvement par la suite.

Le premier projet, nommé REVALP est un programme Interreg II. Ce programme a pour objectif de valoriser et faciliter l'accès aux sources d'informations et services de la partie Nord-Ouest de l'arc alpin sur Internet. Notre participation dans ce programme consiste dans la réalisation d'un moteur de recherche acceptant des requêtes en langage naturel, ceci en français et en italien. Nous développons donc des analyseurs dont le but est d'identifier un ensemble de mots clés et leurs relations logiques permettant l'accès à une base de données. Cette analyse se situe au-dessus d'un système dont les fonctions sont l'indexation et la recherche dans un ensemble restreint de sites WEB. Nos devons mettre en place des analyseurs ayant un degré de robustesse satisfaisant dans un contexte où le lexique et la forme des phrases ne sont pas complètement définis. Ce projet est actuellement dans sa phase initiale. Nous comptons explorer l'utilisation de l'unification et l'introduction du flou pour tenir compte des approximations.

Le second projet s'inscrit dans le domaine de l'aide à la classification automatique de documents. Nous proposons dans un premier temps d'ajouter un outil de surlignage aux navigateurs WEB pour repousser les limites actuelles des signets (ou bookmarks). En effet, un signet ne stocke usuellement que l'adresse et le titre d'un document ; la recherche d'informations et la navigation sont donc difficiles. Associer les parties surlignées devrait tout d'abord améliorer cette situation. Ensuite, les parties surlignées aident l'usager à comprendre en quoi le document est intéressant. Le seul indice actuellement implanté par les navigateurs consiste à distinguer les liens déjà visités par une couleur spéciale. Finalement, le surlignage facilite le conseil de lecture. Il est fréquent de recommander une adresse par courrier électronique, mais pouvoir annoter le document recommandé serait certainement appréciable. Nous proposons d'utiliser un « Xpointer » pour identifier une partie surlignée. Ces Xpointer peuvent être augmentés du commentaire de l'usager, et d'autres métadonnées associées au surlignage. Cette représentation sous forme d'adresse étendue autorise l'utilisation des programmes actuels pour envoyer l'adresse d'un document surligné à d'autres usagers. Il est aussi possible de publier les adresses de documents surlignés. Un prototype opérationnel intégré à Internet Explorer sous Windows a été réalisé.

Nous nous intéressons de manière plus large à la classification automatique des documents annotés. Le nombre des informations accumulées par les usagers augmente en effet significativement avec le temps, et la classification manuelle devient vite complexe. Nous envisageons donc une approche par apprentissage pour classer automatiquement des textes. Pendant l'apprentissage, l'usager annote les documents avec des métadonnées (sujet, type). Ensuite, sur présentation d'un nouveau document, le système devra proposer automatiquement à l'usager des métadonnées.

Si la plupart des techniques de classification de textes travaillent sur le contenu des documents, deux recherches récentes ont obtenu des résultats encourageants en utilisant également le contexte des documents. Cette idée nous semble très pertinente pour déterminer le type d'un document et nous comptons l'explorer.

Dans tous les cas, les techniques de classification automatique requièrent la sélection des termes importants. Ces termes décident de l'appartenance d'un document à une catégorie. La technique du surlignage semble ici trouver un champ d'application idéal puisque les termes importants peuvent être désignés par l'usager lui-même. Le surlignage permet aussi à l'usager de communiquer avec le système, par exemple pour justifier un choix de classification.

Livres

FEATURE SELECTION FOR KNOWLEDGE DISCOVERY AND DATA MINING

HUAN LIU, HIROSHI MOTODA

Kluwer 1998

Note de lecture par Yves Kodratoff — LRI

Cet ouvrage donne une excellente présentation détaillée des diverses méthodes de sélection des descripteurs, issues aussi bien des statistiques que de l'apprentissage symbolique.

Les données « réelles » contiennent toujours des descripteurs inutiles et il n'existe pas de méthode d'induction qui ne soit plus ou moins sensible à la présence de ces descripteurs. De plus, il arrive aussi très souvent que les données présentent des centaines de descripteurs et les temps de calcul deviennent alors prohibitifs.

C'est pourquoi, il est à peu près impossible d'envisager une application sérieuse des méthodes d'Extraction des Connaissances à partir des Données (ECD, appelée le plus souvent « Data Mining ») sans une sélection préalable des descripteurs.

Le « débutant » trouvera une excellente description systématique de ces méthodes, et le spécialiste aura plaisir à grappiller ici et là une méthode qu'il ignorait encore : je recommande le livre à tous.

Du côté des éditeurs:

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

Roland Ducournau, Jérôme Euzenat, Gérald Masini, Amedeo Napoli, éditeurs, Langages et Modèles à objets, *INRIA. 1998*

Vous avez *aimé* un *livre*

Vous voulez faire partager votre intérêt, deux options sont possibles. Vous pouvez en faire une critique d'une ou deux pages et l'envoyer à Philippe Laublet.
Mais vous pouvez aussi lui faire parvenir une courte note, de quelques lignes, indiquant en quoi il vous semble digne d'attention pour les membres de l'AFIA. Nous la publierons volontiers.

CALENDRIER

Congrès/Colloques/Écoles d'été...

21-22 octobre 1999, LFA'99, rencontres francophones sur la logique floue et ses applications, Valenciennes, France.
Objectifs : Cette année LFA s'ouvre à d'autres théories de l'incertain, telles que la théorie des fonctions de croyance de Dempster-Shafer ou les approches qualitatives. Tous les domaines où ces théories sont utilisées sont concernés, tels que la commande automatique, l'aide à la décision, la classification, la reconnaissance des formes et l'analyse de données, le traitement et l'interprétation des images, la fusion d'informations, l'intelligence artificielle, les systèmes d'information, etc. Tous les travaux théoriques, méthodologiques, ainsi que les applications, peuvent être soumis. Comme chaque année, la priorité est donnée aux jeunes chercheurs afin qu'ils puissent présenter leurs travaux. Le comité de programme sera particulièrement attentif au caractère novateur et didactique des contributions des chercheurs universitaires. Des exposés traitant d'expériences concrètes et d'applications dans le domaine industriel sont vivement souhaités.
Contact : Isabelle Bloch (pdt. comité de prgm.) lfa99@tsi.enst.fr
Page d'accueil : **http://www.tsi.enst.fr/lfa99/**

21-22 octobre 1999, 2e conférence du chapitre français de l'ISKO « L'indexation à l'heure de l'Internet », Lyon.
Objectifs : Regroupée au sein du chapitre français de l'ISKO *(International Society of Knowledge Organisation)* depuis 1996, la communauté francophone de recherche dans ce domaine organise des rencontres depuis 1997. La conférence organisée en 1999, à Lyon, avec l'ENSSIB et ERSICO-Lyon 3, sera l'occasion de s'informer et de confronter des travaux concernant les développements les plus récents en matière de modèles, d'outils et de réalisations sur l'indexation.
Contact : Metzger Jean Paul, Professeur à l'université Jean Moulin Lyon 3, Président du comité d'organisation
Courriel : **isko@enssib.fr**

NOVEMBRE 1999

22-26 novembre 1999, IHM'99, 11e Conférence Internationale Francophone sur l'Interaction Homme-Machine, Montpellier, France.
Objectifs : L'interaction homme-machine change de nature avec l'introduction de nouvelles techniques d'interaction accessibles à tous et intervenant dans l'activité humaine, individuelle ou collective, professionnelle ou privée, concernant tous les âges et tous les types d'utilisateurs. IHM'99 fera le point sur les développements les plus récents et les travaux en cours en matière de concepts, modèles, méthodes, outils, réalisations, expérimentations de nouvelles formes d'interaction et de développement de systèmes interactifs individuels ou collectifs. IHM'99 sollicite en particulier la présentation de travaux où les disciplines connexes à l'interaction homme-machine contribuent à l'évolution de l'interaction pour tous.
Réception des articles longs : 30 Avril 1999
Réception des autres soumissions : 1er juin 1999
Notification aux auteurs : 9 Juillet 1999
Réception des versions définitives : 1er Octobre 1999
Président de la conférence : Jocelyne Nanard, LIRMM, Univ. Montpellier II
Page d'accueil : **http://www-ihm.lri.fr/ihm99/**

FÉVRIER 2000

1-3 février 2000, RFIA'2000, 12e Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle, Paris.
Thèmes : RFIA est le congrès par excellence des chercheurs dans les domaines de la Reconnaissance des Formes et de l'Intelligence Artificielle. Il a pour but de promouvoir toutes les recherches dans ces domaines et offre une occasion unique de rassemblement et d'échanges scientifiques indispensables entre les chercheurs de ces deux importantes communautés. RFIA'2000, la douzième édition de ce congrès, se tiendra du 1er au 3 février 2000 à Paris, aux Salons de l'Aveyron situés dans un quartier en pleine expansion à proximité de la Bibliothèque Nationale de France et du parc de Bercy. Il aura pour spécificité de s'internationaliser en s'ouvrant à la francophonie et nous espérons que les chercheurs, ingénieurs et étudiants francophones de tous pays répondront en nombre à cette ouverture et à cet appel. Les communications peuvent concerner des recherches de toute nature : travaux méthodologiques, théoriques ou expérimentaux, développement d'algorithmes, d'outils, de systèmes ou d'applications, dans l'un des domaines de la Reconnaissance des Formes et de l'Intelligence Artificielle. Ces travaux doivent être originaux, mais ils peuvent être soumis simultanément à un congrès non francophone.
Déclaration d'intention : 15 mai 1999
Réception des soumissions : 15 juin 1999
Notification aux auteurs : 25 octobre 1999
Réception des versions définitives : 23 novembre 1999
Co-présidents Comité de programme : Rachid DERICHE et Marie-Christine ROUSSET
Courriel : rfia2000@tsi.enst.fr
Page d'accueil : **http://www.tsi.enst.fr/rfia2000/**

MAI 2000

23-26 Mai 2000, COOP'2000, conférence internationale sur la conception des systèmes coopératifs, Sophia-Antipolis.
Objectif : L'objectif des conférences COOP est de contribuer à résoudre les problèmes liés à la conception de systèmes coopératifs, et à l'intégration de ces systèmes dans les organisations. Les conférences visent à promouvoir : *(a)* une meilleure compréhension des processus de travail coopératif homme-homme et homme-machine ; *(b)* le développement de modèles de la coopération et du travail coopératif à partir de différents points de vue ; *(c)* le développement de méthodologies de conception appropriées et de nouvelles fonctionnalités pour les systèmes coopératifs.
Réception des soumissions : 15 Octobre 1999 (papiers), 31 Octobre 1999 (workshops), 15 Décembre (posters et colloque doctoral)
Contact : Monique.Simonetti@sophia.inria.fr
Page d'accueil : http://www-sop.inria.fr/acacia/Coop/Coop2000/

Nous annonçons dans cette rubrique les manifestations IA (i) *se déroulant en France ou* (ii) *francophones et se déroulant hors de France.*
Pour inclure une annonce, envoyer un court texte (disquette Mac ou mail avec fichier attaché) à Jean Charlet (voir « l'ours » page 2 de ce bulletin).

SOMMAIRES DES REVUES

Merci à Monique Baron et à Jean Charlet pour leur contribution à cette rubrique !

N'hésitez pas à envoyer un message à Brigitte Grau (grau@limsi.fr) pour lui indiquer toute suggestion permettant d'améliorer cette rubrique

REVUE D'INTELLIGENCE ARTIFICIELLE

http://www.editions-hermes.fr/riap.htm

REVUE D'INTELLIGENCE ARTIFICIELLE V13 1999
RAISONNEMENT A PARTIR DE CAS
• Introduction, *ISABELLE BICHINDARITZ, ALAIN MILLE, AMEDEO NAPOLI*
• Raisonnement à partir de cas et résolution de problèmes dans une représentation par objets, *JEAN LIEBER, AMEDEO NAPOLI*
• Une approche formelle du raisonnement à partir de cas dans une logique de descriptions, *SYLVIE SALOTTI, VERONIQUE VENTOS*
• Acquisition d'expérience par raisonnement à partir de cas pour la recherche dans un espace d'états, *JEAN-MAURICE BLIN, VERONIQUE MASSON, RENE QUINIOU*
• Raisonnement fondé sur l'expérience: un nouveau paradigme en supervision industrielle, *ALAIN MILLE, BEATRICE FUCHS, BENOIT CHIRON*
• Etude en ergonomie cognitive sur la réutilisation en conception : quelles leçons pour le raisonnement à partir de cas ?, *WILLEMIEN VISSER*

ARTIFICIAL INTELLIGENCE

http://www.elsevier.nl:80/inca/publications/store/5/0/5/6/0/1/

ARTIFICIAL INTELLIGENCE V108 N°1-2 MARCH 1999
• Functional dependencies in Horn theories, *TOSHIHIDE IBARAKI, ALEXANDER KOGAN, KAZUHISA MAKINO*
• Linear-time algorithms for testing the realisability of line drawings of curved objects, *MARTIN C. COOPER*
• On the complexity of qualitative spatial reasoning: A maximal tractable fragment of the region connection calculus, *JOCHEN RENZ, BERNHARD NEBEL*
• Explanatory update theory: applications of counterfactual reasoning to causation, *CHARLES L. ORTIZ JR.*
• A foundational theory of belief and belief change, *ALEXANDER BOCHMAN*

ARTIFICIAL INTELLIGENCE V109 N°1-2 APRIL 1999
• Modal logics, description logics and arithmetic reasoning, *HANS JÜRGEN OHLBACH, JANA KOEHLER*
• iCTRL: Intensional conformal text representation language, *GABOR REDEY*
• A directly cautious theory of defeasible consequence for default logic via the notion of general extension, *G. ALDO ANTONELLI*
• Using regression-match graphs to control search in planning, *DREW MCDERMOTT*
• Non-uniform time sharing in the concurrent execution of constraint solving, *RICHARD KRAJCOVIECH, MARGARETA KOTOCOVA*
• Representation of propositional expert systems as partial functions, *ROBERT M. COLOMB*
• Learning by discovering concept hierarchies, *BLAZ ZUPAN, MARKO BOHANEC, JANEZ DEMSAR, IVAN BRATKO*
• Constraint satisfaction over connected row-convex constraints, *YVES DEVILLE, OLIVIER BARETTE, PASCAL VAN HENTENRYCK*
• Computational complexity of relating time points with intervals, *PETER JONSSON, THOMAS DRAKENGREN, CHRISTER BÄCKSTRÖM*
• Preferred answer sets for extended logic programs, *GERHARD BREWKA, THOMAS EITER*

ARTIFICIAL INTELLIGENCE V110 N°1 MAY 1999
• Emergent cooperative goal-satisfaction in large scale automated-agent systems, *ONN SHEHORY, SARIT KRAUS, OSHER YADGAR*
• Computing intersections of Horn theories for reasoning with models, *THOMAS EITER, TOSHIHIDE IBARAKI, KAZUHISA MAKINO*
• Statics and dynamics of induced systems, *MICHAEL FREUND*

AI MAGAZINE

http://aaai.org/Magazine/

AI MAGAZINE V20 N°1 SPRING 1999
SPECIAL ARTICLES ON INNOVATIVE APPLICATIONS
• The Innovative Applications of Artificial Intelligence Conference, *BRUCE BUCHANAN, SAM UTHURUSAMYR*
• Knowledge-Based Avoidance of Drug-Resistant HIV Mutants, *RICHARD H. LATHROP, NICHOLAS R. STEFFEN, MIRIAM P. RAPHAEL, SOPHIA DEEDS-RUBIN, MICHAEL J. PAZZANI, PAUL J. CIMOCH, DARRYL M. SEE, JEREMIAH G. TILLES*
• Automated Intelligent Pilots for Combat Flight Simulation, *RANDOLPH M. JONES, JOHN E. LAIRD, PAUL E. NIELSEN, KAREN J. COULTER, PATRICK KENNY, FRANK V. KOSS*
• A New Technique Enables Dynamic Replanning and Rescheduling of Aeromedical Evacuation, *ALEXANDER KOTT, VICTOR SAKS, ALBERT MERCER*
• The NASD Regulation Advanced-Detection System (ADS), *J. DALE KIRKLAND, TED E. SENATOR, JAMES J. HAYDEN, TOMASZ DYBALA, HENRY G. GOLDBERG, PING SHYR*
• Turbine Engine Diagnostics (TED), *RICHARD HELFMAN, ED BAUR, JOHN DUMER, TIM HANRATTY, HOLLY INGHAM*

ARTICLES

• Automated Deduction: Looking Ahead, *DONALD W. LOVELAND*

• Practically Coordinating, *EDMUND H. DURFEE*

BOOK REVIEW, OPINION, AND WORKSHOP REPORTS

• AI Game-Playing Techniques, *DANA S. NAU*

• Applications of Ontologies and Problem-Solving Methods, *ASUNCION GOMEZ-PEREZ, V. RICHARD BENJAMINS*

• Case-Based Reasoning Integrations, *DAVID W. AHA, JODY DANIELS*

• Learning for Text Categorization, *MEHRAN SAHAMI*

• Predicting the Future: AI Approaches to Time-Series Problems, *ANDREA DANYLUK, TOM FAWCETT, FOSTER PROVOST*

• Software Tools for Developing Agents, *BRIAN LOGAN, JEREMY BAXTER*

• The Distributed Data-Mining Workshop, *HILLOL KARGUPTA , PHILIP CHAN*

• Review of Affective Computing, AARON SLOMAN

• Response to Sloman's Review of Affective Computing, *ROSALIND W. PICARD*

• Review of One Jump Ahead: Challenging Human Supremacy in Checkers, *PETER F. PATEL-SCHNEIDER*

COGNITIVE SCIENCE

COGNITIVE SCIENCE V22 N°4 OCTOBER-DECEMBER 1999

• The Representation and Processing of Co-Reference in Discourse, *PETER C. GORDON, RANDALL HENDRICK*

• Distributional Information: A Powerful Cue for Acquiring Syntactic Categories, *MARTIN REDINGTON, NICK CHATER, STEVEN FINCH*

• On the Role of Mathematics in Explaining the Material World: Mental Models for Proportional Reasoning, *DANIEL L SCHWARTZ, JOYCE L. MOORE*

• Pertinence Generation in Radiological Diagnosis Spreading Activation and the Nature of Expertise, *ERIC RAUFASTE, HELENE EYROLLE, CLAUDETTE MARINE*

COGNITIVE SCIENCE V23 N°1 JANUARY-MARCH 1999

• From Pixels to People: A Model of Familiar Face Recognition, *A. MIKE BURTON, VICKI BRUCE, P. J. B. HANCOCK*

• Encoding Categorical and Coordinate Spatial Relations Without Input-Output Correlations: New Simulation Models, *DAVID P. BAKER, CHRISTOPHER F. CHABRIS, STEPHEN M. KOSSLYN*

• Mapping across Domains Without Feedback: A Neural Network Model of Transfer of Implicit Knowledge, *ZOLTAN DIENES, GERRY T. M ALTMANN, SHI-JI GAO*

• Attention Metaphors: How Metaphors Guide the Cognitive Psychology of Attention, *DIEGO FERNANDEZ DUQUE, MARK L. JOHNSON*

APPLIED ARTIFICIAL INTELLIGENCE

http://www.tandf.co.uk/JNLS/aai.htm

APPLIED ARTIFICIAL INTELLIGENCE V13 N°4-5 1999

SPECIAL ISSUE: ANIMATED INTERFACE AGENTS: MAKING THEM INTELLIGENT

Editorial, E. André

• Animated agents for procedural training in virtual reality: perception, cognition, and motor control, *J. RICKEL, W. LEWIS JOHNSON*

• Deictic believability: coordinated gesture, locomotion, and speech in lifelike pedagogical agents, *J. C. LESTER, J. L. VOERMAN, S. G. TOWNS, C. B. CALLAWAY*

• Employing AI methods to control the behavior of animated interface agents, *E. ANDRE, T. RIST, J. MULLER*

• Mind model for multimodal communicative creatures and humanoids, *K. R. THORISSON*

• Animated interface agent applying atms-based multimodal input interpretation, *Y. KONO, T. YANO, T. IKEDA, T. CHINO, K. SUZUKI, H. KANAZAWA*

• The power of a nod and a glance: envelope vs. emotional feedback in animated conversational agents, *J. CASSELL, K. R. THORISSON*

vConferring human action recognition skills to life-like agents, *L. EMERING, R. BOULIC, D. THALMANN*

COMPUTATIONAL INTELLIGENCE

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

COMPUTATIONAL INTELLIGENCE V15 N°2 1999

• Introduction to the Special Issue on PACLING'97, *KIYOSHI JUN-ICHI NAKAMURA*

• Retrieving Domain-Specific Collocations by Co-occurrences and Word Order Constraints, *SAYORI SHIMOHATA, TOSHIYUKI SUGIO, JUNJI NAGATA*

• Resolving Ambiguous Segmentation of Korean Compound Nouns Using Statistics and Rules, *BO-HYUN YUN, YONG-JAE KWAK, HAE-CHANG RIM*

• A Pattern-Based Approach Using Compound Unit Recognition and Its Hybridization with Rule-Based Translation, *HANMIN JUNG, SANGHWA YUH, TAEWAN KIM, SANGKYU PARK*

• English-Thai Translation: Initial Experiments with a Multiphase Translation System, *KANLAYA NARUEDOMKUL, NICK CERCONE, BOONCHAROEN, SIRINAOVAKUL*

• Cooperative Understanding of Utterances and Gestures in a Dialogue-Based Problem Solving System, *TSUTOMU ENDO, TSUNEO KAGAWA*

INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES

INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES V50 N°3 MARCH 1999

vDesigning claims for reuse in interactive systems design, *A. G. SUTCLIFFE, J. M. CARROLL*

• User-interface modelling-adding usability to use cases, *MAGNUS LIF*

• A review of user-interface design guidelines for public information kiosk systems, *M. C. MAGUIRE*

INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES V50 N°4 APRIL 1999

• Depth of processing and design-assessment of ecological interfaces: Task analysis, *PATRICE TERRIER, JEAN-MARIE CELLIER*

• Facilitating navigation in information spaces: Road-signs on the World Wide Web, *CHRISTOPHER S. CAMPBELL, PAUL P. MAGLIO*

• Annotation technology, *ILIA A. OVSIANNIKOV, MICHAEL A. ARBIB, THOMAS H. MCNEILL*

INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES V50 N°5 MAY 1999

• On the generation of interactive iconic environments, *GENNARO COSTAGLIOLA, SERGIO OREFICE, GIUSEPPE POLESE, MAURIZIO TUCCI, GENOVEFFA TORTORA*

• Expert error in trouble-shooting: an exploratory study in electronics, *DENIS BESNARD, MIREILLE BASTIEN-TONIAZZO*

• Compositional modelling of reflective agents, *FRANCES M. T. BRAZIER, JAN TREUR*

INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES V50 N°6 JUNE 1999

• Editorial: Perceptual Control Theory and its application, *M. M. TAYLOR*

• Models and their worlds, *W. THOMAS BOURBON, WILLIAM T. POWERS*

• A model of kinesthetically and visually controlled arm movement, *WILLIAM. T. POWERS*

• PERCOLATe: Perceptual control analysis of tasks, *RICHARD S. MARKEN*

• Perceptual control and layered protocols in interface design: I. Fundamental concepts, *P. S. E. FARRELL, J. G. HOLLANDS, M. M. TAYLOR, H. D. GAMBLE*

• Perceptual control and layered protocols in interface design: II. The general protocol grammar, *M. M. TAYLOR, P. S. E. FARRELL, J. G. HOLLANDS*

• Towards explaining the behaviour of novice users, R. HAAKMA

• Testing the self as a control system: Theoretical and methodological issues, *R. J. ROBERTSON, D. M. GOLDSTEIN, M. MERMEL, M. MUSGRAVE*

• Simulating arcs and rings in gatherings, *CHARLES W. TUCKER, DAVID SCHWEINGRUBER, CLARK MCPHAIL*

COMPUTATIONAL LINGUISTICS

http://mitpress.mit.edu/journal-contents.tcl?issn=08912017

COMPUTATIONAL LINGUISTICS V25 N°1 MARCH 1999

• A Process Model for Recognizing Communicative Acts and Modeling Negotiation Subdialogues, *SANDRA CARBERRY, LYNN LAMBERT*

• Default Representation in Constraint-based Frameworks, *ALEX LASCARIDES, ANN COPESTAKE*

• Bitext Maps and Alignment via Pattern Recognition, *I. DAN MELAMED*

• Statistical Properties of Probabilistic Context-Free Grammars, *ZHIYI CHI*

BOOK REVIEWS

• Survey of the State of the Art in Human Language Technology, Editorial Board: *RONALD COLE,* (Editor-in-Chief), *JOSEPH MARIANI, HANS USZKOREIT, ANNIE ZAENEN, AND VICTOR ZUE, by VAROL AKMAN*

• Exploring Textual Data By *LUDVIC LEBART, ANDRE SALEM, AND LISETTE BARRY, by DOUGLAS BIBER*

• Linguistic Databases By *JOHN NERBONNE* (editor), by *JÖRG TIEDEMANN*

NATURAL LANGUAGE ENGINEERING

http://www.cup.cam.ac.uk/scripts/webjrn1.asp?mnemonic=nle

NATURAL LANGUAGE ENGINEERING V4 N°4 DECEMBER 1998

Linguistic issues in the development of ReGra: A grammar checker for Brazilian Portuguese, *RONALDO TEIXEIRA MARTINS, RICARDO HASEGAWA, MARIA DAS GRAÇAS VOLPE NUNES, GISELE MONTILHA, OSVALDO NOVAIS DE OLIVEIRA JR.*

Splitting-merging model of Chinese word tokenization and segmentation, *YUAN YAO, KIM TEN LUA*

Finding a domain-appropriate sense inventory for semantically tagging a corpus, *ALESSANDRO CUCCHIARELLI, PAOLA VELARDI*

Understanding pronominal references to quantified expressions, *ELISE H. TURNER, COLLETTE M. MATTHIAS*

BOOK REVIEWS

Comparing English Worldwide: The International Corpus of English, *SIDNEY GREENBAUM* (ed.), by *GEOFFREY SAMPSON*

Analogical Natural Language Processing, *DANIEL JONES*, by *JAMES HAMMERTON*

Quantifiers, Logic, and Language, *JAAP VAN DER DOES, JAN VAN EIJK* (eds.), by *VAROL AKMAN*

The Role of Argument Structure in Grammar: Evidence from Romance, *ALEX ALSINA, by MAYUMI MASUKO*

Practical Guide to Syntactic Analysis, *GEORGIA M. GREEN, JERRY L. MORGAN, by YVONNE CANNING*

Machine Translation: an Introductory Guide, *D. ARNOLD, L. BALKAN, R.*

LEE HUMPHRIES, S. MEIJER, L. SADLER, by YORICK WILKS
The Uncertain Reasoner's Companion: A Mathematical Perspective, *J. B. PARIS*, by *MALCOLM FARROW*
Logic, Language and Computation, vol 1., *JERRY SELIGMAN AND DAG WESTERSTAHL*, (eds.), by *MICHEL EYTAN*

COMPUTER SPEECH AND LANGUAGE

COMPUTER SPEECH AND LANGUAGE V13 N°2 APRIL 1999

• Interpolation of n-gram and mutual-information based trigger pair language models for Mandarin speech recognition, *Z. GUODONG, L. KIMTENG*
• Multiple pronunciation dictionary using HMM-state confusion characteristics, *Y. WAKITA, H. SINGER, Y. SAGISAKA*
• Evaluating the pronunciation component of text-to-speech systems for English: a performance comparison of different approaches, *R. I. DAMPER, Y. MARCHAND, M. J. ADAMSON, K. GUSTAFSON*
• Stochastically-based semantic analysis for machine translation, *W. MINKER, M. GAVALDA, A. WAIBEL*
• A Bayesian triphone model, *JI MING, F. J. SMITH*

MACHINE LEARNING

MACHINE LEARNING V34 N°1-3 FEBRUARY 1998

• Guest Editors' Introduction: Machine Learning and Natural Language, *CLAIRE CARDIE, RAYMOND J. MOONEY*
• Forgetting Exceptions is Harmful in Language Learning, *WALTER DAELEMANS, ANTAL VAN DEN BOSCH, JAKUB ZAVREL*
• Similarity-Based Models of Word Cooccurrence Probabilities, *IDO DAGAN, LILLIAN LEE, FERNANDO C. N. PEREIRA*
• An Efficient, Probabilistically Sound Algorithm for Segmentation and Word Discovery, *MICHAEL R. BRENT*
• A Winnow-Based Approach to Context-Sensitive Spelling Correction, *ANDREW R. GOLDING, DAN ROTH*
• Using Decision Trees to Construct a Practical Parser, *MASAHIKO HARUNO, SATOSHI SHIRAI, YOSHIFUMI OOYAMA*
• Learning to Parse Natural Language with Maximum Entropy Models, *ADWAIT RATNAPARKHI*
• Statistical Models for Text Segmentation, *DOUG BEEFERMAN, ADAM BERGER, JOHN LAFFERTY*
• An Algorithm that Learns What's in a Name, *DANIEL M. BIKEL, RICHARD SCHWARTZ, RALPH M. WEISCHEDEL*
• Learning Information Extraction Rules for Semi-Structured and Free Text, *STEPHEN SODERLAND*

MACHINE LEARNING V35 N°1 APRIL 1998

• An Experimental Evaluation of Integrating Machine Learning with Knowledge Acquisition, *G. I. WEBB, J. WELLS AND Z. ZHENG*
• Concept Learning and Feature Selection Based on Square-Error Clustering, B. MIRKIN
• An Efficient Method to Estimate Bagging's Generalization Error, *D. H. WOLPERT, W. G. MACREADY*
• Learning to Take Actions, *R. KHARDON*

NEURAL NETWORKS

http://www.elsevier.nl/cite/54/show/jrnl_index.htt

NEURAL NETWORKS V12 N°2 MARCH-1998

• Topographic shear and the relation of ocular dominance columns to orientation columns in primate and cat visual cortex, *RICHARD J. WOOD, ERIC L. SCHWARTZ*
• A subthreshold MOS circuit for the Lotka--Volterra neural network producing the winners-share-all solution, *T. ASAI, T. FUKAI, S. TANAKA*
• Language acquisition from sparse input without error feedback, *R.F. HADLEY, V.C. CARDEI*
• Embedding recurrent neural networks into predator--prey models, *YVES MOREAU, STEPHANE LOUIES, JOOS VANDEWALLE, LEON BRENIG*
• Results from the Gardner--Derrida--Mottishaw theory of associative memory, *HIDEYUKI KOYAMA, NORIO FUJIE, HIROYUKI SEYAMA*
• Analysis and optimal design of continuous neural networks with applications to associative memory, *MIAO ZHENJIANG, YUAN BAOZONG*
• On delayed impulsive Hopfield neural networks, *ZHI-HONG GUAN, GUANRONG CHEN*
• Improved bidirectional retrieval of sparse patterns stored by Hebbian learning, *FRIEDRICH T. SOMMER, GÜNTHER PALM*
• Matrix logarithm parametrizations for neural network covariance models, *PETER M. WILLIAMS*
• Model selection in neural networks, *ULRICH ANDERS, OLAF KORN*
• ANNSyS: an Analog Neural Network Synthesis System, *ISMET BAYRAKTARGLU, ARIF SELÇUK ÖGRENCI, GÜNHAN DÜNDAR*
• A loss function approach to model selection in nonlinear principal components, *ANDREW R. WEBB*
• A comfortable brain-interface to video displays, *MASAHIDE NOMURA*
• Multi-sensor integration for on-line tool wear estimation through radial basis function networks and fuzzy neural network, *R.J. KUO, P.H. COHEN*
• Application of neural networks to the modelling of some constitutive laws, *S. PERNOT, C.-H. LAMARQUE*
• Comments on: A unified algorithm for principal and minor components extraction, *FA-LONG LUO, ROLF UNBEHAUEN*
• Response, Chen Tianping, *SHUNICHI AMARI*

NEURAL NETWORKS V12 N°3 APRIL 1999

• Evaluating the Vapnik--Chervonenkis dimension of artificial neural networks using the Poincaré polynomial, *MARTHA A. CARTER, MARK E. OXLEY*
• Modeling the precedence effect for speech using the gamma filter, *ODELIA SCHWARTZ, JOHN G. HARRIS, JOSE C. PRINCIPE*
• The eigenspace separation transform for neural-network classifiers, *DON TORRIERI*
• Nonlinear independent component analysis: Existence and uniqueness results, *AAPO HYVÄRINEN, PETTERI PAJUNEN*
• Dynamics of neural networks with a central element , *Y.B. KAZANOVICH, R.M. BORISYUK*
• Global dynamics of bidirectional associative memory neural networks involving transmission delays and dead zones, *V. SREE HARI RAO, BH.R.M. PHANEENDRA,*
• Neural network approximation of continuous functionals and continuous functions on compactifications, *M.B. STINCHCOMBE*
• Template-based procedures for neural network interpretation, *J.A. ALEXANDER, M.C. MOZER*
• A neural network for enhancing boundaries and surfaces in synthetic aperture radar images, *ENNIO MINGOLLA, WILLIAM ROSS, STEPHEN GROSSBERG*
• An NN-based approach for tuning servocontrollers, *ELDER M. HEMERLY, CAIRO L. NASCIMENTO, JR*
• A neural bridge from syntactic to statistical pattern recognition, *FRANK T. ALLEN, JASON M. KINSER, H. JOHN CAULFIELD*
• A global learning algorithm for a RBF network, *QIUMING ZHU, YAO CAI, LUZHENG LIU*
• Hierarchical classification with a competitive evolutionary neural tree, *R.G. ADAMS, K. BUTCHART, N. DAVEY*
• Fat neural network for recognition of position-normalised objects, *D. DOLLFUS, L. BEAUFORT*

ARTIFICIAL INTELLIGENCE IN MEDICINE

ARTIFICIAL INTELLIGENCE IN MEDICINE V16 N° 1 1999

• Selected techniques for data mining in medicine, *NADA LAVRA*
• Analysing and improving the diagnosis of ischaemic heart disease with machine learning, *MATJA KUKAR, IGOR KONONENKO, CIRIL GROELJ, KATARINA KRALJ, JURE FETTICH*
• Two-Stage Machine Learning model for guideline development, *SUBRAMANI MANI, WILLIAM R. SHANKLE, MALCOLM B. DICK, MICHAEL J. PAZZANI*
• Medical data mining using evolutionary computation, *PO SHUN NGAN, MAN LEUNG WONG, WAI LAM, KWONG SAK LEUNG, JACK C.Y. CHENG*
• Dynamic and static approaches to clinical data mining, *DAVID MCSHERRY*

ARTIFICIAL INTELLIGENCE IN MEDICINE V16 N°2 1999

• Fuzzy diagnosis, *LUDMILA I. KUNCHEVA, FRIEDRICH STEIMANN*
• A fuzzy clustering based segmentation system as support to diagnosis in medical imaging, *FRANCESCO MASULLI, ANDREA SCHENONE*
• Obtaining interpretable fuzzy classification rules from medical data, *DETLEF NAUCK, RUDOLF KRUSE*
• Fuzzy pre-processing of gold standards as applied to biomedical spectra classification, *NICOLINO J. PIZZI*
• Diagnostic monitoring in anaesthesia using fuzzy trend templates for matching temporal patterns, *ANDREW LOWE, MICHAEL J. HARRISON, RICHARD W. JONES*

FUZZY SETS AND SYSTEMS

http://www.elsevier.nl/cite/60/show/toc.htt

FUZZY SETS AND SYSTEMS V103 N°2 APRIL 1998

• Soft computing for feature analysis, *NIKHIL RANJAN PAL*
• Voting in fuzzy rule-based systems for pattern classification problems, *HISAO ISHIBUCHI, TOMOHARU NAKASHIMA, TAKEHIKO MORISAWA*
• Fuzzy clustering analysis for optimizing fuzzy membership functions, *MUSONG CHEN, SHINN-WEN WANG*
• Patterns and symbols treated in a three-layered associative memory network, *Y. HATTORI, T. FURUHASH*
• FIRE operators for image processing, *FABRIZIO RUSS*
• Fuzzy reasoning for image compression using adaptive triangular plane patches, *LISONG WANG, LIFENG HE, ATSUKO MUTOH, TSUYOSI NAKAMURA, HIDENORI ITOH*
• Restoration of gray images based on a genetic algorithm with Laplacian constraint, *YEN-WEI CHEN, ZENSHO NAKAO, KOUICHI ARAKAKI, XUE FANG, SHINICHI TAMURA*
• Two new neural network approaches to two-dimensional CT image reconstruction, *FATH EL ALEM F. ALI, ZENSHO NAKAO, YEN-WEI CHEN*
• Design of image exploring agent using genetic programming, *MARIO KÖPPEN, BERTRAM NICKOLAY*
• A method for still image interpretation relying on a multi-algorithms fusion scheme. Application to human face characterization, *JEAN FIGUE, MICHEL GRABISCH, MARIE PIERRE CHARBONNEL*
• Pattern recognition with neural networks combined by genetic algorithm, *SUNG-BAE CHO*
• From hybrid adjustable neuro-fuzzy systems to adaptive connectionist-based systems for phoneme and word recognition, *N.K. KASABOV, R.I. KILGOUR, S.J. SINCLAIR*

FUZZY SETS AND SYSTEMS V103 N° 3 MAY 1999

• Ranking of Indian coals via fuzzy multi attribute decision making, *V. RAVI, P.J. REDDY*
• Fuzzy direct adaptive control for a class of nonlinear systems, *TIANYOU CHAI, SHAOCHENG TONG*

• Finding relevant attributes and membership functions, *TZUNG-PEI HONG, JYH-BIN CHEN*
• Consumer surplus and producer surplus in fuzzy sense, *KWEIMEI WU*
• Consumer surplus and producer surplus for fuzzy demand and fuzzy supply, *JING-SHING YAO, KWEIMEI WU*
• Fuzzy distances for proximity characterization under uncertainty, *K. SRIDHARAN, H.E. STEPHANOU*
• Convergence theorems for fuzzy random variables and fuzzy martingales, *YUHU FENG*
• Entropy variation on the fuzzy numbers with arithmetic operations, *WEN-JUNE WANG, CHIH-HUI CHIU*
• Minimizing unconstrained fuzzy functions, *K.DAVID JAMISON, WELDON A. LODWICK*
• The fuzzy integral on product spaces for NSA measures, *E. SUAREZ DIAZ, F. SUAREZ GARCIA*
• Novel neural algorithms based on fuzzy rules for solving fuzzy relation equations: Part II, *XIAOZHONG LI, DA RUAN*
• Independence concepts in possibility theory: Part II, *LUIS M. DE CAMPOS, JUAN F. HUETE*
• Stability confidence for fuzzy matrices, *WEN-JUNE WANG, SHAO-FENG YAN*
• Completely induced L-fuzzy topological spaces, *HALIS AYGÜN, M.W. WARNER, ND S.R.T. KUDRI*
• Hom functors in the fuzzy category Fm, *FUZHENG PAN*
• (q)-level subset, *SANDEEP KUMAR BHAKAT*
• Connectedness and -connectedness in fuzzy bitopological spaces, *S. SAMPATH KUMAR*
• An advanced fuzzy controller, *RUSTOM MAMLOOK, CHIN-WANG TAO, WILEY E. THOMPSON*

FUZZY SETS AND SYSTEMS V104 N°1 MAY 1999

• Quasi- and pseudo-inverses of monotone functions, and the construction of t-norms, *ERICH PETER KLEMENT, RADKO MESIAR, ENDRE PAP*
• A note to a construction of t-norms based on pseudo-inverses of monotone functions, *PETER VICENIK*
• Smoothly generated Archimedean approximation of continuous triangular norms, *SANDOR JENEI, ENDRE PAP*
• A note on a nilpotent lower bound of nilpotent triangular norms, *VLADIMIR MARKO, RADKO MESIAR*
• Diagonals of continuous triangular norms, *RADKO MESIAR, MIRKO NAVARA*
• On reversible triangular norms, *JANOS FODOR, SANDOR JENEI*
• Generalized t-norm structures, *C.A. DROSSOS*
• Triangular norms on product lattices, *BERNARD DE BAETS, RADKO MESIAR*
• Operations fitting triangular-norm-based biresiduation, *RADKO MESIAR, VILEM NOVAK*
• On some solutions of the distributivity equation, *T. CALVO*
• Convex combinations in terms of triangular norms: A characterization of idempotent, bisymmetrical and self-dual compensatory operators, *BERNHARD MOSER, ELENA TSIPORKOVA, ERICH PETER KLEMENT*
• Triangular norm-based iterative compensatory operators, *ANNA KOLESAROVA, MAGDA KOMORNIKOVA*
• On a type of monotonicity for multidimensional operators, *G. MAYOR, T. CALVO*
• Van Melle's combining function in MYCIN is a representable uninorm: An alternative proof, *BERNARD DE BAETS, JANOS FODOR*
• A note on a nilpotent lower bound of nilpotent triangular norms [Fuzzy Sets and Systems 104 (1999) 27-34], *RADKO MESIAR*

FUZZY SETS AND SYSTEMS V104 N°2 JUNE 1999

• Conceptual fuzzy sets and their connectives, *A.O. ARIGONI, L. GOVERNATORI, A. ROSSI*
• A new approach to the identification of a fuzzy model, *MINKEE PARK, SEUNGHWAN JI, EUNTAI KIM, MIGNON PARK*
• Simplifying fuzzy modeling by both gray relational analysis and data transformation methods, *YO-PING HUANG, HUNG-CHI CHU*
• Fuzzy modeling with hybrid systems, *A.F. GOMEZ-SKARMETA, F. JIMENEZ*
• New methods for students' evaluation using fuzzy sets, *SHYI-MING CHEN, CHIA-HOANG LEE*
• Multi-layered handwriting recognition approach, *ASHUTOSH MALAVIYA, LILIANE PETERS*
• An accurate and cost-effective COG defuzzifier without the multiplier and the divider, *DAIJIN KIM, IN-HYUN CHO*
• Fuzzy control of unknown multiple-input-multiple-output plants, *CHUN-SHIEN LI, ROLAND PRIEMER*
• Sugeno controllers with a bounded number of rules are nowhere dense, *BERNHARD MOSER*
• Stable fuzzy adaptive control for a class of nonlinear systems, *YUEZ-HONG TANG, NAIYAO ZHANG, YANDA LI*
• A new fuzzy model of pattern recognition and hitch diagnoses of complex systems, *XIAODONG LIU*
vAn efficient computational algorithm for min-max operations, *XINGHU ZHANG, SHAOHUA TAN, CHANG-CHIEH HANG, PEI-ZHUANG WANG*
• Distance measure and induced fuzzy entropy, *JIULUN FAN, WEIXIN XIE*
• Existence and comparison theorems to Volterra fuzzy integral equation in (En,D), *SHIJI SONG, QIN-YU LIU, QI-CHUN XU*
• On the fuzzification of the classical definition of preference structure, *AHMED BUFARDI*
• Continuity in Zadeh's compositional rule of inference, *SANDOR JENEI*
• Fuzzy proximity ordered spaces, *M.Y. BAKIER, K. EL-SAADY*

FUZZY SETS AND SYSTEMS V104 N°3 JUNE 1999

• A representation theorem for fuzzy subsystems of a fuzzy partial algebra, *V.E.S. NISTALA MURTHY, YOGESWARA VADDIPARTHI*

• Categories isomorphic to FNS, *DEXUE ZHANG, LUOSHAN XU*
• Alternative characterizations of FNCS, *R. LOWEN, LUOSHAN XU*
• A short note on fuzzy neighbourhood convergence spaces, *N. BLASCO, LUOSHAN XU*
• Closure operators and fuzzy connectedness, *I.W. ALDERTON, G. CASTELLINI*
• Exponentiable objects in the category of topological molecular lattices, *YONG-MING LI*
• Generalized fuzzy filter and limit structure, *YOON JIN LEE*
• Subspaces of smooth fuzzy topologies and initial smooth fuzzy structures, *WERNER PEETERS*

FUZZY SETS AND SYSTEMS V105 N°1 JULY 1999

• Decision support for fuzzy comprehensive evaluation of urban development, *SHAN FENG, LI D. XU*
• Economic order quantity in fuzzy sense for inventory without backorder model, *HUEY-MING LEE, JING-SHING YAO*
• Stability analysis of fuzzy control systems, *A. KANDEL, Y. LUO, Y.-Q. ZHANG*
• A multi-stage system in compilation environments, *VICENÇ TORRA, ANGELA SODAN*
• Flexible stability criteria for a linguistic fuzzy dynamic system, *WEN-JUNE WANG, SHAO-FENG YAN, CHIH-HUI CHIU*
• Dynamic fuzzy data analysis based on similarity between functions, *A. JOENTGEN, L. MIKENINA, R. WEBER, H.-J. ZIMMERMANN*
• Fuzzy clustering with structural constraints, *VICTOR CHEPOI, DAN DUMITRESCU*
• Comparison of fuzzy partitions based on their -cuts, *SLAVKA BODJANOVA*
• A fuzzy neural network algorithm for multigroup classification, *RALF ÖSTERMARK*
• Automatic fault detection in gear boxes by dynamic fuzzy data analysis, *A. JOENTGEN, L. MIKENINA, R. WEBER, A. ZEUGNER, H.-J. ZIMMERMANN*
• Numerical solutions of fuzzy differential equations, *MING MA, MENAHEM FRIEDMAN, ABRAHAM KANDEL*
• Probability density functions of fuzzy random variables, *HSIEN-CHUNG WU*
• Fuzzy retraction and folding of fuzzy-orientable compact manifold, *M. ELGHOUL*
• Embedding problem of noncompact fuzzy number space E~(I), *CONGXIN WU, ZHANG BOKAN*
• Fuzzy multiplication semigroup, *PRATYAYANANDA DAS*
• Fubini theorem for generalized fuzzy number-valued integrals, *SUNG-JIN CHO, BYUNG-SOO LEE, GUE-MYUNG LEE, DO-SANG KIM, YEOUL-OUK SONG*
• On product of fuzzy subgroups, *ASOK KUMER RAY*
• Comments on « Frequency/time domain methods for solutions of N-order fuzzy differential equations » by *ZHANG, WANG AND LIU, JAMES J. BUCKLEY, THOMAS FEURING*
• Characterisation of fuzzy semiprime ideals of the ring of integers Z, *T.K. DUTTA, B.K. BISWAS*

FUZZY SETS AND SYSTEMS V105 N°2 JULY 1999

• Contributions to fuzzy analysis, *WERNER GÄHLER, SIEGFRIED GÄHLER*
• General approach to continuity measures, *MARK BURGIN*
• Fuzzy smoothness and sequences of fuzzy smooth functions, *M. KALINA*
• Introduction to fuzzy partial differential equations, *JAMES J. BUCKLEY, THOMAS FEURING*
• Extremal problems of approximation theory in fuzzy context, *SVETLANA V. ASMUSS, ALEXANDER P. OSTAK*
• Fuzzy topological vector spaces over , *J.J. BUCKLEY, AIMIN YAN*
• Some questions in fuzzy topology, *T. KUBIAK, M.A. DE PRADA VICENTE*
• On the L-fuzzy Brouwer fixed point theorem, *TOMASZ KUBIAK, DEXUE ZHANG*
• A further investigation for fuzzy measures on metric spaces, *QINGSHAN JIANG, SHENGRUI WANG, DIEMEL ZIOU*
• On the Lp space of observables, *BELOSLAV RIEAN*
• The Jegorov theorem on MV algebras, *MARIA JUREKOVA*

THE JOURNAL OF LOGIC PROGRAMMING

THE JOURNAL OF LOGIC PROGRAMMING V39 N°1-3 APRIL 1999

• Automated verification of Prolog programs, *BAUDOUIN LE CHARLIER, CHRISTOPHE LECLÈRE, SABINA ROSSI, AGOSTINO CORTESI*
• Abstract diagnosis, *MARCO COMINI, GIORGIO LEVI, MARIA CHIARA MEO, GIULIANA VITIELLO*
• Termination of logic programs with delay declarations, *ELENA MARCHIORI, FRANK TEUSINK*
• Verification of logic programs, *DINO PEDRESCHI, SALVATORE RUGGIERI*
• Opium: an extendable trace analyzer for Prolog, *MIREILLE DUCASSE*
• Layered modes, *SANDRO ETALLE, MAURIZIO GABBRIELLI*

THE JOURNAL OF LOGIC PROGRAMMING V40 N°1 JULY 1999

• Semantics of (disjunctive) logic programs based on partial evaluation, *STEFAN BRASS, JÜRGEN DIX*
• An approach to declarative programming based on a rewriting logic, *J.C. GONZALEZ-MORENO, M.T. HORTALA-GONZALEZ, F.J. LOPEZ-FRAGUAS, M. RODRIGUEZ-ARTALEJO*
• An algebraic framework for the definition of compositional semantics of normal logic programs, *PAQUI LUCIO, FERNANDO OREJAS, ELVIRA PINO*

Une société d'agents temporels pour la supervision de systèmes industriels.

Mohamad K. Allouche

Thèse d'informatique de l'Université Jean Monnet (Saint Etienne) et de l'École Nationale Supérieure des Mines de Saint Etienne., soutenue le 15 octobre 1998.

La complexité croissante des systèmes industriels en termes d'expansion, d'hétérogénéité et de décentralisation entraîne de plus en plus de contraintes dans leur fonctionnement. Ces contraintes ont souvent une composante temporelle. Des techniques telles que les systèmes multi-agents, sont bien adaptées au caractère hétérogène et décentralisé de ces systèmes, elles sont de plus en plus souvent utilisées dans les applications telles que : Contrôle, Supervision, Simulation, Pilotage... Des travaux existent sur la représentation du temps dans la modélisation et le raisonnement sur les systèmes industriels. Dans le contexte multi-agents, peu de travaux prennent en compte ce facteur temporel explicitement et il est quasiment absent de tous les aspects sociaux de l'agent. Dans cette thèse, nous proposons et définissons une société d'agents temporels qui permet la prise en compte explicite du temps à la fois dans le raisonnement individuel et le raisonnement social au sein de l'agent. L'accent est mis d'une part sur l'interaction et d'autre part sur les différents types de dépendances qui existent entre agents et leur évolution au cours du temps. Ce modèle générique, a été appliqué à la supervision par reconnaissance de scénarios temporels, domaine dans lequel l'exécution des différentes tâches est extrêmement contrainte par le temps. Le système STARS est une implémentation du modèle. Il permet une reconnaissance distribuée des scénarios par un ensemble d'agents utilisant un raisonnement temporel.

Mots clés : systèmes multi-agents, intelligence artificielle distribuée, réseaux de dépendances, raisonnement temporel, supervision.

Adresse : Mohamad K. Allouche, Pavillon A. Pouliot — Université Laval, Ste-Foy, Québec, Canada
Mél : Mohamad.Allouche@ift.ulaval.ca

Une approche ordinale de la décision dans l'incertain, axiomatisation, représentation logique et application à la décision séquentielle

Régis Sabbadin

Thèse de l'Université Paul Sabatier (Toulouse III), soutenue le 9 décembre 1998 à l'Université Paul Sabatier.

La première partie de cette thèse consiste en un état de l'art de la décision dans l'incertain, considérée du point de vue de l'Intelligence Artificielle. En premier lieu sont décrites certaines approches classiques de la décision dans l'incertain, dont la théorie de l'utilité espérée puis certaines approches non-classiques, numériques ou ordinales. Ensuite, nous évoquons deux domaines où la décision dans l'incertain interagit avec l'IA : les processus décisionnels markoviens et leur application à la planification dans l'incertain et la représentation logique des préférences. La seconde partie constitue l'apport spécifique de cette thèse. Elle est divisée en trois sous-parties reprenant les thèmes de la première partie :

1) Nous étudions des critères qualitatifs de décision dans l'incertain prenant leurs valeurs dans une échelle finie, totalement ordonnée et nous en proposons une justification axiomatique. Ces critères sont fondés sur une intégrale de Sugeno qui peut être considérée comme une contrepartie qualitative de l'intégrale de Choquet. Parmi les critères axiomatisés on retrouve, entre autres, les deux fonctions d'utilité qualitative possibiliste proposées par Dubois et Prade.

2) Nous étudions une contrepartie possibiliste des processus décisionnels markoviens, totalement et partiellement observables et leur application à la planification sous incertitude et nous proposons un certain nombre d'algorithmes de résolution de type « programmation dynamique ».

3) Nous proposons enfin un langage structuré de représentation des problèmes de décision sous incertitude qualitative. Ce langage est fondé sur la logique propositionnelle (valuée) et les « systèmes de maintien de la cohérence fondés sur les hypothèses » (ATMS). Le langage proposé s'adapte à la représentation de problèmes de décision sous incertitude, non seulement lorsque celle-ci est du type « possibiliste », mais aussi lorsqu'elle est représentée par des « fonctions de croyance ». Des méthodes et algorithmes de résolution sont proposés, utilisant des procédures de recherche de modèles du type Davis et Putnam.

Mots clés : Aide à la décision, intelligence artificielle, incertitude et préférences ordinales.

Adresse : R. Sabbadin, IRIT-UPS, Bureau 326, 118 route de Narbonne, 31062 Toulouse Cedex
Mél : sabbadin@irit.fr
Adresse Web : http://www.irit.fr/

Analyse de génomes microbiens, apports de la classification pyramidale.

Jean-Christophe Aude

Thèse de l'université Paris IX Dauphine, soutenue le 18 janvier 1999.

Cette thèse s'inscrit dans le cadre des méthodes d'analyse de données, plus particulièrement de la classification automatique. La forte augmentation des performances des systèmes de séquençage automatique a contribué à la croissance exponentielle de la taille des banques de séquences nucléiques et protéiques. Les séquences complètes d'une dizaine de génomes microbiens sont ainsi déjà disponibles. Les informations contenues dans ces banques permettent d'envisager une classification globale des gènes, ce qui apporte une connaissance fondamentale sur l'organisation, l'évolution et la fonction. L'analyse par des moyens traditionnels ne peut plus répondre aux exigences d'une analyse systématique. L'objet de cette thèse est la conception et l'implémentation d'un ensemble de méthodes permettant la classification inter- et intra-génomes. Pour résoudre ce problème nous nous sommes plus particulièrement focalisé sur les points suivants :

- La détermination d'un indice de dissimilarité entre les séquences protéiques. Ceci nous a conduit à définir de nouvelles méthodes d'estimation de la qualité d'un score par une étude approfondie de la méthode du Z-score.
- La classification pyramidale est une généralisation de la classification hiérarchique. Nous présentons l'intérêt de cette méthode dans le cadre de l'analyse de séquences biologiques. Enfin, nous avons étudié un algorithme de construction pyramidale par optimisation de la matrice robinsonienne, en utilisant les régressions isotones.
- Une méthode de classification composite, qui consiste à utiliser l'algorithme du lien simple, puis à effectuer une classification hiérarchique des sous-graphes obtenus, et enfin réaliser une classification pyramidale sur chaque composante.

Ce travail est complété par la présentation des résultats induits par la méthode, « les premières lois de la génomique ». Pour réaliser cette étude, nous avons développé plusieurs outils informatiques : PLATOO, une bibliothèque de fonctions C permettant l'analyse des résultats de comparaisons de séquences calculés par le logiciel LASSAP ; J/DrawPyr, des programmes écrits en langages C et Java permettant le dessin et la manipulation interactive d'une pyramide.

Contact : Fabrice Rossi
WEB : http://www.ceremade.dauphine.fr/~rossi/lise/seminaire.html
et http://cao.inria.fr/~aude/These
Mél : rossi@ceremade.dauphine.fr

Représentations de connaissances : de l'approximation à la confrontation

Jérôme Euzenat

Habilitation à diriger des recherches de l'Université Joseph Fourier, soutenue le 20 janvier 1999 à Grenoble.

Résumé.

Un formalisme de représentation de connaissances a pour but de permettre la modélisation d'un domaine particulier. Bien entendu, il existe divers langages de représentation de connaissances et, au sein d'un même langage, divers modèles peuvent représenter un même domaine. Ce mémoire est consacré à l'étude des rapports entre des représentations multiples de la même situation. Il présente les travaux de l'auteur entre 1992 et 1998 en progressant de la notion d'approximation, qui fonde la représentation, vers la confrontation entre les diverses représentations. Tout d'abord la notion d'approximation au sein des représentations de connaissance par objets est mise en avant, en particulier en ce qui concerne l'ensemble des mécanismes tirant parti de la structure taxonomique (classification, catégorisation, inférence de taxonomie). À partir de la notion de système classificatoire qui permet de rendre compte de ces mécanismes de manière unique, on montre comment un système de représentation de connaissances peut être construit. Le deuxième chapitre introduit la possibilité de tirer parti de taxonomies multiples (sur le même

ensemble d'objets) dans un système de représentation de connaissances. La multiplicité des représentations taxonomiques est alors introduite et justifiée. Ces taxonomies multiples sont replacées dans le cadre des systèmes classificatoires présentés auparavant. La notion de granularité, qui fait l'objet du troisième chapitre, concerne la comparaison de représentations diverses de la même situation. À la différence des autres chapitres, il n'est pas situé dans le cadre des représentations de connaissances par objets mais dans celui des algèbres de relations binaires utilisées pour représenter le temps et l'espace. Le quatrième chapitre, enfin, va vers la confrontation des différentes représentations pour en tirer le meilleur parti (obtenir une représentation consensuelle ou tout simplement une représentation cohérente). Le but des travaux qui y sont présentés est de développer un système d'aide à la construction collaborative de bases de connaissance consensuelles. À cette fin, les utilisateurs veulent mettre dans une base commune le contenu de leurs bases de connaissances personnelles. Pour cela, deux problèmes particuliers sont traités : la conception d'un mécanisme de révision, pour les représentations de connaissances par objets (permettant aux utilisateurs de traiter les problèmes d'incohérence) et la conception d'un protocole de soumission de connaissances (garantissant l'obtention d'une base commune consensuelle). Cet aperçu partiel des travaux possibles dans l'étude des relations entre représentations est limité, mais il met en évidence le caractère non impératif des solutions proposées qui s'appliquent bien au cadre où le modélisateur interagit avec le système de représentation.

Mots clés : représentation de connaissances, approximation, bases de connaissances, modélisation, représentation par objets

Contact : Jérôme Euzenat, INRIA Rhône-Alpes, 655, avenue de l'Europe,
38330 Montbonnot St Martin
Mél : Jerome.Euzenat@inrialpes.fr
WEB ftp://ftp.inrialpes.fr/pub/sherpa/theses/hdr-euzenat.ps.gz

Validité de l'approche classification dans la réduction statistique de la complexité de #SAT

Valérie Rouat

Thèse de l'Université de Rennes 1, soutenue le 22 janvier 1999

Le problème abordé dans cette thèse est celui du dénombrement des solutions du problème de la satisfiabilité d'une formule booléenne sous forme normale conjonctive (problème #SAT). Ce problème est #P-complet et les résultats actuels montrent que la résolution de #SAT, même approchée, est, dans le pire des cas, exponentielle en temps. Notre approche utilise une algorithmique issue de l'analyse combinatoire des données et s'appuie sur une représentation ensembliste, géométrique et logique des clauses définie par une bijection entre clauses et sous-ensemble de l'ensemble des interprétations. Une telle représentation permet une vision synthétique du problème #SAT et la prise en compte de caractéristiques statistiques globales de l'instance. En appliquant le principe général « diviser pour résoudre », la méthode de résolution approchée de #SAT que nous proposons permet de réduire de façon considérable la complexité algorithmique du problème. Elle est fondée sur la segmentation d'une sériation établie sur la table d'incidence associée à la formule. Nous montrons, dans le cas difficile des instances SAT aléatoires, l'intérêt de la sériation et de sa meilleure coupure en deux parties connexes et de tailles comparables. De plus, nous définissons la notion d'indépendance en probabilité entre clauses et entre ensembles de clauses. La méthode proposée est validée à la fois théoriquement et par une vaste expérimentation. Par ailleurs, nous nous intéressons à la distribution de probabilité du nombre de solutions d'une instance kSAT aléatoire. L'espérance et la variance du nombre de solutions sont connues, nous montrons en utilisant des tests statistiques que la distribution suit une loi log-normale.

Mots clés : satisfiabilité, théorie de la complexité, dénombrement de solutions, problèmes #P-complets, analyse combinatoire des données, classification.

Adresse : Valérie Rouat, IRISA, Campus universitaire de Beaulieu,
35042 Rennes Cedex, France.
Mél. : Valerie.Rouat@irisa.fr
Adresse Web : http://www.irisa.fr/aida/Valerie.Rouat/
Thèse au format électronique : http://www.irisa.fr/EXTERNE/bibli/theses/theses99/rouat/rouat.html

Contribution à la représentation et à la vérification des modèles de connaissances produits en ingénierie d'ensembles mécaniques.

Florence Sellini

Thèse de l'École Centrale de Paris, soutenue le 5 mars 1999.

La capitalisation du savoir-faire, de la gestion des connaissances désigne la préservation, l'enrichissement, le partage, la réutilisation d'un patrimoine construit depuis que l'entreprise existe. Le travail présenté se situe à l'intersection des travaux sur les mémoires d'entreprise, des travaux sur la modélisation des connaissances attachées à un produit et des travaux concernant la validation des systèmes à base de connaissances (SBC). Notre étude se place dans le domaine des systèmes d'aide à la conception d'ensembles mécaniques, fondés sur les connaissances de savoir-faire. C'est une démarche qui vise à coupler la capitalisation des connaissances du savoir-faire en conception, avec la puissance des modeleurs géométriques existants (systèmes CAO). Cette modélisation du savoir-faire est constituée d'une part, de la représentation du produit à concevoir (modèle produit) et d'autre part, de la représentation du processus de conception. L'objectif est de diminuer les temps de reconception, en réutilisant ce qui est déjà éprouvé et stable, pour se concentrer uniquement sur les nouveaux choix à gérer pour une nouvelle solution. Notre travail de recherche s'inscrit en complément de cette démarche à partir d'un besoin exprimé, celui de la vérification de l'ensemble des connaissances « mises en boîte ». Dans ce cadre, notre contribution concerne le contrôle a priori de la modélisation et le champ de l'étude a été restreint au modèle de représentation du produit. Elle se situe dans une double problématique. Les modèles ont été enrichis de façon à rendre les connaissances explicites pour permettre une action de vérification plus efficace. À partir de cette évolution, des mécanismes de vérification ont été spécifiés, en ayant au préalable défini les connaissances nécessaires (règles d'écriture, cas d'incohérences). Ces mécanismes ont été mis en cohérence pour qu'ils soient constitutifs d'une démarche d'édition et de contrôle de modèles. Cela pourrait constituer un premier pas vers la construction d'un atelier d'édition et de maintenance de modèles de connaissances, directement utilisable par l'expert-concepteur. Celui-ci pourrait ainsi modifier les connaissances représentées dans les modèles de l'application en ayant la garantie d'être assisté pour sa mise à jour correcte de la BC métier.

Mots clés : représentation de connaissances, vérification, modèles objets et contraintes, savoir-faire en conception mécanique, mémoire d'entreprise.

Contact : Florence Sellini
PSA Peugeot Citroën & ISMCM-CESTI
Mél : Florence. Sellini@wanadoo.fr

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

• saisir sous traitement de texte compatible avec Word si possible, en times 10, interligne moyen (ou équivalent);
• présentation (pour un doctorat d'Université) :

<Titre de la thèse>
<auteur>
Thèse de l'Université <université>,
soutenue le <date> [à <lieu de soutenance>
si ≠ de <université>]
[préparée à <lieu de préparation> si ≠ de <université>]

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

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

APPELS À DOSSIER

Dossier « Plates-formes de développement de systèmes multi-agents »

à paraître dans le Bulletin de l'AFIA n° 39 en octobre1999

Coordinateurs : le groupe ASA (Architecture des Systèmes à Agents) de l'AFIA et plus particulièrement Olivier Boissier (SIC - ENSM.SE), Zahia Guessoum (LERI-Reims, LIP6) et Michel Occello (MAGMA LEIBNIZ – IMAG)

Objectif

Les systèmes multi-agents proposent un ensemble de modèles et d'outils pour le développement de systèmes à plusieurs composantes autonomes pouvant coopérer entre elles. Ce domaine de recherche propose de multiples approches. Les architectures logicielles qui en découlent sont donc nombreuses et variées. Dans de nombreux cas, les propositions sont conceptuelles, sans être étayées par des réalisations pratiques permettant de les valider ; l'une des raisons principales est la difficulté de réalisation de systèmes multi-agents.

Ce dossier a pour but de faire le point sur les différentes plates-formes de développement de systèmes multi-agents que propose la communauté francophone actuellement et montrer l'apport de ces plates-formes par rapport aux outils classiques (objets, systèmes experts, ...). Il s'agit d'une part de présenter les principales composantes d'une plate-forme de développement de systèmes multi-agents et la différence avec les différentes techniques existantes, et d'autre part, de cerner les caractéristiques des domaines d'application de telles plates-formes.

Qui peut soumettre ?

Toute équipe de recherche de la communauté francophone ayant développé ou étant en train de développer une plate-forme de construction de systèmes multi-agents.

Présentation des contributions

Les contributions s'organiseront selon un questionnaire élaboré par le groupe ASA. Celui-ci suit le schéma suivant :

- *Caractéristiques générales,* dont l'objectif est de fournir une fiche d'identité de la plate-forme,
- *Modèles multi-agents :* description des concepts multi-agents disponibles sur la plate-forme,
- *Caractéristiques physiques :* description des caractéristiques physiques des systèmes multi-agents réalisables au travers de la plate-forme,
- *Environnement de développement,*
- *Environnement d'exécution.*

Le questionnaire est disponible à : http://www-poleia.lip6.fr/~guessoum/asa/criteres.html

Format des contributions

Les contributions devront faire environ une page et sont à envoyer avant le 15 août 1999 sous forme d'un fichier texte ou word à Zahia.Guessoum@lip6.fr. Le texte ne doit comporter ni caractère de tabulation, ni tableau, ni aucune mise en forme autre que gras ou italique.

En principe, vous avez dû signaler votre intention de contribuer à ce dossier avant le 1er juillet 1999 à l'un des trois coordonnateurs (boissier@emse.fr, Zahia.Guessoum@lip6.fr, Michel.Occello@imag.fr).

Appel à contributions

Dossier IA et EIAH
« Intelligence Artificielle et Environnements Informatiques pour l'Apprentissage Humain »

Coordinateurs : Nicolas Balacheff (LEIBNIZ — IMAG), Monique Baron (SysDeF-LIP6)
Mél. : Nicolas.Balacheff@imag.fr, Monique.Baron@lip6.fr

Date de parution prévue : janvier 2000

Objectifs

Depuis le dossier EIAO (« Environnements Interactifs d'Apprentissage avec Ordinateur ») rassemblé par J.-F. Nicaud en avril 1992 dans le bulletin n° 9, trois éditions des Journées Francophones EIAO ont eu lieu (1993, 1995 et 1997, à l'ENS de Cachan), et la France accueille en juillet 1999 la conférence internationale AI-ED'99 (Artificial Intelligence in Education), au Mans (comité d'organisation présidé par M. Vivet).

Les problématiques de recherche en EIAO ont évolué, ces dernières années, pour prendre en compte les développements des TIC (technologies de l'information et de la communication) et de leurs usages (réels ou prévisibles), qui concernent en particulier le multimédia et le multimodal, d'une part, ainsi que la communication à distance et en réseau, d'autre part. C'est cette évolution que veut marquer le nouveau sigle EIAH, qui fait référence à la conception de systèmes complexes, qui peuvent être distribués dans l'espace et dans le temps, pour favoriser des apprentissages chez des « apprenants » humains en situation de formation initiale, continue ou professionnelle.

Ce dossier est l'occasion de faire le point sur les problématiques et sur les travaux situés à la croisée de ce domaine pluridisciplinaire et de l'Intelligence Artificielle. Toute contribution des chercheurs francophones du secteur privé ou public est bienvenue.

Merci d'envoyer, dès que possible, votre intention de participer à ce dossier, aux deux coordinateurs, par courrier électronique.

Présentation des contributions

Les contributions expliqueront la problématique visée et présenteront les concepts, méthodes, techniques, expériences mises en oeuvre. Elles pourront être organisées selon le schéma suivant :

- identification de l'équipe (adresse, téléphone, mél)
- chercheurs concernés par le thème
- cadre général et objectifs, problématique
- panorama des travaux effectués par l'équipe sur le thème
- bibliographie des travaux de l'équipe sur ce thème (réf. principales récentes)

Le volume souhaité des contributions est d'environ une page, elles sont à envoyer

***** avant le 1er novembre 1999 *****

aux deux coordinateurs, par courrier électronique, au format Word (compatible Word Mac si possible) ou RTF dépourvu de toute mise en forme (ni tableau ni tabulation…) hormis le gras et l'italique.

Adhésion et abonnement ❑ Demande ❑ Renouvellement

Nom : Prénom :
Affiliation :
Adresse postale :
N° de téléphone : N° de télécopie :
Adresse électronique :
Activités :
à titre professionnel amateur *(rayer la mention inutile)*

- ❑ Adhésion simple 1999 (l'abonnement au bulletin inclus) : ..... 290 francs
- ❑ Adhésion doctorant ou étudiant (sur justificatif) 1999 : ..... 100 francs
- ❑ Adhésion de soutien1999 (l'abonnement au bulletin inclus) : ..... 580 francs
- ❑ Abonnement au bulletin sans adhésion (1999) : ..... 265 francs
- ❑ Adhésion 1998 (incluant l'abonnement au bulletin)
  + abonnement à la Revue de l'Intelligence Artificielle : .....(individu) 745 francs
  (réglé par institution) 940 francs
- ❑ adhésion au collège IADSMA : ajouter ..... 100 francs

(si vous avez déjà payé votre cotisation AFIA et que vous souhaitez adhérer au collège, ne payez que 100 F)

Veuillez trouver un chèque de ..... francs

Adhésion Personne morale ❑ Demande ❑ Renouvellement

Organisme : cachet de l'organisme:

Nom du représentant : Prénom :
Fonction :
Adresse postale :
N° de téléphone : N° de télécopie :
Adresse électronique :

Tarifs 1999

Laboratoires universitaires : **1507,50 F TTC** (1250 F HT)
Personnes morales non universitaires : **3135,60 F TTC** (2600 F HT)

- ❑ je joins un bon de commande pour une somme de _____ francs
- ❑ j'accepte que les renseignements ci-dessus apparaissent dans l'annuaire de l'AFIA
- ❑ j'accepte que les renseignements ci-dessus soient communiqués a l'ECCAI dans le cadre de la constitution en cours d'un fichier IA européen.
- ❑ je m'oppose à toute diffusion des renseignements ci-dessus.

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

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

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

Les dossiers du Bulletin de l'AFIA

IA et WEB	Bulletin n°38
Mémoires d'entreprises	Bulletin n°36
IA et logique	Bulletin n°35
Ingénierie des connaissances	Bulletin n°34
IA et Télécommunications	Bulletin n°33
IA et Terminologie	Bulletin n°32
Décision et IA	Bulletin n°31
Raisonnement IA et Image	Bulletin n°30
Raisonnement temporel et spatial	Bulletin n°29
Systèmes Multi-agents	Bulletin n°28
IA et robotique	Bulletin n°27
I.A . et biologie moléculaire	Bulletin n°26
I.A. et droit	Bulletin n°25
I.A. et fusion de données	Bulletin n°24
I.A. et musique	Bulletin n°23
Apprentissage	Bulletin n°22
Les explications dans les SBC	Bulletin n°20
Pétrole-Chimie	Bulletin n°19
Le raisonnement à partir de cas	Bulletin n°18
I.A. et temps-réel	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

SOMMAIRE DU BULLETIN N°38

La vie de l'A.F.I.A. 4
Débats 8
IA et Internet 11
Présentation de Laboratoires 13
IA et WEB 19
Livres 41
Calendrier 42
Sommaire des revues 43
Résumés habilitations et thèses 50
Appels à dossiers 54

CALENDRIER DE PARUTION DU BULLETIN DE L'AFIA

Hiver
Réception des contributions:
15 décembre
Sortie le **31 janvier**

Eté
Réception des contributions:
15 juin
Sortie **31 juillet**

Printemps
Réception des contributions:
15 mars
Sortie le **30 avril**

Automne
Réception des contributions:
30 septembre
Sortie le **31 octobre**