Partie I Présentation du laboratoire page 3

Présentation
- Équipes
- Environnement universitaire
- Production scientifique
- Contrats et les relations industrielles
- Recrutements
- Budget
- Administration et technique
- Implication externe

5

Organi-gramme

17

Partie II Activités de recherche page 19

A^3
- Membres de l'équipe
- Fondements scientifiques et orientations de recherche
- Composante Apprentissage Numérique
- Composante Apprentissage Symbolique
- Approches Transversales
- Animation de la recherche
- Publications principales

21

LCR
- Membres de l'équipe
- **Combinatoire algébrique**
- **Logique et Programmation**
- **Spécification et Vérification**
- **Bases et Entrepôts de Données**
- Séminaires d'équipe

69

OCAD
- Membres de l'équipe
- Présentation scientifique générale
- **Optimisation Combinatoire**
- **Algorithmique Distribuée**
- Séminaire

115

RCLN
- Membres de l'équipe
- Orientation scientifique générale
- Thèmes de recherche
- Bilan et perspectives de recherche
- Synthèse
- Animation de la recherche
- Formation à la recherche
- Publications principales

149

Partie III Activités du laboratoire page 183

Formation par la recherche
- Master Informatique
- Master Mathématique Informatique
- Thèses et Habilitations

185

Organisation de conférences

203

Partie IV Publications page 205

Partie V Acronymes page 251

Partie I
Présentation du laboratoire

Chapitre 1
Le LIPN : état des lieux et évolutions

Sommaire

1.1 Équipes **6**
1.2 Environnement universitaire **8**
1.3 Production scientifique **8**
1.4 Contrats et les relations industrielles **9**
1.5 Recrutements **11**
1.6 Budget **11**
1.7 Administration et technique **12**
1.7.1 Locaux 12
1.7.2 Environnement technique 12
1.8 Implication externe **14**

Le Laboratoire d'Informatique de l'Université Paris-Nord (LIPN[1]) est associé au CNRS depuis janvier 1992 et a le statut d'UMR depuis janvier 2001. Le LIPN poursuit ses recherches en automatisation du raisonnement autour de ses axes forts, l'*Optimisation Combinatoire*, l'*Informatique Fondamentale* et l'*Intelligence Artificielle* en s'appuyant

[1] http://www-lipn.univ-paris13.fr. (On trouvera en annexe la liste des acronymes et leur signification)

sur les compétences de ses membres, en particulier en Algorithmique, Logique, Langage Naturel, Apprentissage. Ces recherches sont effectuées dans quatre équipes par 57 chercheurs et enseignants-chercheurs permanents, en poste pour la plupart à l'université Paris 13 (sur l'Institut Galilée, dans les IUT de Villetaneuse et de Bobigny). Au-delà de ces chercheurs et enseignants-chercheurs permanents, c'est en fait une centaine de personnes qui participent aux activités du laboratoire.

Le développement du laboratoire s'est accompagné ces dernières années d'un recentrage des recherches autour des thèmes principaux tout en s'ouvrant à de nouvelles problématiques prometteuses. Nous privilégions ainsi les recrutements dans nos domaines forts afin d'acquérir les masses critiques suffisantes pour une recherche de pointe. De même, nous profitons de la mise en place de masters sur l'université Paris 13 (master «Modélisation Informatique des Connaissances et du Raisonnement», master «Mathématique-Informatique») et du master effectué en collaboration avec l'université d'Antananarivo (Madagascar) pour améliorer l'offre de formation à la recherche et mieux accueillir les étudiants en thèse. Ces conditions sont en effet nécessaires pour une production scientifique qualitativement et quantitativement importante. Ce sont aussi les éléments indispensables au questionnement permanent que nous devons avoir sur les objectifs de nos recherches pour lesquelles nous cherchons à conserver l'équilibre théorie / application actuel. Cet équilibre, concrétisé par un budget au 2/3 provenant de contrats, apparaît globalement dans chaque thématique. Le LIPN a ainsi choisi de s'impliquer dans des projets de grande envergure (projets européens du 6ème PCRDT, pôles de compétitivité, ANR), tout en continuant à participer activement aux projets sur l'université Paris 13, en particulier dans le domaine du traitement des langues ou en apprentissage en collaborant avec d'autres laboratoires de l'université (LLI, Lim&Bio, LEEC).

1.1 Les équipes

Depuis 2002, le LIPN a, par ses recrutements et la mise en place de projets, renforcé les thématiques principales de trois de ses équipes, et restructuré l'équipe «Apprentissage, Diagnostic et Agents» (Adage) en une équipe orientée apprentissage («Apprentissage Artificiel et Applications», A^3) après plusieurs départs de permanents. Le Laboratoire est maintenant structuré en quatre équipes dont les sujets de recherche présentent de nombreux points de rencontre :

- «Apprentissage Artificiel et Applications» (A^3), qui développent des méthodes d'apprentissage innovantes en réponse à de nombreux besoins applicatifs. Les méthodes utilisées sont symbolique, numériques ou transversales.
- «Logique, Calcul et Raisonnement» (LCR), dont les intérêts se sont recentrés essentiellement sur la logique linéaire et ses applications en informatique, les spécifications de systèmes et l'aide à la modélisation, et la combinatoire algébrique.
- «Optimisation Combinatoire et Algorithmique Distribuée» (OCAD), où sont abordés des problèmes liés à l'optimisation en nombres entiers, la satisfaction et l'optimisation des systèmes de contraintes, la combinatoire et l'algorithmique pour les systèmes parallèles et distribués.
- «Représentation des Connaissances et Langage Naturel» (RCLN), qui a pour objectif d'améliorer les techniques de représentation de la sémantique des langues à la fois pour obtenir une meilleure adéquation des modèles formels et pour développer des outils efficaces d'accès à l'information textuelle ; son approche pluridisciplinaire lui confère une place significative dans le développement actuel des Sciences Cognitives.

Le laboratoire s'était fixé comme orientations scientifiques dix champs pour la période 2003-2006. On trouvera ci-dessous quelques éléments factuels indiquant dans quelle

mesure ceux-ci ont pû être mis en chantier. Ces points sont repris et développés dans les sections consacrées aux équipes.

Méthodes exactes et heuristiques pour l'optimisation combinatoire (OCAD) : deux recrutements de MCF ont eu lieu sur ce thème (un nouveau recrutement MCF s'effectue à la rentrée 2007), auquel s'ajoute l'intégration au laboratoire d'un professeur assistant de l'ESSEC ; ces recrutements contrebalancent les trois départs sur la même période. Une ACI a été obtenue dont les résultats sont notables. Le laboratoire a soutenu directement ce sujet par l'octroi d'un fonds supplémentaire. Ce domaine est partie intégrante du master recherche mathématique-informatique conjoint avec le LAGA ouvert en 2005, c'est aussi le cas dans les deux points suivants.

Combinatoire, analyse de structures discrètes et algorithmique distribuée (OCAD) : deux MCF et un CR ont été recrutés sur ce thème, auxquels il faut ajouter un PU en combinatoire algébrique (cf point suivant). Les recherches abordées portent sur les propriétés de familles d'objets combinatoires ainsi que sur l'analyse perturbative.

Combinatoire et Physique (LCR) : l'équipe est maintenant bien constituée après le recrutement d'un PU (et d'un MCF à la rentrée 2007). Celle-ci a constitué un groupe de travail en interaction avec le LAGA et le laboratoire d'informatique de Marne-la-Vallée. La principale activité en combinatoire et physique a été de créer des algèbres (associatives, de Hopf et de Lie) adaptées à certains problèmes de physique mathématique, de déterminer leurs déformations quantiques et de calculer leurs constantes de structure. Cette équipe a eu le soutien pendant trois mois d'un chercheur associé au CNRS.

Logique linéaire, interactivité et modélisations (LCR) : l'équipe s'est renforcée après les arrivées de trois MCF et deux CR CNRS. Elle a accueilli ou accueille quatre post-doc sur cette période. Le groupe participe à 1 ANR et 2 ACI pour lesquels des avancées significatives ont été obtenues en particulier dans la caractérisation de la complexité implicite des programmes. Depuis deux ans, les approches à base de logique linéaire, pour lesquelles le groupe a une compétence reconnue, ont été confrontées à d'autres approches (quasi-interprétations, modèle BSS) après la venue de J.Y. Moyen d'abord comme Post-doc puis comme MCF. Ce dernier recrutement renforce de plus les interactions avec la thématique Spécification.

Spécification, modélisation, validation, et certification de logiciels (LCR) : 1 PU et 1 MCF sont venus compléter le groupe. L'équipe a maintenant un potentiel fort qui lui permet de s'inscrire dans les recherches les plus en pointe en méthodologie de spécification, en vérification modulaire ou dans les techniques de raffinement de spécifications.

Analyse du contenu textuel (RCLN) : Recrutement de 1 PU, 1 CR, 1 MCF. L'équipe participe activement à de nombreux projets en extraction et analyse de données textuelles. Il faut noter que celle-ci est maître d'œuvre dans un PPF débuté en 2004 et comprenant deux autres laboratoires de l'Université Paris 13. Rappelons enfin qu'Adeline Nazarenko a reçu la médaille de bronze CNRS en 2004.

Ingénierie des connaissances à partir de textes (RCLN) : La mutation d'une MCF a compensé un départ. Sur ce sujet d'actualité, le groupe contribue activement aux recherches au travers de développements logiciels comme Terminae, de collaborations extérieures importantes et d'un projet ANR. Le logiciel Terminae, développé depuis quelques années au LIPN, est maintenant diffusé et exploité par d'autres laboratoires.

Analyse sémantique et apprentissage (RCLN + Adage/A^3) : le recrutement récent de MCF à la frontière de ces deux domaines et la restructuration de l'équipe Adage en une équipe «Apprentissage» ont permis de nouer des collaborations fructueuses, notamment dans le projet Infom@gic du pôle de compétitivité CAP DIGITAL auquel les deux équipes RCLN et A^3 participent conjointement.

Apprentissage et exploration de données (Adage, maintenant équipe A³) : ce domaine s'est considérablement développé au LIPN tant quantitativement (2 MCF et 1 PU) que qualitativement. Les recherches qui sont menées ont naturellement abouti à des projets d'envergure (ANR, pôle de compétitivité, projets industriels) avec des collaborations variées (de l'étude du comportement de fourmis à l'analyse d'un réseau de capteurs).

Diagnostic autonome, évolutif et distribué (Adage) : la mutation de Ph. Dague au LRI en 2005 a conduit le laboratoire à recentrer la thématique de l'équipe autour de l'apprentissage. Si ce thème ne fait plus partie des orientations stratégiques, il n'en reste pas moins un domaine naturel d'applications pour les recherches en apprentissage.

De nouveaux domaines de recherche ont émergé ces dernières années, parmi lesquelles on peut citer :

- le développement d'un langage de programmation fortement typé pour le web (avec une ANR et un projet Marie Curie),
- l'enrichissement des langages d'entrepôts de données par intégration de données hétérogènes.

Ayant acquis des compétences et les masses critiques nécessaires, le laboratoire compte poursuivre ces recherches pour le prochain quadriennal dans les domaines suivants :

- *Méthodes exactes et heuristiques pour l'optimisation combinatoire*
- *Combinatoire, analyse de structures discrètes et algorithmique distribuée*
- *Logique linéaire, interactivité et modélisations*
- *Spécification, modélisation, validation, et certification de logiciels*
- *Langage de programmation fortement typé pour le web*
- *Langages d'entrepôts de données par intégration de données hétérogènes*
- *Ingénierie des connaissances à partir de textes*
- *Analyse sémantique et apprentissage*
- *Apprentissage et exploration de données*

1.2 L'environnement universitaire

Le LIPN s'inscrit dans la politique scientifique de l'université en participant à un projet de l'université pour le contrat quadriennal 2005-2008, après avoir participé à 2 projets importants pour le contrat quadriennal précédent. La création de la Maison des Sciences de l'Homme Paris Nord a ouvert des perspectives de collaboration intéressantes d'une part dans le domaine du langage entre le Laboratoire de Linguistique Informatique (LLI) et l'équipe RCLN, d'autre part dans le domaine de l'image avec Younès Bennani de l'équipe A³. Les équipes RCLN et Adage-A³ avaient participé également activement à la mise en place du laboratoire d'informatique médicale et bioinformatique (Lim&Bio) qui a été en 2002 l'initiateur d'une équipe-projet du CNRS. Fort de ce contexte, le LIPN est maître d'œuvre pour le plan quadriennal 2005-2008 d'un programme pluri-formation sur le thème de la linguistique informatique appliquée au domaine des textes médicaux, en collaboration avec le LLI et le Lim&Bio.

1.3 La production scientifique

La production scientifique a été de grande qualité, tant en termes de publications dans des revues de haut niveau ou dans les conférences internationales les plus importantes dans les différents thèmes de recherche, qu'en termes de diffusion de la connaissance. Les informations données en figure 1.1 quantifient cette production. On trouvera en annexe la

liste exhaustive des publications. Notons que le nombre de publications dans des revues nationales ou internationales était d'une moyenne de 17 par an pour la période 1997-2000, de 19,5 pour la période 2001-2003 et est de plus de 25 depuis 2004.

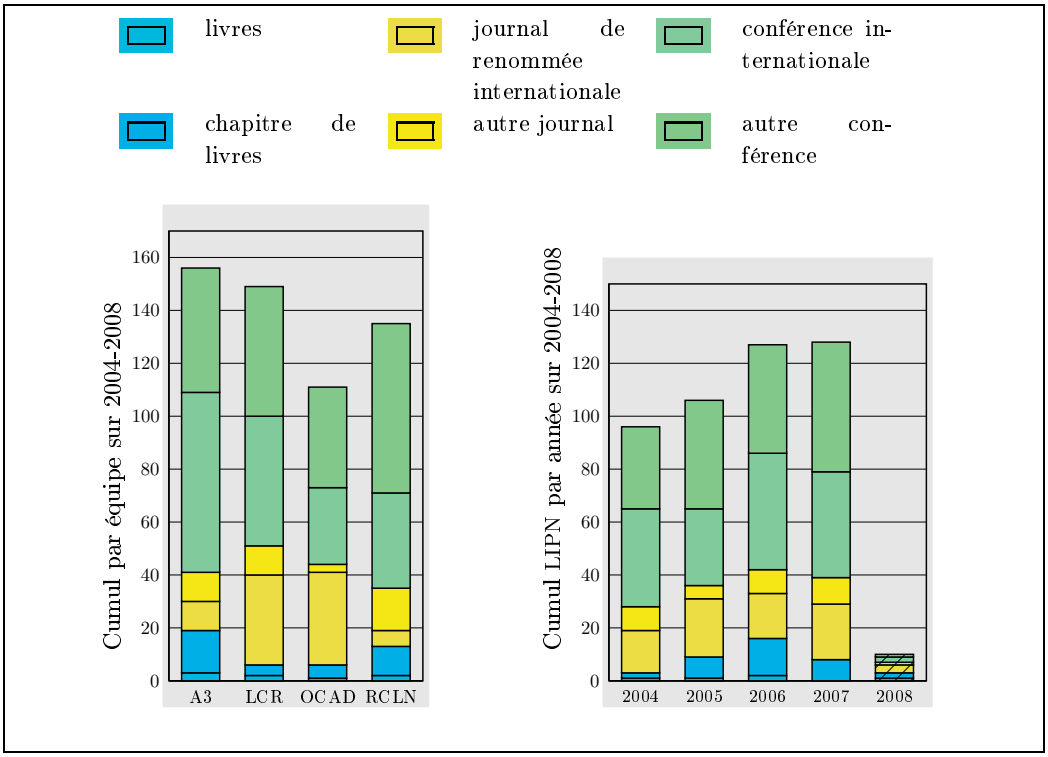


FIG. 1.1: Production scientifique principale (2008 en cours)

1.4 Les contrats et les relations industrielles

Le LIPN développe des recherches fondamentales tout en intensifiant une politique de coopération avec les grands organismes et les industriels, politique qui s'est traduite par un accroissement de sa participation dans les actions RNRT et RNTL, et maintenant sur les projets ANR, des contrats avec France Télécom R&D et différents ministères, et une collaboration avec des entreprises dans le cadre soit de projets de recherche, soit de bourses CIFRE. Le LIPN participe aussi de manière importante à de grands projets régionaux (INFOM@GIC dans le pôle de compétitivité CAP DIGITAL) ou nationaux (projet Quaero, ..)

On trouvera ci-dessous la liste des principaux contrats industriels ou académiques du LIPN obtenus ou achevés sur ces quatre dernières années. Ceux-ci sont donnés avec un mot clé caractérisant le champ disciplinaire. Tous les champs de recherche du laboratoire sont couverts. Ces projets sont décrits plus avant dans les chapitres relatifs aux équipes concernées. Remarquons que le LIPN a avec succès proposé des projets dans le cadre des appels de l'ANR lancés en 2005, 2006 et 2007. Neuf projets ont ainsi été retenus, dont un dans le cadre des appels ANR 2007 (les résultats ne sont pas tous connus). Il est à noter que deux de ces projets correspondent aux appels «Jeunes chercheuses et jeunes chercheurs», témoignant de la vitalité des recherches entreprises par les nouveaux permanents du laboratoire. Ces résultats renouvellent les projets précédents du laboratoire,

qui s'inscrivaient dans les cadres RNTL ou RNRT. Les appels européens du 6ème PCRDT ne permettaient guère une participation active en recherche amont, contrairement aux appels du 5ème PCRDT, d'où leur importance globale relative. Enfin, on peut sans doute regretter le faible nombre de contrats proprement industriels, à compenser par la quantité notable de conventions CIFRE pour la poursuite de thèses en milieu industriel.

Contrats européens	
IP IDD	*Diagnostic*
STREP ALVIS	*Traitement automatique des langues*
Marie Curie (TOK) WEBSICOLA	*Sûreté du web*
Contrats MENESR	
RNTL Technolangue ATONANT	*Ingénierie des Connaissances*
Technolangue CESART	*Traitement automatique des langues*
ACI Sécurité Informatique	*Logique et complexité implicite*
ACI Nouvelles interfaces des mathématiques	*Logique et programmation*
ACI Planification	*Optimisation combinatoire*
RNRT IP-SIG	*Qualité de service dans les réseaux*
RNTL ExtraPloDocs	*Traitement automatique des langues*
RNTL RESEDA	*Diagnostic*
Pôle de compétitivité INFOM@GIC	*Traitement automatique des langues et apprentissage*
Contrats ANR	
ANR SILLAGES	*Apprentissage*
ANR NOCOST	*Complexité implicite*
ANR MICRAC	*Traitement automatique des langues*
ANR CERPAN	*Grille de calcul*
ANR ARA Sécurité	*Grille de calcul*
ANR PROG. FONC. POUR LE WEB	*Sûreté du web*
ANR SEPTIA	*Apprentissage*
ANR OPTIMA	*Web et Sciences Humaines*
ANR CADI	*Apprentissage*
RNTL TEXTCOOP	*Traitement automatique des langues*
RNTL DAFOE	*Ingénierie des Connaissances et Traitement automatique des langues*
Contrats CNRS	
AS TCAN Traitement des connaissances	*Traitement automatique des langues*
Contrats CIFRE et contrats industriels	
Convention Cifre GEOBS	*Fouille de données*
Convention Cifre SNCF	*Optimisation combinatoire*
Convention Cifre Sinequa	*Traitement automatique des langues*
Convention Cifre Elda	*Traitement automatique des langues*
Convention Cifre Numsight	*Apprentissage pour le marketing*
Contrat EDF	*Résolution de systèmes de contraintes*
Contrat Numsight	*Apprentissage pour le web mining*
Contrat de veille SNCF	*Optimisation combinatoire*
Contrat de collaboration EDF	*Résolution de systèmes de contraintes*

1.5 Les recrutements

Nous poursuivons une politique de recrutement en synergie avec la commission de spécialistes 27ième section de l'université Paris 13, qui prend en compte au mieux les intérêts du laboratoire, par le renforcement de ses points forts tout en enrichissant les thèmes de recherche actuels. Chaque équipe a eu ainsi l'opportunité de recruter presque chaque année professeur ou maître de conférences. Par ailleurs, trois des quatre équipes ont maintenant (au moins) un chargé de recherche CNRS : Après Patrick Baillot en logique et Cyril Banderier en analyse d'algorithmes, Thierry Poibeau (thèse en contrat CIFRE à Thalès et au LIPN dans le domaine de l'informatique linguistique) a été affecté par le CNRS au LIPN à la rentrée 2003, Paulin Jacobé de Naurois a été admis au CNRS en 2006 (thèse en cotutelle Loria Nancy et université de Hong-Kong dans le domaine de l'informatique théorique, ATER au LIPN en 2005-2006). Outre les post-doctorants recrutés dans le cadre des projets ANR, le CNRS a octroyé un post-doctorat d'un an chacune de ces trois dernières années (Nadia Pisanti, 2004-2005 ; Jean-Yves Moyen, 2005-2006 ; Jean-François Culus, 2006-2007). En sus des professeurs invités (une dizaine de mois par an), des chercheurs associés ont aussi été accueillis au LIPN (Alfredo Viola pour 6 mois en 2003 ; Todor Popov sur 3 mois en 2004). Enfin les délégations CNRS ont été régulières (Jean-Marie Le Bars du Greyc a été accueilli deux ans ; Adeline Nazarenko a bénéficié d'un an, Frédéric Toumazet et Younès Bennani de deux ans). Ces mouvements et recrutements ont entraîné l'accroissement du nombre global de titulaires au sein du LIPN, qui était de quarante et un en 2000 et atteint soixante-trois à cette rentrée universitaire 2007-2008.

Corps	2002 (situation)	2006 (situation)	Flux
CR	2	4	2
PU	13	15	1
MCF	32	36	2
Ingénieurs	1	1+1/2	1/2
Administration.	3	3	0
Postdocs	3	4	1
Ingénieurs (CDD)	0	3	3
Doctorants	21	32	11
Total	*75*	*99*	*24*

1.6 Budget du laboratoire

Le laboratoire a un budget total supérieur à 700 Keuros par an, hors salaires des permanents et dépenses liées à l'infrastructure couvertes par l'université. La Fig. 1.2 indique les ressources directes et reports. Un tiers de ce budget provient de fonds récurrents (ministère et CNRS, cf Fig. 1.3), le reste provient majoritairement de projets de recherche nationaux et européens, reflétant au niveau financier les contrats mentionnés par ailleurs. Il est indispensable pour le laboratoire de continuer à avoir un financement récurrent conséquent afin de prévenir les fluctuations fortes engendrées par les contrats financièrement importants mais non lissés sur les différentes années. Ce dernier point explique les reports constatés d'une année sur l'autre, ceux-ci étant beaucoup plus importants en ce qui concerne les contrats que pour les crédits récurrents. La Fig. 1.4 gauche donne enfin l'évolution depuis 2000 du budget total ainsi que des reports. On peut d'une part constater que ces reports évoluent moins vite que l'accroissement budgétaire. Il nous faudra toutefois veiller à être plus réactif dans la gestion financière des contrats. Enfin l'accroissement substantiel du budget, presque doublé en sept ans, a induit une charge

administrative lourde même si la gestion effectuée par le SAIC de l'université Paris 13 est efficace.

Les dépenses (400 Keuros par an, cf Fig. 1.4 droite), concernent principalement les missions des chercheurs (visites de travail, conférences), l'équipement du laboratoire, ainsi que les charges salariales. Celles-ci ont augmenté sensiblement par rapport aux années antérieures. Ce phénomène devrait s'amplifier rapidement avec un financement de post-doctorants, de vacataires, voire d'allocations de thèse, à travers les projets ANR ou accompagnant les pôles de compétitivité mis en place depuis 2006. Le pic de dépense salariale de l'année 2004 se justifie conjoncturellement par des contrats européens importants arrivant à échéance (ce qui explique aussi les moindres dépenses engagées sur 2005).

1.7 Situation administrative et technique

La gestion du laboratoire est effectuée par trois personnels administratifs (deux CNRS et un de l'université) et deux personnels techniques (un CNRS et un université). Un ingénieur a été octroyé au titre de la recherche et est partagé avec le laboratoire de mathématiques (LAGA). La situation est correcte mais reste anormale pour certains aspects : d'une part il nous faut constater que la gestion technique est encore faite en partie par des enseignants-chercheurs, sans doute techniquement compétents mais dont ce ne devrait pas être la tâche ordinaire ; d'autre part le laboratoire ayant fortement cru ces dernières années tant en termes de permanents qu'en termes de projets gérés, la situation administrative devient délicate.

1.7.1 Locaux

Le LIPN occupe actuellement des couloirs de l'Institut Galilée (1432 m^2 SHON). Toutefois, est inscrit au contrat de plan Etat-Région 2007-2012 la construction pour 11 Meuros d'un bâtiment commun pour le LAGA et le LIPN. Cette construction nous permettra de déserrer la contrainte surface actuelle et devrait aussi aboutir à une mutualisation des moyens entre les deux laboratoires. Le LIPN occupe actuellement 3,5 couloirs de l'Institut Galilée : les premier et second étages du bâtiment A, le troisième étage du bâtiment B et une partie du deuxième étage du bâtiment B (l'autre partie étant occupée par la reprographie de l'institut). Ces locaux se répartissent en plusieurs catégories :

- 40 bureaux de taille variable, accueillant 1 à 4 personnes chacun ;
- une salle de repos (courrier, machine à café, ...) ;
- une salle de séminaires et réunions ;
- une salle serveurs climatisée ;
- une salle dédiée à l'archivage et à l'entrepôt des matériels ;
- 3 salles en libre service comportant imprimantes et photocopieurs.

1.7.2 Environnement technique

Le système informatique du LIPN se décompose en deux parties principales, la partie machines utilisateurs et la partie réseau et serveurs, plus un cluster de machines.

- Le laboratoire dispose de 82 postes clients fixes sous Linux (debian), 16 Mac, 3 machines sous windows et 7 autres machines utilisant d'autres versions de Linux. Les clients sous debian sont installés avec une procédure automatique mise au point par l'équipe système. Ceci assure une homogénéité du parc. Les autres postes clients

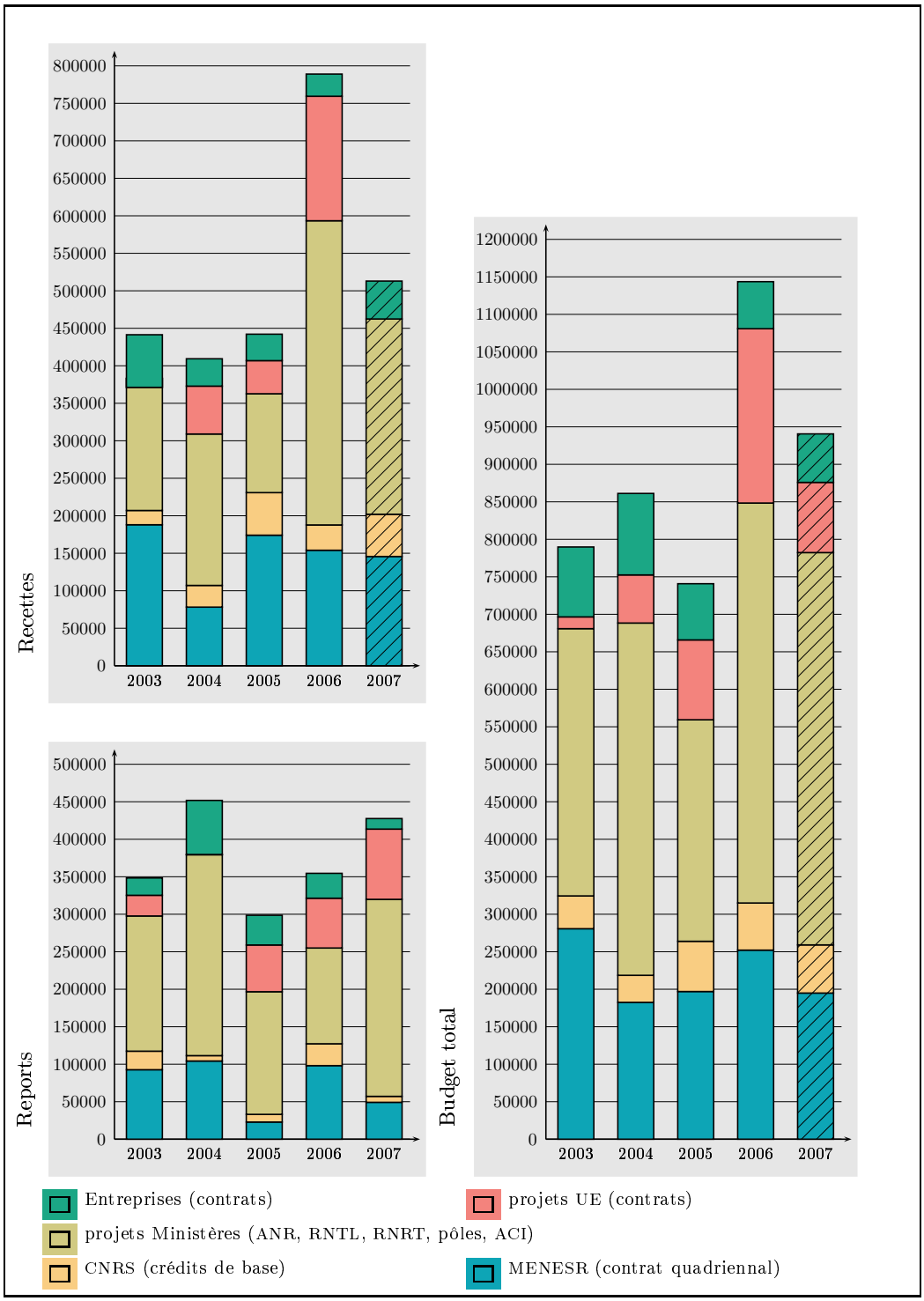


Fig. 1.2: Données financières du LIPN sur 2003-2006 et provisoire 2007 (euros)

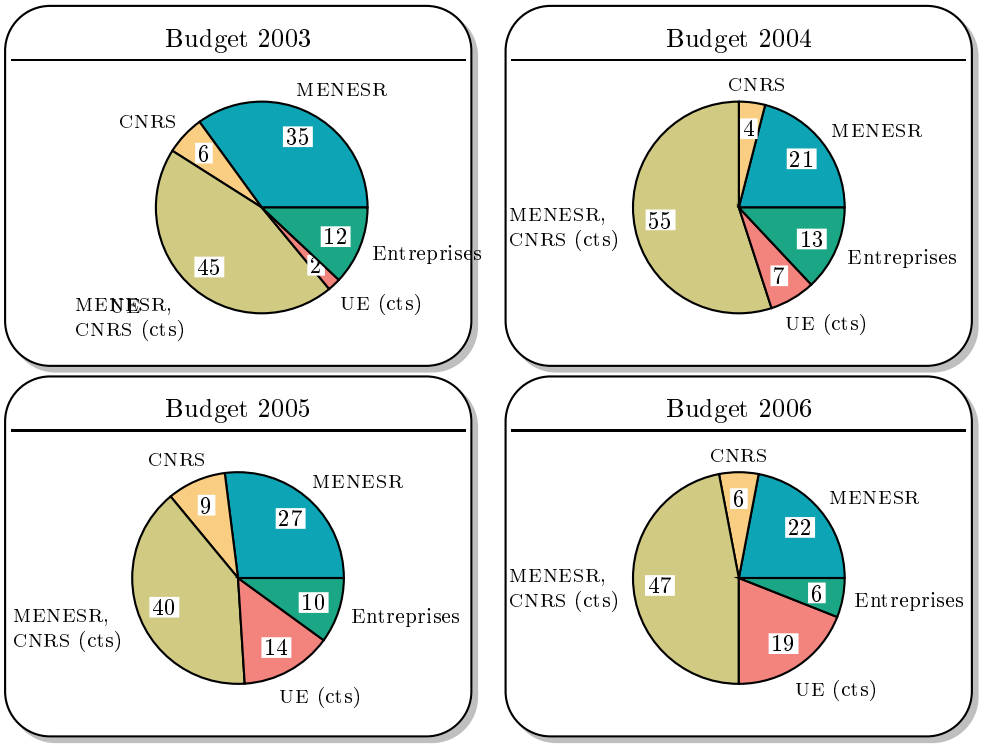


Fig. 1.3: Répartition des ressources budgétaires (cts : contrats)

et les portables sont administrés par leur utilisateur. Les postes dans les bureaux sont connectés via des câbles ethernet 100 Mb/s. Chaque bureau comporte 3 à 5 prises réseau. Le matériel est renouvelé régulièrement.

– Les différents serveurs sont installés dans un local climatisé. Ils sont reliés à des switchs connectés par rocade optique à 1 Gb/s. Dans l'état actuel, le laboratoire dispose de 4 serveurs de fichiers, un serveur de mail (dupliqué), un serveur web, un firewall et un serveur de sauvegarde. Les serveurs de fichiers seront prochainement remplacés par un serveur puissant redondant avec une baie de disques. Le plan du réseau est présenté dans la figure 1.5.
– Le *cluster* comporte actuellement 18 machines. Il est partagé par le LIPN et le LAGA. Ce cluster, physiquement situé dans les locaux du LAGA, a été installé, configuré et testé depuis plusieurs mois. Les membres des deux laboratoires peuvent y accéder depuis novembre 2006.

1.8 L'implication dans et hors de l'université

Le LIPN est bien intégré dans l'environnement universitaire de Paris 13, à l'Institut Galilée et à l'IUT de Villetaneuse. Les différentes formations en informatique ont ainsi été confortées dans le cadre du passage au système Licence-Master-Doctorat (licences professionnelles à l'IUT, licence-maîtrise-master, spécialité « Ingénierie Logicielle » de la formation d'ingénieurs à l'Institut Galilée), ce qui assure dans une certaine mesure un vivier d'étudiants à accueillir en stage ou en thèse. Concrètement, les membres du laboratoire prennent une part active dans la vie de l'université. Jacqueline Vauzeilles a été vice-présidente du Conseil Scientifique de l'Université de septembre 1999 à janvier

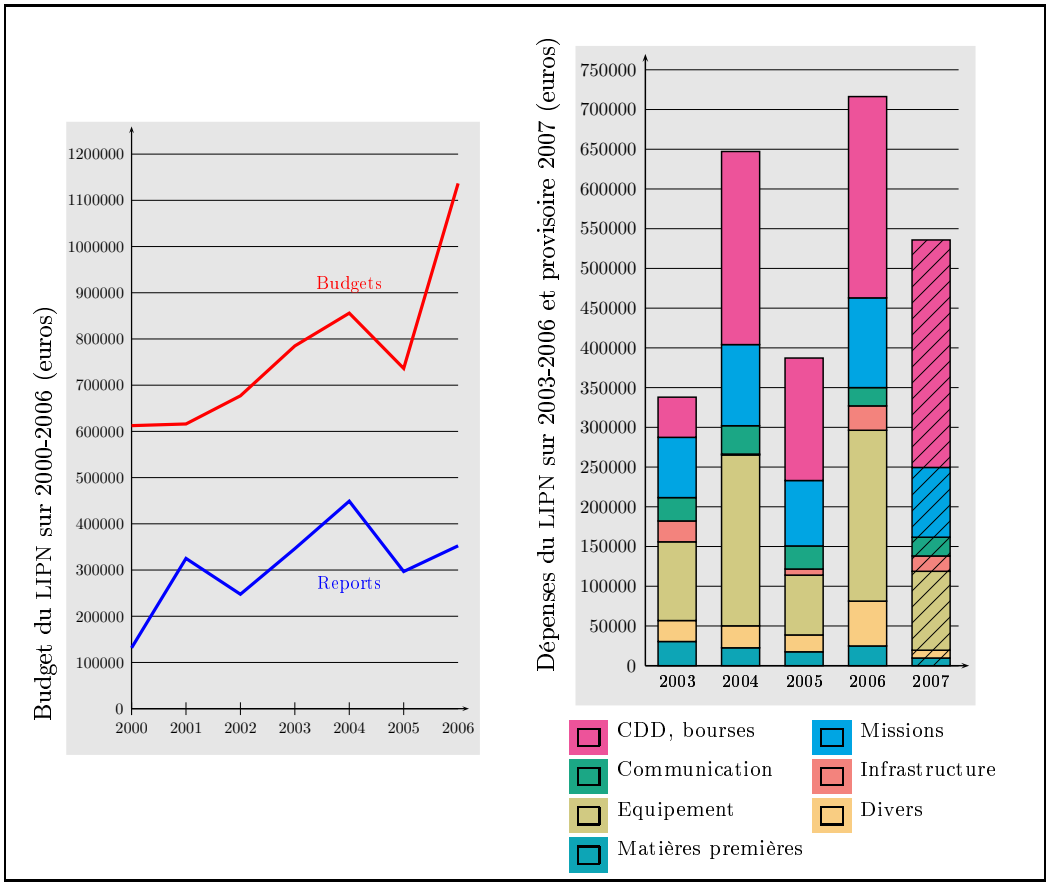


FIG. 1.4: Evolution budgétaire 2000-2006 et dépenses du LIPN sur 2003-2006 et provisoire 2007 (euros)

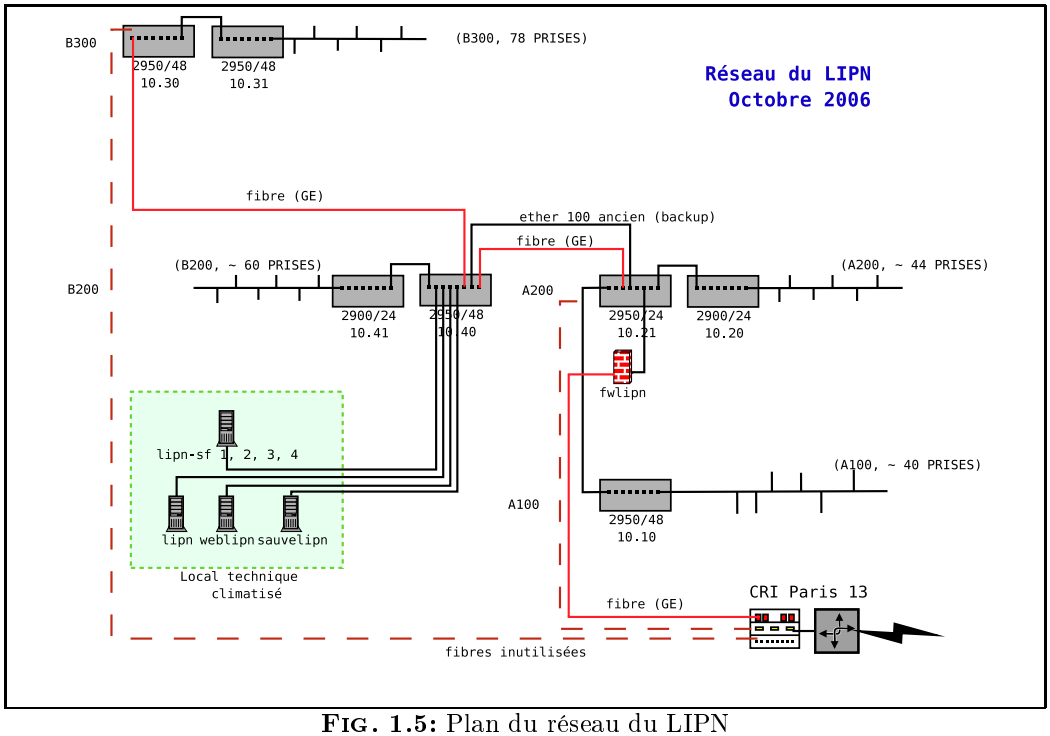


Fig. 1.5: Plan du réseau du LIPN

2003. Christophe Fouqueré et Françoise Gayral ont siégé au Conseil d'Administration de l'Université jusque fin 2004, Marc Champesme et Emmanuel Viennet depuis cette date ; Jacqueline Vauzeilles fait partie du Conseil Scientifique depuis 1998, Christophe Tollu en a été membre de 2002 à 2004, Christophe Fouqueré en fait partie depuis 2005 ; Pascal Coupey a fait partie du Conseil des Etudes et de la Vie Universitaire jusque fin 2004 et Christine Choppy l'a relayé au début de 2005. Les membres du laboratoire assurent également des responsabilités pédagogiques et scientifiques : Daniel Kayser a été président du département d'informatique de l'Institut Galilée jusqu'en décembre 2003 et Christine Choppy a pris sa succession, François Lévy est chef du département informatique de l'IUT de Villetaneuse. Younès Bennani est responsable du master informatique et Gérard Duchamp est co-responsable du master mathématique-informatique. Christophe Fouqueré a été directeur de la recherche de l'Institut Galilée de 2003 à 2004 et Jacqueline Vauzeilles est maintenant élue à cette fonction, Philippe Dague a été directeur des relations internationales et industrielles de l'Institut Galilée de 2004 à son départ. Enfin, Jacqueline Vauzeilles est membre du bureau de l'université depuis juin 2006, Christophe Fouqueré et Emmanuel Viennet sont chargés de mission informatique pour l'université depuis respectivement 2004 et 2006.

Notons par ailleurs que J. Vauzeilles et S. Salotti font ou ont fait partie du Conseil National des Universités, section 27. J. Vauzeilles est de plus Chargée de mission auprès de la MSTP (Nov. 2006-Juillet 2007) : responsable de l'évaluation des masters pour la DSPT9 - évaluation de projets internationaux, de laboratoires - participation aux comités de visite de l'IRCCyN (Nantes), du LIA (Avignon), du L3I (La Rochelle). Elle est membre des comités de visite des Ecoles Doctorales à Rouen et à Caen pour le département STIC de l'AERES (sept./oct. 2007). Elle est enfin coordinatrice scientifique du CSD1 (STIC) de l'USAR depuis juin 2007.

Chapitre 2
Organigramme

Services informatique et administratif du LIPN :

	Nom	*Prénom*	*Situation*	*Institution*
	Giraud	Jacqueline	Tech.	Univ Paris 13
	Guéveneux	Brigitte	ai	CNRS
	Wilk	Antonia	ajt	CNRS
	Lecocq	Michel	ie	Univ Paris 13
	Sow	Mamadou	ie	CNRS

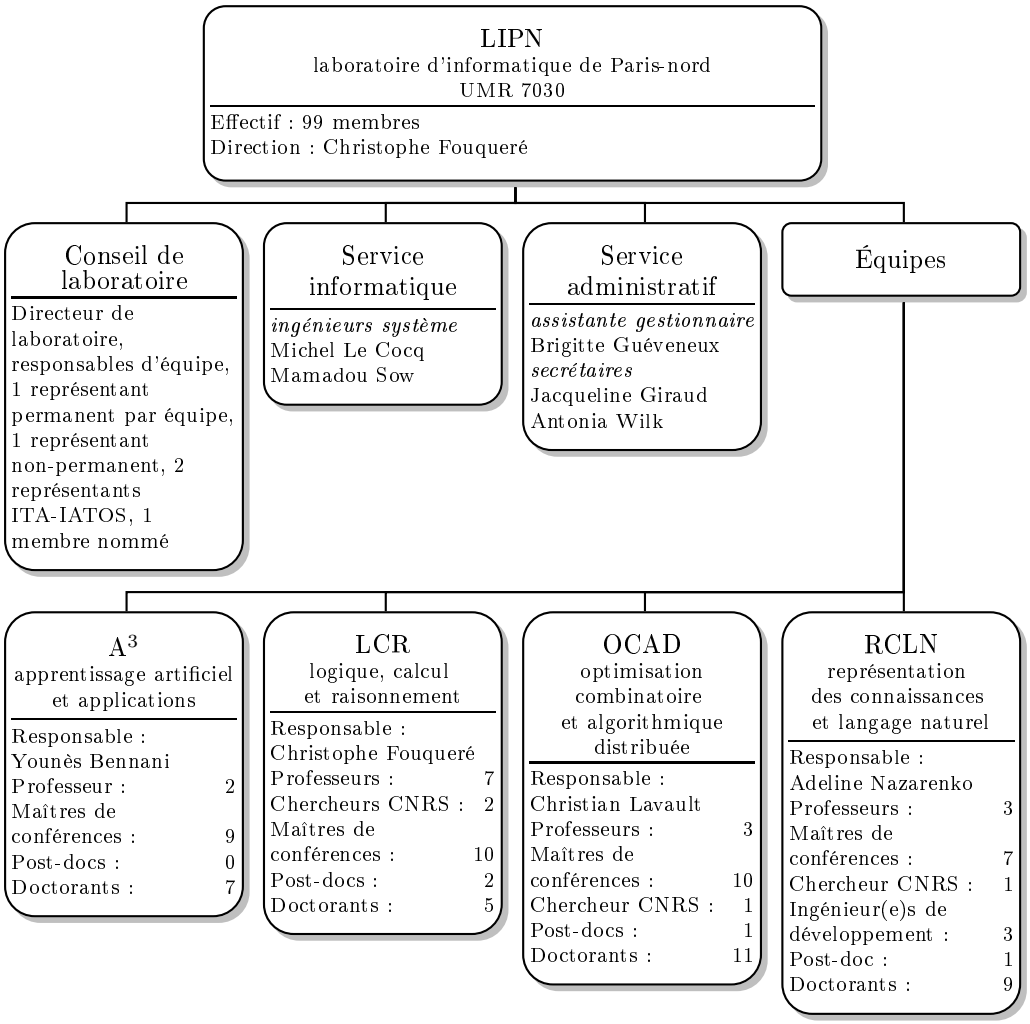


FIG. 2.1: Organigramme du laboratoire

Partie II
Activités de recherche

Chapitre 1
A^3 (Apprentissage Artificiel et Applications)

Sommaire

1.1 Membres de l'équipe 22
1.1.1 Evolution des membres de l'équipe 24
1.1.2 Historique 25
1.2 Fondements scientifiques et orientations de recherche . . . 26
1.3 Composante Apprentissage Numérique 28
1.3.1 Approches statistiques 28
1.3.2 Approches connexionnistes 29
1.3.3 Approches inspirées des comportements d'insectes sociaux . 33
1.3.4 Domaines d'applications 35
1.3.5 Résultats significatifs 35
1.3.6 Perspectives de recherche 36
1.4 Composante Apprentissage Symbolique 37
1.4.1 Les axes de recherche 37
1.4.2 Résultats significatifs 41
1.4.3 Domaines d'applications 41
1.4.4 Perspectives de recherche 41

1.5 Approches Transversales 43
1.5.1 Les axes de recherche 44
1.5.2 Domaines d'applications 45
1.5.3 Résultats significatifs 46
1.5.4 Perspectives de recherche 46
1.6 Animation de la recherche 47
1.6.1 Formation à la recherche 47
1.6.2 Valorisation de la recherche 51
1.6.3 Diffusion de la recherche 62
1.6.4 Evaluation de la recherche 63
1.6.5 Collaborations 65
1.7 Publications principales 65

1.1 Membres de l'équipe

Permanents

	Nom	*Prénom*	*Situation*	*Institution*
	Alphonse	Erick	MCF	IG
	Azzag	Hanane	MCF	IUTV
	Bennani	Younès	PU	IG
	Bouthinon	Dominique	MCF	IUTV
	Champesme	Marc	MCF	IG
	Gérard	Pierre	MCF	IUTV
	Guérif	Sébastien	MCF	IG
	Kanawati	Rushed	MCF	IUTV
	Osmani	Aomar	MCF	IUTV
	Rouveirol	Céline	PU	IG
	Soldano	Henry	MCF	IG
	Viennet	Emmanuel	MCF	IUTV

Doctorants

Nom	*Prénom*	*Situation*	*Institution*

	Aseervatham	Sujeevan	ATER	IG
	Cabanes	Guénaël	AM	
	Dehni	Lahcène	ATER	IG
	Elati	Mohamed	ATER	IG
	Grozavu	Nistor	CDD	
	Karoui	Hager	ATER	IUTV
	Rogovschi	Nicoleta	CDD	

A³

Membres associés

Nom	*Prénom*	*Situation*	*Institution*
Benabdeslem	Khalid	MCF	Université Lyon 1
Bossaert	Fabrice	MCF	IUT Bobigny
Krief	Francine	PU	Université Bordeaux 1
Lebbah	Mustapha	MCF	IUT Villetaneuse
Zucker	Jean-Daniel	DR	IRD Bobigny

personnels invités du 22/06/2005 au 1/9/2006

Prénom	*Nom*	*Origine*	*Financement*
2005			
Jose	Neuman de Souza	Federal University of Ceara (Brazil)	professseur invité
2006			
Jose	Neuman de Souza	Federal University of Ceara (Brazil)	professseur invité
2007			
Stefan	Kramer	TU Munich (Allemagne)	professseur invité
2008			
Attilio	Giordana	Université de Turin (Italie)	professseur invité

professeur invité : 1 mois d'invitation par l'université Paris 13

1.1.1 Evolution des membres de l'équipe

Départs

- A. El Fallah, MCF (thème SMA), promotion PU à Paris 10 (2002)
- A. Guerra-Hernandez, MCF, Univ. de La Cruz (Mexique), (Thèse soutenue en décembre 2003)
- K. Benabdeslem, MCF, Univ. Lyon 1, (Thèse soutenue en décembre 2003)
- F. Zehraoui, MCF, Univ. Evry, (Thèse soutenue en octobre 2004)
- H. Kadri-Dahmani, Post-doc, (Thèse soutenue en novembre 2005)
- F. Krief, MCF (thème diagnostic), promotion PU à Bordeaux 1 (2004)
- Ph. Dague, PU (thème diagnostic), mutation à Paris 11 (2005)
- A. Pauchet, (thème SMA), Post-doc chez France Télécom (2006)
- Z. Jrad, (thème SMA), Post-doc à l'INRIA (2006)
- H. Moungla, (thème SMA), Post-doc chez France Télécom (2006)

Arrivées

- P. Gérard, MCF IUTV (Apprentissage par renforcement), 2003 (thèse Paris 6)
- H. Karoui, doctorant, ATER (RaPC), octobre 2003
- S. Aseervatham, doctorant, AM (Apprentissage statistique), octobre 2004
- L. Dehni, doctorant, CDD ANR (Apprentissage connexionniste), octobre 2004
- E. Alphonse, MCF IG (Apprentissage symbolique), 2005 (thèse Paris 11)
- H. Azzag, MCF IUTV (Apprentissage bio-inspiré), 2006 (thèse Tours)
- C. Rouveirol, PU IG (Apprentissage symbolique), 2006 (auparavant MCF Paris 11, LRI Orsay)
- M. Elati, doctorant, AM (Apprentissage symbolique), octobre 2006
- N. Rogovschi, doctorante, CDD ANR (Apprentissage et réseaux sociaux), octobre 2006
- N. Grozavu, doctorant, CDD ANR (Apprentissage connexionniste), octobre 2006

1.1.2 Historique

L'équipe A³ (Apprentissage Artificiel & Applications) est issue d'une réorganisation des thèmes de recherche de l'équipe Adage (Apprentissage, Diagnostic et Agents) qui est devenue nécessaire après les départs respectifs d'Amal El Fallah (thème Agents) à Paris 10 et de Philippe Dague (thème Diagnostic) à Paris 11. Cette restructuration a été effectuée avec la volonté de tirer parti de la complémentarité de ses thèmes de recherche et de mettre en œuvre leur coopération afin de rendre plus visible cet axe de recherche au niveau national et international. En effet, le rayonnement du thème «Apprentissage Artificiel» du LIPN, se heurtait au manque d'une structure fédératrice et dynamique. La réorganisation des activités de recherche sur l'apprentissage artificiel à Paris-Nord (LIPN et ses proches voisins) selon une structure d'équipe de recherche fédératrice est un moyen prometteur pour renforcer la visibilité scientifique de ce thème de recherche. Cette structure a pour vocation de s'appuyer sur la qualité scientifique et la dynamique des chercheurs sur ce thème, les coopérations scientifiques sous forme de plusieurs axes de recherche inter-équipes du LIPN, des axes de recherche inter-laboratoires présents ou en cours d'élaboration. En effet, ces collaborations donnent naissance à une structure scientifique fédératrice des activités de recherche autour de l'apprentissage artificiel à Paris-Nord ayant un noyau au LIPN. Il s'agit d'un des premiers pôles sur cette thématique en Ile de France. Il faut souligner aussi que l'axe « Apprentissage Artificiel» est un des deux axes de recherche du LIPN présents dans le pôle de compétitivité Cap Digital-IMVN (Image, Multimédia et Vie Numérique).

Parallèlement à la prolongation des recherches au sein de chaque thème de l'équipe A³, plusieurs axes de recherche inter-équipes sont présents. En effet, la problématique de l'apprentissage est déjà largement partagée avec l'équipe RCLN, en particulier dans le cadre du projet InfoMagic (Pôle de compétivité Cap Digital), tout comme la classification hybride combinant des réseaux connexionnistes à un système de raisonnement à partir de cas. La résolution de problèmes de satisfaction de contraintes, utile notamment en raisonnement spatio-temporel, donne lieu à une coopération avec l'équipe OCAD sur le problème MAX-CSP. Enfin, des chercheurs en apprentissage symbolique de l'équipe participent à l'EPML IAPuces dirigée par Jean-Daniel Zucker, du Lim&Bio Bobigny de Paris-Nord. Soulignons enfin qu'une caractéristique essentielle du mode de fonctionnement de cette équipe est que les recherches y sont alimentées et évaluées grâce à de nombreuses applications, concernant la génomique, la parole, l'image, les télécommunications, Internet, le e-commerce et le e-learning, etc, donnant lieu pour la plupart à des contrats industriels.

1.2 Fondements scientifiques et orientations de recherche

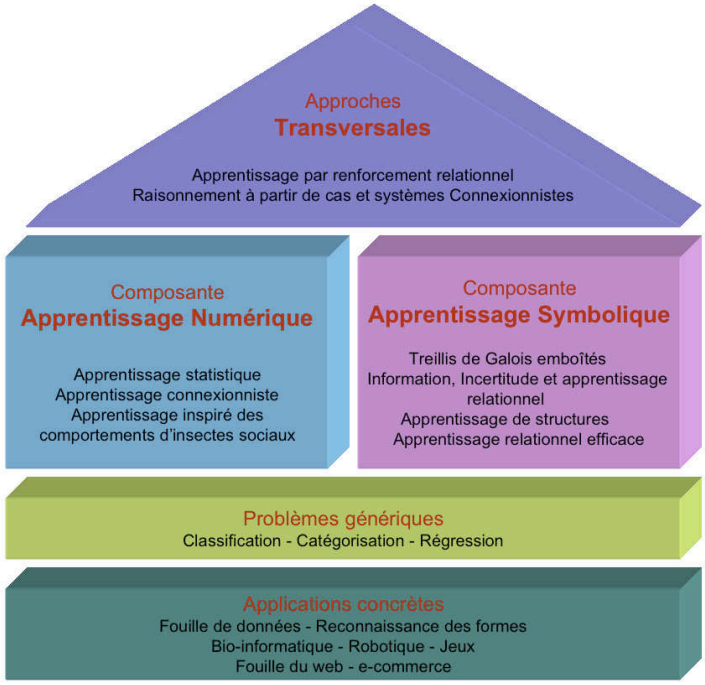
L'apprentissage artificiel est une discipline scientifique qui recouvre plusieurs aspects d'études mathématiques, statistiques et algorithmiques. Ses applications sont très stratégiques, notamment pour le traitement de données scientifiques, pour l'ingénierie, et pour les systèmes d'information. Les techniques d'apprentissage artificiel permettent de capturer des relations entre des événements ou des signaux caractéristiques d'un phénomène connu par des exemples (observations). Ces méthodes sont utilisées le plus souvent pour des problèmes pour lesquels on est capable de réaliser un grand nombre de mesures et dont on ne connaît pas la loi sous-jacente. Cela couvre une large gamme de problèmes pouvant aller de la modélisation de phénomènes naturels à l'identification et la commande de processus industriels en passant par la reconnaissance des formes et l'exploration des données. Les recherches dans ce domaine sont extrêmement actives depuis les années 80. Des laboratoires tant industriels qu'universitaires participent à cette recherche. Différentes voies ont déjà été explorées et les systèmes d'apprentissage sont devenus des outils incorporés dans différents logiciels commercialisés pour les statistiques, la fouille de données (Data Mining), le traitement du signal et l'identification de systèmes.

A³

Les activités de recherche au sein de l'équipe A³ sont structurées autour de deux grandes composantes :

- Composante Apprentissage numérique,
- Composante Apprentissage symbolique,

Afin de tirer parti de la complémentarité de ses différents thèmes de recherche et de mettre en oeuvre leur coopération des approches transversales (symbolico-numériques) sont aussi explorées. Ces recherches sont alimentées, coordonnées et évaluées grâce à diverses applications dans le domaine de la reconnaissance des formes et de la fouille de données.



Orientations scientifiques de l'équipe A³

L'équipe A³ développe des recherches fondamentales et intensifie une politique de coopération avec les grands organismes et les industriels, politique qui se traduit par un accroissement de sa participation dans les actions ANR, RNRT et RNTL, les contrats avec les différents ministères, et une collaboration avec des entreprises dans le cadre soit de projets européens, soit de bourses CIFRE.

1.3 Composante Apprentissage Numérique

Participants : S. Aseervatham, H. Azzag, Y. Bennani, L. Dehni, N. Grozavu, S. Guérif, N. Rogovschi, E. Viennet

L'apprentissage numérique consiste à découvrir un concept sous une forme géométrique simple (hyperplan). Il se présente sous la forme d'optimisation d'un critère (fonction de coût) interprétable sur l'ensemble d'apprentissage. Nos activités de recherche pour ces méthodes sont centrées sur les apports des techniques d'apprentissage numériques (statistiques et bio-inspirées) à l'analyse, l'exploration et l'interprétation de données complexes et de taille importante. Notre travail de recherche s'oriente plus particulièrement vers l'extraction de connaissances à partir de données et l'aide à la décision. Nos objectifs et activités de recherche s'articulent autour de deux dimensions complémentaires : théorique et pratique. Les études théoriques portent sur la conception de nouveaux algorithmes d'apprentissage et le développement d'approches modulaires et hybrides faisant coopérer différentes techniques. La dimension pratique s'attache à valoriser les activités de recherche en intégrant les résultats dans des systèmes opérationnels. Cette intégration se fait dans le cadre d'applications réelles dans différents domaines : parole, image, biométrie, diagnostic, Data Mining, Web Mining, Marketing. Notre problème central est donc celui de l'analyse et l'exploration de données afin d'en comprendre le sens, de déceler des relations entre des événements, d'en déduire des modèles de comportement. Il s'agit de la transformation de données en connaissances qui est et sera de plus en plus touchée par la nécessité de développer, de maîtriser et de choisir de nouveaux outils d'analyse et d'exploration des données. Le principal objectif de ces outils est celui d'un processus automatisé qui prend ses sources dans les données et dont le but est l'extraction de connaissances pour l'aide à la décision.

A³

Actuellement, nous développons les axes de l'apprentissage non-supervisé, l'apprentissage semi-supervisé et la visualisation de données structurées en séquences (temporelle, spatio-temporelles et évolutives).

Nous poursuivons aussi nos travaux sur la combinaison de modèles et la fusion de données et de décisions pour des problèmes de classement, de classification et de régression. Nous continuons également à travailler sur les techniques de sélection, d'extraction et de pondération de caractéristiques en apprentissage non-supervisé. Il nous paraît important de poursuivre l'étude en la généralisant à la modélisation de données dans de grandes bases de données issues des entrepôts de données. Nous envisageons de développer des activités sur l'exploration de données textuelles. Les objectifs visés concernent l'accès à l'information textuelle (recherche d'information, navigation dans les documents, filtrage et routage de documents), la gestion de données textuelles (structuration de documents et indexation, classification de textes), et l'analyse de contenu (extraction d'information, acquisition de terminologie spécialisée).

Ces recherches sont validées sur des problèmes de reconnaissance des formes et de fouille de données, en particulier sur des données temporelles, spatio-temporelles et évolutives (traces de navigations sur Internet, données comportementales, données de consommation, données de marketing, réseaux de capteurs, ...).

1.3.1 Approches statistiques

Apprentissage à base de noyaux pour le traitement de données textuelles

Cet axe est nouveau dans l'équipe, et concrétise la volonté affichée depuis plusieurs années par le laboratoire de développer des axes de recherche inter-équipes. Nous travaillons sur ce thème en relation avec plusieurs chercheurs de l'équipe RCLN, en

particulier dans le cadre du projet InfoM@gic et du projet Alvis. Une thèse est en cours sur les méthodes à noyaux appliquées au traitement de données structurées et semi-structurées.

Apprentissage semi-supervisé

Dans de nombreuses applications des techniques d'apprentissage, on dispose de grands volumes de données non étiquetées, mais seulement d'un petit échantillon de données étiquetées. Les algorithmes d'apprentissage séparent traditionnellement le traitement des données étiquetées (apprentissage *supervisé*) de celui des données non étiquetées (*non supervisé*). L'équipe A³ travaille depuis plusieurs années sur l'apprentissage non supervisé, en particulier via les cartes topologiques de Kohonen, et sur les classifieurs supervisés, en particulier les machines à vecteurs de support (SVM). Nous étudions actuellement comment aborder le cas semi-supervisé : modèles transductifs proposés par Vapnik, extraction de traits non supervisés, techniques de co-apprentissage. Nous espérons déboucher sur des approches applicables à des problèmes aussi différents que la fouille de textes ou la correction du biais lors de l'évaluation du risque crédit (traitement des refusés par les établissements bancaires).

A³

Réseaux Sociaux et apprentissage

L'étude des réseaux sociaux est une discipline récente. Longtemps confiné aux mathématiciens (théorie des graphes aléatoires) et aux sociologues, ce domaine a connu un net regain d'intérêt à la fin des années 90, lorsqu'on a remarqué que la plupart des réseaux observés dans la nature ne se comportaient pas comme des graphes aléatoires simples, mais partageaient un certain nombre de propriétés fondamentales. Parmi celles-ci, citons l'invariance de la structure par changement d'échelle, qui se traduit par une distribution des degrés suivant une loi de puissance. L'étude de ces réseaux a bien sûr été fortement stimulée par l'essor du Web, qui offre de nombreuses occasions de créer des réseaux (entre pages Web, entre utilisateurs...) dont l'analyse est fort utile (recherche d information, animation de sites Web, systèmes de recommandations...) mais pose des problèmes techniques mal résolus.

Dans la cadre du projet ANR RNTL Septia (*Solution Automatique pour le Traitement des Interactions Clients*), nous travaillons sur l'apport des techniques d'analyse des réseaux sociaux pour l'apprentissage, dans le contexte de l'élaboration de systèmes de classification de messages et d'aide à la réponse (pour la gestion de centres d'appel). Nous essayons par exemple de comprendre comment définir la notion de communauté d'utilisateurs afin de maximiser les performances d'un système de classification d'utilisateurs (ou des messages émis par ces derniers). Cette problématique a des liens évidents avec les méthodes de clustering non supervisé étudiées depuis plusieurs années dans l'équipe A³, et donne lieu au travail de doctorat de Mlle N. Rogovschi.

1.3.2 Approches connexionnistes

Les réseaux de neurones artificiels (ou réseaux connexionnistes) sont des systèmes numériques permettant la modélisation de processus généraux par l'établissement de modèles fonctionnels [LI-7]. Ceux-ci sont identifiés à partir des entrées-sorties du processus par des algorithmes dits "d'apprentissage" qui s'apparentent à des techniques d'estimation statistiques. Nés en informatique dans le domaine de l'intelligence artificielle, ils ont connu depuis le début des années 80 un développement intensif dû au succès rencontré dans une très large gamme d'applications. Les réseaux connexionnistes offrent

une panoplie de techniques adaptatives pour de nombreux problèmes génériques : la classification, le classement, la modélisation, la prévision. Parmi les domaines où ils ont été utilisés, on peut citer le traitement du signal, de l'image, de la parole, le contrôle, la prévision, l'optimisation combinatoire, le diagnostic, la fouille de données, ... Il s'agit donc d'un outil général pouvant être utilisé dans des cadres différents avec une grande souplesse. Les modèles connexionnistes apparaissent comme une alternative ou plus souvent comme un complément aux techniques classiques. Les recherches dans le domaine sont extrêmement actives actuellement. Plusieurs laboratoires tant industriels qu'universitaires participent à cette recherche en Europe, aux États-Unis et au Japon. Différentes voies ont déjà été explorées et les réseaux connexionnistes sont devenus un outil incorporé dans différents logiciels commercialisés pour les statistiques, la fouille de données, le traitement du signal et l'identification de systèmes.

Classification connexionniste non supervisée et visualisation de données évolutives

A^3

Par opposition aux données temporelles, le recueil des données évolutives ne se fait pas avec une périodicité de la mesure : des mesures peuvent avoir été recueillies à différents moments. Cela prive les données de leur caractère temporel au sens statistique. L'évolution de ces données peut être vue sous forme de trajectoires temporelles (de taille variable et parfois multivariées). La visualisation de ces trajectoires dans l'espace de l'évolution devient difficile à interpréter dès que le nombre de trajectoires est élevé. L'un des objectifs principaux de l'analyse de ce type de données est d'obtenir une classification (ou typologie) évolutive et incrémentale de trajectoires en fonction de leurs positions et leurs évolutions. Cette typologie permettra de diminuer le nombre de ces trajectoires, en assimilant les trajectoires de la même typologie à la trajectoire prototype de la typologie.

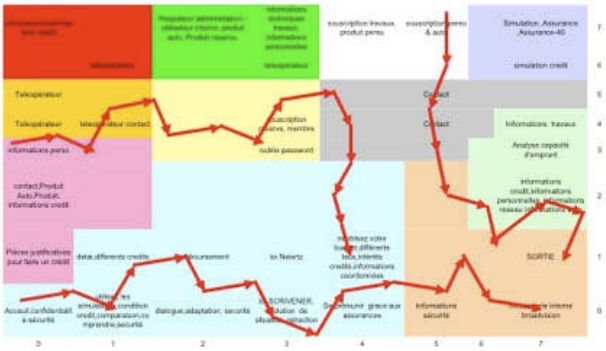


FIG. 1.1: Cartographie comportementale d'un site web et visualisation des trajectoires des Internautes

Nous avons proposé plusieurs approches d'apprentissage non supervisé pour l'exploration de ce type de données [CI-27, RI-24, CI-3]. La validation de ces approches a été effectuée en particulier sur les traces de navigation dans les sites web. La figure 1.1 montre un exemple de cartographie comportementale d'un site web avec la visualisation des navigations sous forme de trajectoires.

Nous travaillons sur d'autres approches à base d'algorithme d'apprentissage non-supervisé pour traiter ce type de données en quatre phases :

– *modélisation globale : recherche d'un modèle unique de représentation qui résume d'une manière globale l'ensemble des données (compromis).*

- *étude de l'interstructure : identification des proximités ou des différences entre les ensembles de données, sans description.*
- *étude de l'intrastructure : analyse fine des ressemblances ou des différences permettant d'expliquer les structures identifiées.*
- *étude de l'évolution des trajectoires : analyse et description de l'évolution du phénomène d'une manière globale ou étude des évolutions de chacun des éléments.*

Réduction des dimensions de l'espace de description

La taille des données peut être mesurée selon deux dimensions, le nombre de variables et le nombre d'exemples. Ces deux dimensions peuvent prendre des valeurs très élevées, ce qui peut poser un problème lors de l'exploration et l'analyse de ces données. Pour cela, il est fondamental de mettre en place des outils de traitement de données permettant une meilleure compréhension de la valeur des connaissances disponibles dans ces données. La réduction des dimensions est l'une des plus vieilles approches permettant d'apporter des éléments de réponse à ce problème. Son objectif est de sélectionner ou d'extraire un sous-ensemble optimal de caractéristiques pertinentes pour un critère fixé auparavant. La sélection de ce sous-ensemble de caractéristiques permet d'éliminer les informations non-pertinentes et redondantes selon le critère utilisé. Cette sélection/extraction permet donc de réduire la dimension de l'espace des exemples et rendre l'ensemble des données plus représentatif du problème. En effet, les principaux objectifs de la réduction de la dimension sont :

- *faciliter la visualisation et la compréhension des données,*
- *réduire l'espace de stockage nécessaire,*
- *réduire le temps d'apprentissage et d'utilisation,*
- *identifier les facteurs pertinents.*

Les algorithmes d'apprentissage artificiel requièrent typiquement peu de traits (features) ou de variables (attributs) très significatifs caractérisant le processus étudié. Dans le domaine de la reconnaissance des formes et de la fouille de données, il pourrait encore être bénéfique d'incorporer un module de réduction de la dimension dans le système global avec comme objectif d'enlever toute information inconséquente et redondante. Cela a un effet important sur la performance du système. En effet le nombre de caractéristiques utilisées est directement lié à l'erreur finale. L'importance de chaque caractéristique dépend de la taille de la base d'apprentissage (pour un échantillon de petite taille, l'élimination d'une caractéristique importante peut diminuer l'erreur). Il faut aussi noter que des caractéristiques individuellement peu pertinentes peuvent être très informatives si on les utilise conjointement.

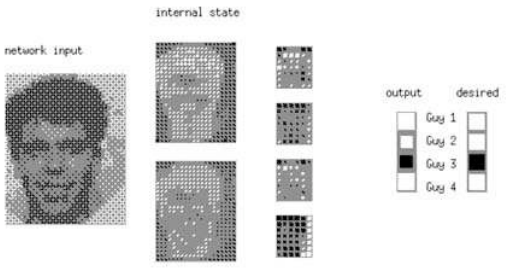


FIG. 1.2: Architecture d'un système connexionniste pour la reconnaissance faciale avant l'optimisation

La réduction de la dimension est un problème complexe qui permet de réduire le volume d'informations à traiter et faciliter le processus de l'apprentissage. Nous

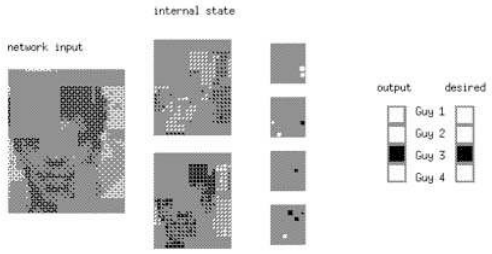


FIG. 1.3: Architecture d'un système connexionniste pour la reconnaissance faciale après optimisation par la mesure de pertinence HVS proposée par Y. Bennani et M. Yacoub en 2001. Les zones grises représentent les neurones supprimés.

A³

pouvons classer toutes les techniques mathématiques de réduction des dimensions en deux grandes catégories :

- *la sélection de variables : qui consiste à choisir des caractéristiques dans l'espace de mesure,*
- *l'extraction de traits : qui vise à sélectionner des caractéristiques dans un espace transformé (dans un espace de projection)*

La première catégorie est appropriée quand l'acquisition de mesures des formes est coûteuse. Ainsi l'objectif principal de la sélection de caractéristiques dans ce cas est de réduire le nombre de mesures requises. Par contre, les techniques d'extraction de traits (deuxième catégorie) utilisent toute l'information contenue dans les formes pour la compresser et produire un vecteur de plus petite dimension. Ces techniques projettent un vecteur forme de l'espace de représentation dans un espace de dimension plus petite. Les systèmes d'apprentissage connexionniste sont un bon exemple de cette catégorie. En effet, les modèles connexionnistes connus pour une tâche de discrimination fournissent un système avec des aptitudes intéressantes pour l'analyse du processus. Les cellules cachées d'un Perceptron multi-couches apprennent comment extraire les caractéristiques significatives du signal d'entrée. Nous avons proposé dans cette direction plusieurs mesures de pertinence, en apprentissage supervisé et non supervisé [CO-86, CI-110].

Combinaison de modèles et fusion de décisions

Un de nos thèmes de recherche majeur est la conception et la mise en œuvre de systèmes qui font coopérer différents modèles entre eux et avec des techniques plus classiques dans les différents domaines d'application. Ces systèmes constituent une autre génération des modèles connexionnistes et apparaissent comme la seule solution pour la réalisation de tâches complexes et de grande taille. Les études théoriques portent sur les algorithmes d'apprentissage pour ce type de système, la décomposition de tâches, la conception de chaînes de traitement optimales. Nous travaillons sur la possibilité de coopération des algorithmes connexionnistes et d'autres méthodes plus classiques comme les Chaînes de Markov Cachées, afin de concevoir des systèmes modulaires hybrides plus robustes.

Classification topographique évolutive et stabilité

L'apprentissage non-supervisé évolutif est un problème assez récent et représente une voie de recherche très prometteuse ouvrant d'excellentes perspectives. En effet, l'apprentissage évolutif fait partie des sujets difficiles en apprentissage artificiel, essentiellement parce qu'il n'existe pas de théorie de l'apprentissage statistique réellement dévolue dans le cas où les exemples sont fournis en ligne et les concepts

de classes varient dans le temps. Ces systèmes permettront de découvrir un espace topologique (évolutif) d'un ensemble de données. Cet espace topographique préserve la notion de voisinage entre les exemples décrivant le processus à modéliser. Nous avons étudié la possibilité de rendre évolutive la découverte de cet espace topographique : i.e. le nombre de regroupement homogène (clusters) change dans le temps en fonction des informations complémentaires qui arrivent, la classification se poursuit dans le temps. Pour cet axe de recherche, nous travaillons aussi sur les critères d'évaluation et de stabilité [CI-160]. En effet, les algorithmes de classification sont aujourd'hui très nombreux, mais on n'a presque jamais de raison « théorique » de choisir l'un plutôt que l'autre. La stabilité est une propriété très souhaitable pour ce type d'algorithmes Ces recherches sont validées en particulier sur des données temporelles, spatio-temporelles, évolutives (traces de navigations sur Internet, données comportementales, données RFID, données de consommation, données de marketing, réseaux de capteurs, ...).

1.3.3 Approches inspirées des comportements d'insectes sociaux

Notre travail dans cette direction est consacré à l'étude de solutions originales de classification non supervisée. Ces solutions s'inspirent des propriétés du vivant pour proposer des constructions originales, dotées de capacités étonnantes et aptes à résoudre des problèmes difficiles vis-à-vis desquels les approches plus traditionnelles rencontrent des difficultés. Le principe des méthodes que nous développons se base sur des études menées en éthologie et consiste ainsi à extraire de ces études un modèle réalisable du point de vue informatique.

Ce nouveau domaine de recherche a pour objet de transformer la connaissance que les éthologues ont des capacités collectives de résolution de problèmes des insectes sociaux en techniques artificielles de résolution de problèmes en informatique. Comment un ensemble d'entités individuellement simples au fonctionnement décentralisé est-il capable collectivement de fournir une réponse complexe? Ce principe de plusieurs entités simples en interaction qui émerge à un niveau collectif en une structure complexe (qualifiée d'intelligente) est présent dans différentes disciplines : biologie, physique, informatique, etc. Dans ces domaines, différents modèles ont vu le jour pour décrire et analyser de tels systèmes.

C'est dans ce vaste domaine de l'informatique biomimétique dans lequel nous tentons de contribuer en apportant de nouveaux modèles bio-inspirés pour résoudre des problèmes de classification non supervisée et hiérarchique sur des grandes masses de données. Les résultats obtenus par des simulations permettent ensuite d'évaluer la qualité du modèle relativement aux études biologiques. Les conséquences peuvent être utiles aux informaticiens afin d'améliorer le modèle lui même ou aux éthologues qui peuvent développer et tester leurs théories du comportement animal.

Simulation et visualisation 3D de la dynamique des blattes

Dans le cadre du projet *ANR Sillages* nous nous sommes intéressés à un processus d'auto-organisation qui est l'agrégation, et sa capacité à générer des organisations spatio-temporelles chez les espèces grégaires, telles que la blatte américaine *Periplaneta americana*. Cette dernière sera utilisée par les éthologues comme modèle expérimental.

Afin de faciliter le travail des éthologues, nous avons développé un outil qui simule le comportement des blattes *Periplaneta americana*. Notre but est que les chercheurs

puissent mieux cerner la dynamique de ces insectes et découvrent de nouvelles hypothèses les concernant.

Comme nous le montrerons dans la section suivante, ce simulateur nous a permis de découvrir que le comportement collectif des blattes peut construire de nouvelles méthodes de classification des données.

Notre simulateur modélise le dispositif expérimental présenté dans la figure 1.4.

Ce modèle utilise une chaîne de *Markov* qui simule le déplacement de chaque blatte suivant sa position courante : vers les abris ou vers l'environnement extérieur. Tous ces déplacements vont dépendre de probabilités liées à la qualité de l'abri, sa taille ou encore au nombre de blattes présentes sous l'abri à un instant t.

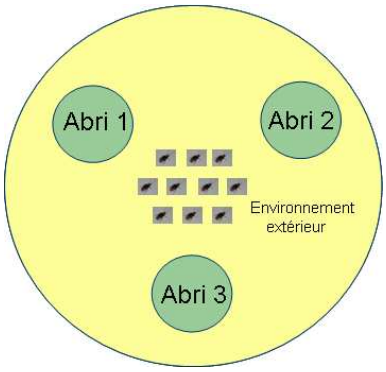


FIG. 1.4: Dispositif expérimental pour tester les blattes réel

Nous avons également greffer à notre simulateur un outil de visualisation 3D permettant de voir les déplacements de blattes de manière plus réaliste (voir figure 1.5).

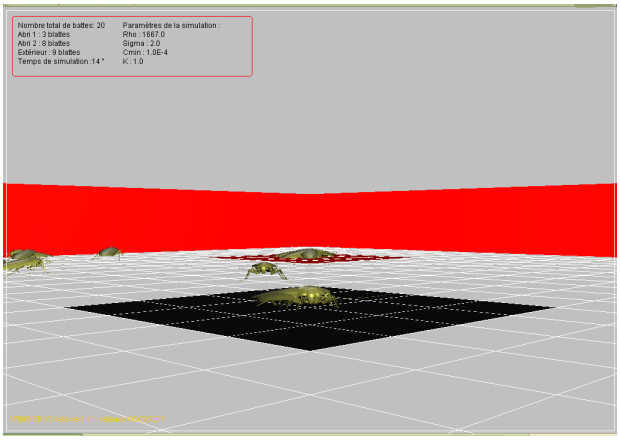


FIG. 1.5: Visualisation 3D de la dynamique du comportement des blattes

Classification non supervisée inspirée des comportements d'insectes sociaux

Cet axe a pour objectif de transformer la connaissance que les éthologues ont des capacités collectives de résolution de problèmes chez les blattes américaines, en techniques artificielles de Clustering. Nous avons ainsi imaginé un système dans lequel il s'agirait de modéliser la manière dont les blattes forment des agrégats

sous chaque abri, et d'utiliser ce comportement pour partitionner des données en différentes classes. Chaque blatte va ainsi représenter une donnée à classer et les abris les classes initiales.

Ces blattes (artificielles) vont ensuite construire de manière similaire un partitionnement en appliquant certaines règles comportementales. Les déplacements des blattes vers les différentes classes dépendent de la similarité entre les données.

Les résultats que nous avons obtenus sont très intéressants, nous avons pu les tester sur plusieurs types de données car notre approche ne dépend pas du type de données à classer : il suffit uniquement de définir une fonction de similarité pour ces nouvelles données.

Il reste beaucoup de problèmes à résoudre pour ce type d'approche et les résultats obtenus jusqu'ici sont plus qu'encourageants [CI-171]. Ce nouveau modèle se placera ainsi dans la lignée des travaux précédents sur les algorithmes de classification s'inspirant des insectes sociaux et plus généralement des systèmes biologiques.

Génération de graphes par des fourmis artificielles A^3

En collaboration avec le laboratoire d'informatique de l'Université de Tours, nous avons développé un algorithme incrémental nommé AntGraph pour la construction de graphes de voisinage. Il s'inspire du comportement d'auto-assemblage observé chez une population de fourmis réelles où ces dernières se fixent progressivement à un support fixe puis successivement aux fourmis déjà fixées afin de créer une structure vivante.

Nous utilisons ainsi une approche à base de fourmis artificielles où chaque fourmi représente une donnée. Ce comportement a été précédemment utilisé pour construire de manière incrémentale un graphe à partir d'une mesure de similarité entre les données. Il s'agit en effet d'une généralisation de l'algorithme AntTree proposé par Hanane Azzag dans sa thèse.

Par cette méthode, le graphe obtenu est plus représentatif de la similarité existante entre les données avec en plus une complexité inférieure à celle observée dans les méthodes classiques (Voisins Relatifs, graphes de Gabriel, triangulation de Delaunay). Les résultats obtenus sont meilleurs, notamment en terme de temps de calcul [CI-146].

1.3.4 Domaines d'applications

Nous travaillons sur différents domaines d'applications, toutefois nous avons développé ces dernières années un thème principal qui est l'analyse et la visualisation de données structurées en séquences :

- *Exploration des traces de navigations sur Internet ;*
- *Analyse et visualisation des données comportementales : Marketing, RFID ;*
- *Diagnostic des systèmes complexes ;*
- *Identification multi-modale : vocale et visuelle.*

1.3.5 Résultats significatifs

Parmi les nombreux résultas de nos activités de recherche, nous pouvons mettre en avant les cinq contributions suivantes :

- *Mesures de pertinence pour la sélection de caractéristiques et l'optimisation des architectures connexionnistes ;*

- *Algorithmes de pondération des variables, durant l'apprentissage, en fonction de leur pouvoir discriminant ;*
- *Classification incrémentale de données évolutives, séquences ;*
- *Algorithmes de classification topographique à deux niveaux simultanés ;*
- *Sélection/pondération de variables en apprentissage non-supervisé à deux niveaux.*

1.3.6 Perspectives de recherche

Nous continuons actuellement nos travaux sur la combinaison de modèles et la fusion de données et de décisions pour des problèmes de classement, de classification et de régression. Nous continuons également à travailler sur les techniques de sélection et d'extraction de caractéristiques dans les réseaux connexionnistes. Il nous paraît important de poursuivre l'étude en la généralisant à la modélisation de données dans de grandes bases de données issues du web ou des entrepôts de données. Nous envisageons aussi de développer des activités sur l'exploration de données textuelles. Les objectifs visés concernent l'accès à l'information textuelle (recherche d'information, navigation dans les documents, filtrage et routage de documents), la gestion de données textuelles (structuration de documents et indexation, classification de textes), et l'analyse de contenu (extraction d'information, acquisition de terminologie spécialisée). La problématique de la classification topographique évolutive sera aussi abordée dans le cadre de la classification à base de noyaux, pour laquelle nous définirons un critère d'évaluation prenant en compte son double objectif (évolution et topographie). Nous proposerons aussi dans un contexte d'apprentissage non supervisé, des noyaux pour traiter des données structurées (séquences, arbres, graphes, ...). Enfin, nous souhaitons apporter quelques contributions permettant de faire naître un cadre théorique pour l'apprentissage non-supervisé évolutif.

1.4 Composante Apprentissage Symbolique

Participants : E. Alphonse, D. Bouthinon, M. Champesme, M. Elati, A. Osmani, C. Rouveirol, H. Soldano.

Cet axe a pour but d'appréhender des quantités de données importantes (cf. fouille de données) décrites de manière symbolique, i.e., telle que les exemples et les hypothèses sont représentés par des formules logiques. Dans ce cas, les exemples ainsi que les hypothèses candidates pour le concept cible sont organisés dans des espaces fortement structurés. Les techniques d'apprentissage symbolique, par opposition aux techniques d'apprentissage numérique, vont explorer l'espace des hypothèses en tirant parti de la structure de l'espace de recherche imposée par la relation structurant l'espace, afin d'identifier une hypothèse satisfaisant des critères de couverture de l'ensemble d'apprentissage.

L'équipe couvre un large spectre de techniques d'apprentissage (apprentissage supervisé, apprentissage descriptif), et conçoit des systèmes d'apprentissage pour des langages de concept allant de la logique attribut-valeur à des restrictions de la logique des prédicats comme Datalog non récursif avec négation. Orthogonalement, des travaux sont menés sur un nouveau cadre d'apprentissage permettant de formaliser l'apprentissage en présence d'exemples incomplets ou incertains.

A³

Nous avons poursuivi des travaux dans le domaine de l'apprentissage attribut-valeur, sur le thème des treillis de Galois emboîtés et développé et étudié les performances d'un système d'apprentissage multi-agent. L'équipe est également active dans le domaine de l'apprentissage de structures (séquences, graphes) et en Programmation Logique Inductive, domaine dédié à l'Apprentissage Symbolique dans des langages de concepts restrictions de la logique du premier ordre. Ces langages de concept complexes sont requis dans le cadre d'applications en fouille de données issues du Web (link mining, graph mining), de développements en Bases de Données Inductives ou d'applications dans les domaines du Traitement Automatique de la Langue, de la Bio-chimie ou de la Bioinformatique.

Dans ces langages, les mécanismes associés à l'apprentissage sont coûteux (voire indécidables), la recherche heuristique joue donc un rôle primordial. Nous avons développé une nouvelle stratégie d'apprentissage relationnel efficace, fondées sur un changement de représentation, appelé propositionnalisation, qui projette un problème d'apprentissage en un problème attribut-valeur. Enfin, nous étudions les phénomènes de transition de phase liés à l'apprentissage relationnel et leurs effets sur l'efficacité de la recherche d'hypothèses, comme l'apparition d'importants phénomènes de plateaux au cours de la recherche d'une hypothèse candidate, rendant "aveugle" le système d'apprentissage.

1.4.1 Les axes de recherche

Treillis de Galois emboîtés - constructions de taxonomies

Notre travail utilise essentiellement des structures latticielles basées sur des correspondances de Galois définies entre les objets observés et leur description. L'inconvénient principal de ces structures est leur grande taille (en nombre de nœuds) qui en rend très difficile l'exploration directe par l'utilisateur, nous avons défini une notion de «projection» sur l'ensemble des objets que nous utilisons pour faire varier la granularité de cette structure en offrant à l'utilisateur la possibilité d'explorer les données selon un point de vue d'abord grossier (pour appréhender globalement l'ensemble des données) puis de plus en plus détaillé à mesure que l'intérêt de l'utilisateur se focalise.

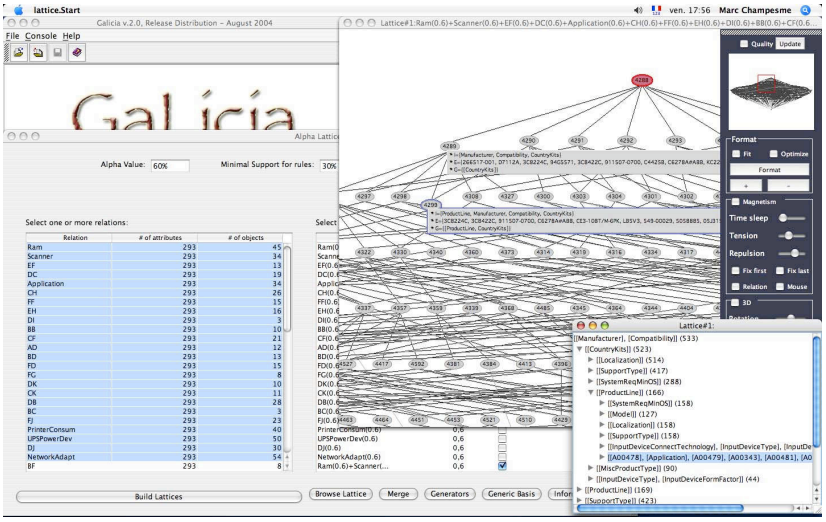


FIG. 1.6: La plate-forme Galicia étendue pour la prise en compte de la projection alpha (ici, treillis construit sur les données C/NET pour $\alpha = 60\%$)

La notion de projection "alpha" [RE-16] (cas particulier de projection que nous avons définie) permet en outre d'utiliser, quand elle existe, une partition des objets en classes, l'exploration des données étant alors guidée par cette partition. L'utilisation de telles partitions nous permet alors d'ignorer des propriétés marginales des objets, de manière analogue aux théories avec défauts exceptions, mais en restant dans le cadre théorique simple des correspondances de Galois "ordinaires" (i.e. domaine de l'analyse formelle de concept). Notre travail s'intéresse par ailleurs à l'extraction des règles induites par cette structuration.

Apprentissage dans un système multi-agents cognitif

Dans le cadre de la thèse d'A. Guerra Hernandez, nous nous sommes intéressés avec A. Elfallah-Segrouchni, à l'apprentissage symbolique multi-agents. Le SMA considéré est composé d'agents communicants de type BDI (Beliefs, Desires, Intentions). Au niveau collectif, un groupe d'agents BDI améliore son apprentissage de contextes de plans en échangeant des croyances avec d'autres agents [CO-49]. Nous nous intéressons actuellement à l'apprentissage incrémental collectif dans les Systèmes Multi-Agents dans le cas où les agents sont complètement autonomes.

Analyse de l'information mise en jeu au cours de l'apprentissage

Ce thème s'articule autour de 2 axes. D'une part, nous avons proposé une formalisation en logique du premier ordre de différents cadres d'apprentissage dont les données recèlent des formes variées d'incomplétude ou d'ambiguïté. En particulier, nous présentons un nouveau cadre d'apprentissage qui rend compte de situations jusqu'ici inexprimables du fait de la nature de l'incomplétude (incertitude sur les exemples positifs et négatifs + abstraction). Ce cadre permet de plus de représenter tout problème d'apprentissage multi-instances introduit par Dietterich ou d'apprentissage par satisfiabilité introduit par de Raedt.

D'autre part, nous menons une analyse de l'information mise en jeu au cours de l'apprentissage, qui vise à identifier et quantifier l'information fournie à l'apprenant. Le modèle d'information que nous proposons, fondé sur la théorie de l'information, permet de mesurer et comparer l'information apportée par différents types d'ex-

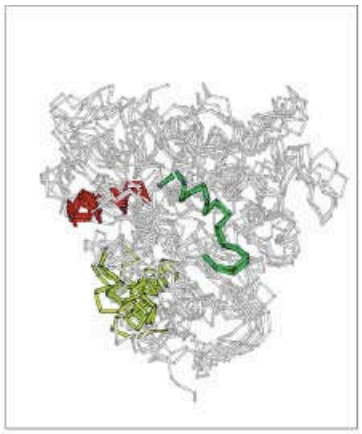
emples incomplets ou ambigus. En particulier, nous mettons en évidence un fort déséquilibre informationnel entre les exemples positifs et négatifs en apprentissage multi-instances. Un approfondissement de cet axe nous a permis d'établir des définitions et propriétés relatives à la corrélation existant entre éléments informationnels. Nous démontrons qu'un ensemble d'éléments informationnels apporte plus d'information que la somme des contributions de chacun d'eux. Enfin nous avons exhibé un critère informationnel permettant de guider des changements dynamiques de représentations au sein d'un espace de représentations possibles.

Apprentissage de structures

Nous nous intéressons à deux types de structures qui apparaissent dans un grand nombre d'applications, notamment en Bioinformatique : les séquences et les graphes. En ce qui concerne la recherche de structures séquentielles, nous avons travaillé sur la détection de *motifs relationnels répétés* dans les séquences, c'est-à-dire de motifs décrivant non seulement leurs composants mais aussi les relations qui les lient. L'algorithme KMRRelat et les conditions de son application à la recherche de motifs structuraux 3D sur les protéines ont fait l'objet de la thèse de N. El-Zant. Des recherches récentes nous ont conduits, avec N. Pisanti, en post-doc CNRS au LIPN, et M. Carpentier (thèse en sciences de la Vie à Paris VI, soutenue en décembre 2005) à une amélioration de cet algorithme permettant d'éviter la génération de (parfois très nombreux) candidats inutiles lors de la recherche de répétitions et aux représentations de ces motifs en vue de leur recherche dans les séquences de structures de protéines [CL-15]. Nous nous intéressons actuellement à la sélection et l'ordonnancement de ces motifs (de différentes longueurs) en vue d'obtenir des signatures structurales pour différentes familles de protéines.

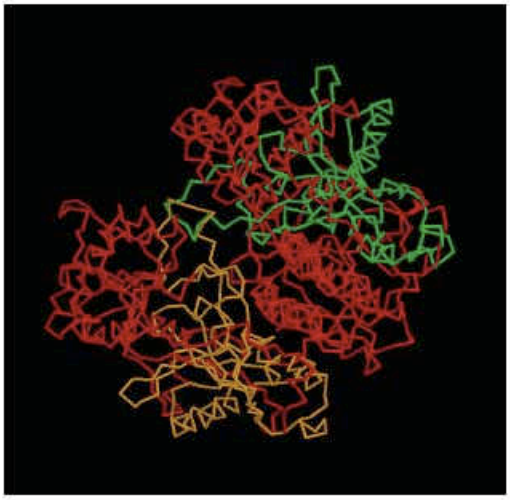
D'autre part, dans le cadre de la thèse de M. Elati, nous abordons l'apprentissage de structures de graphes sous-jacentes à des données propositionnelles. Ce travail est appliqué à l'extraction de réseaux de régulation à partir de données d'expression de gènes, en collaboration avec l'Institut Curie et l'Institut de Génétique Moléculaire de l'Université Paris 11. La détection de relations de régulation repose la plupart du temps sur l'hypothèse qu'un régulateur et son gène cible sont corrélés. Construire un graphe acyclique où chaque lien est une relation orientée de régulation et maximisant une mesure de corrélation donnée est un problème NP-complet, et a donné lieu aux développements de multiples approches heuristiques, dont bayésiennes. Cependant, ces méthodes sont peu adaptées à la détection de régulation coopératives, où plusieurs régulateurs se coordonnent pour affecter un gène cible et pour lesquelles l'expression d'un gène et d'un de ses régulateurs n'est pas nécessairement fortement corrélée. Nous avons développé une approche heuristique originale qui recherche pour chaque gène un réseau "local" représentant l'ensemble de ses régulateurs (activateurs et/ou inhibiteurs). La recherche du meilleur ensemble de régulateurs pour chaque gène cible est restreinte à l'ensemble des régulateurs fréquemment co-exprimés. Cette approche limite drastiquement le nombre d'ensemble de régulateurs candidats à explorer, tout en préservant les régulations coopératives. Elle a été appliquée avec succès à des données d'expression de levure et de patients atteints de tumeurs de vessie. Nous collaborons actuellement avec des statisticiens de l'Université d'Evry pour construire, à partir des réseaux locaux ainsi obtenus, un réseau global reliant les gènes et/ou les régulateurs entre eux, et qui s'appuie sur une information plus complexe que sur la détection de corrélations d'expression directes.

Programmation Logique Inductive



Motifs relationnels "structuraux" de longueur 18 pour les 5 cytochromes P450 4CP4 , 1ROM, 1CPT, 2HPD chain B et 3CPP. Les relations entre éléments (Cα des acides aminés) sont ici leurs distances internes.

Analyse intégrée d'une duplication directe proche interne (DDPI) dans la protéine 1a9x (carbamyl phosphate synthetase) et son gène. La duplication apparaît en vert et en orange dans la figure. La duplication se traduit par deux sous domaines structuraux similaires et symétriques. La duplication 3D est plus étendue que la duplication trouvée en acides aminés, cette dernière étant environ deux fois plus étendue que la duplication nucléique détectée.



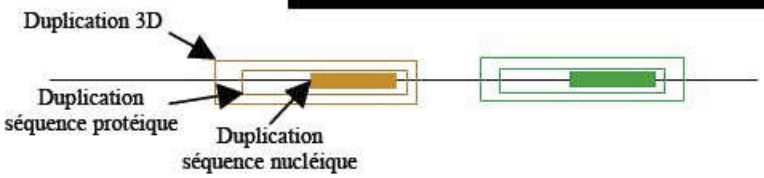


FIG. 1.7: Motifs relationnels structuraux et analyse intégrée d'une duplication 3D

L'équipe aborde dans ce thème les développements algorithmiques pour l'apprentissage dans des fragments de la logique du premier ordre, comme le Datalog non-recursif avec negation, selon deux axes privilégiés : l'utilisation de macro-opérateurs pour l'abstraction de l'espace de recherche, qui bénéficie de l'expertise acquise dans le cadre du thème «Treillis de Galois emboîtés - constructions de taxonomies», et la résolution algébrique du problème d'apprentissage dans des treillis distributifs. Ce deuxième axe s'intéresse à la formalisation et au développement de solveurs algébriques de problèmes d'apprentissage. Cela nous a notamment permis d'identifier des problèmes clés d'apprentissage de type MAX-CSP et CSP pondérés (*Weighted CSP*) qui sont des problèmes déjà étudiés dans l'équipe [CI-22]. Notamment sur ce dernier point, nous avons formalisé et développé une technique de résolution efficace dite «descendante dirigée par les données» dans des restrictions de la logique du premier ordre.

Dans un cadre transversal, nous étudions l'efficacité de ces techniques dans le cadre

empirique de la *transition de phase*. Ce cadre est une étude topologique de la complexité en moyenne de problèmes combinatoires NP-complets en fonction de paramètres d'ordre. Ce travail au sein de l'équipe fait suite aux nombreuses études et publications des différentes communautés d'Optimisation Combinatoire sur les problèmes de type SAT, CSP, TSP etc, avec la même finalité de développer un cadre d'analyse et de comparaison des approches algorithmiques proposées. Notamment, l'étude des transitions de phase permet le développement de générateurs de problèmes difficiles pour tout algorithme afin de faciliter l'analyse des performances des algorithmes en moyenne. Nous avons montré notamment que l'apparition d'une transition de phase (du test de couverture) en PLI était un cas pathologique pour la recherche heuristique des algorithmes d'apprentissage. Cependant, nos premiers résultats montrent que ce cas pathologique n'affecte systématiquement que les stratégies dites « générer-tester » et non pas la stratégie « descendante dirigée par les données ».

1.4.2 Résultats significatifs

- formalisation et développement d'une technique de résolution algébrique dite «descendante dirigée par les données» dans des restrictions de la logique du premier ordre [CI-133] ;
- étude de l'impact de la transition de phase du test de couverture sur la recherche heuristique en Programmation Logique Inductive [RI-82*]

1.4.3 Domaines d'applications

Les domaines visés par les méthodes d'apprentissage symbolique sont multiples, et en pleine extension, car de nombreux utilisateurs réclament des méthodes d'apprentissage dont les résultats sont exprimés sous forme de règles utilisant un vocabulaire explicite pour ces utilisateurs. Par exemple, nous avons travaillé sur des données semi-structurées décrivant des produits informatiques (données C/NET). Les deux autres axes applicatifs privilégiés de l'équipe sont :

- la Bioinformatique. Plus précisément l'équipe collabore avec le laboratoire MIG de l'INRA Jouy-en-Josas, avec l'Atelier de Bioinformatique, de l'Université Paris 6, ainsi qu'avec l'Institut Curie (Paris)
- le traitement de connaissances textuelles, avec le laboratoire MIG de l'INRA Jouy-en-Josas également

Ces domaines d'applications nous permettent d'établir des collaborations au sein de l'Université Paris 13 avec l'équipe RCLN et le Lim&Bio.

1.4.4 Perspectives de recherche

Les perspectives de l'apprentissage à partir d'exemples incomplets et de l'analyse de l'information mise en jeu au cours du processus d'apprentissage, sont à la fois appliquées et théoriques. Sur le plan des applications, la Bioinformatique est un terrain prometteur. En effet les changements de représentations en fonction de critères informationnels semblent intéressants pour guider la recherche de motifs répétés au sein de séquences biologiques : ainsi une séquence de nucléotides (ADN ou ARN) peut être représentée de façon plus abstraite comme une séquence d'acides aminés. Sur le plan théorique, nous nous intéressons à reproduire cette étude en présence d'informations révisables (non monotones).

Une part importante de notre activité est consacrée au développement d'un système hautement configurable d'apprentissage dans des langages de représentation au moins

aussi riches que le Datalog non-récursif. Nous allons valoriser à court terme ce développement en offrant à la fois aux spécialistes et aux praticiens de l'apprentissage une plate-forme pour l'expérimentation et pour la résolution de problèmes d'apprentissage. Cette mise à disposition va permettre de centraliser et de diffuser les différentes spécialisations de la plate-forme effectuées pour les applications ; elle servira également de plate-forme expérimentale pour le développement et la validation d'approches algorithmiques de l'apprentissage développées dans l'équipe A^3 et notamment en Programmation Logique Inductive. Les problèmes ouverts de recherche dans le cadre relationnel que nous comptons aborder avec cette plate-forme sont la sélection/réduction du langage de concepts, en particulier à l'aide de macro-opérateurs, et le traitement de données incomplètes et bruitées.

Enfin, à plus long terme, nous souhaiterions développer un axe de recherche transversal (relationnel et numérique) dans l'équipe sur le thème de de l'*apprentissage relationnel probabiliste*, dont le but est d'effectuer du raisonnement de l'inférence probabiliste sur des représentations logiques, qui nous permettra de modéliser et d'apprendre des connaissances en utilisant la souplesse des représentations probabilistes et l'intelligibilité des représentations logiques.

A^3

1.5 Approches Transversales

Participants : P. Gérard, R. Kanawati, H. Karoui, C. Rouveirol

Cet axe rassemble les deux approches qui font collaborer des techniques d'apprentissage symbolique et numérique, l'apprentissage par renforcement relationnel d'une part, et le raisonnement à partir de cas. L'apprentissage par renforcement (AR) est une approche classiquement numérique. Un système automatique d'apprentissage par renforcement interagit avec son environnement et utilise l'information récoltée pour apprendre un comportement adéquat. Dans chaque situation qui peut se présenter, le système cherche à apprendre quelle action il vaut mieux entreprendre parmi celles qui sont à sa disposition. L'action la meilleure est celle qui devrait lui permettre de maximiser un cumul de récompenses scalaires obtenues sur le long terme. En effet, l'apprentissage se fonde sur une récompense scalaire sporadique, plus facile à définir par un concepteur que le détail des actions à entreprendre dans chaque situation possible. Le système cherche à maximiser le cumul de ces récompenses successives, sans connaître a priori son environnement, en effectuant un compromis entre récompenses immédiates et récompenses attendues dans le futur. Les systèmes d'AR peuvent être envisagés partout où un système autonome doit s'améliorer au fur et à mesure de son expérience. Les domaines d'application les plus courants sont la robotique, les jeux vidéo ou la fabrication industrielle.

A³

Un problème central de l'AR est celui de la généralisation. En effet, dès que la taille des problèmes augmente, il est illusoire de vouloir apprendre quelle action entreprendre dans chacune des multiples situations possibles. Il est alors nécessaire d'utiliser des représentations propres des espaces d'états et d'actions de manière non-extensive. Depuis 2001, l'apprentissage par renforcement relationnel commence à explorer les représentations relationnelles pour remplacer les représentations attributs-valeurs, peu concises, peu aisées à généraliser, mais classiquement utilisées en apprentissage par renforcement. Les états et les actions sont alors représentés par des prédicats relationnels plutôt que par des vecteurs d'attributs. A ce jour, la plupart des travaux en AR relationnel portent sur les systèmes de régression relationnelle employés. Nous envisagerons plutôt l'adaptation des algorithmes classiques d'apprentissage par renforcement aux systèmes de PLI employés.

La deuxième approche transversale developpée dans l'équipe A³ concerne les travaux menés dans le domaine du raisonnement à partir de cas (RàPC). Le RàPC est une méthodologie de résolution de problèmes fondée sur l'adaptation des solutions éprouvées afin de résoudre de nouveaux problèmes similaires à ceux traités dans le passé. Naturellement, l'emploi de cette méthodologie ne peut être envisageable que dans des domaines où des problèmes similaires acceptent des solutions similaires. La figure 1.8 illustre le schéma classique d'une application RàPC : une base de cas dits cas sources regroupe les problèmes résolus et leurs solutions. La résolution d'un nouveau problème passe par cinq étapes successives : 1) l'élaboration qui permet de déterminer les attributs du nouveau problème à résoudre (appelé cas cible), 2) la rémémoration qui permet de trouver les cas sources les plus facilement adaptables par rapport au cas cible, 3) l'adaptation dont le rôle est de concevoir une solution au cas cible à partir des solutions appliquées aux cas sources retenues, 4) la révision de la solution proposée et 5) l'apprentissage qui permet d'ajouter le nouveau cas résolu si opportun.

Ainsi, la relation entre le domaine du RàPC et de celui de l'apprentissage automatique est double : d'une part, la cinquième phase du cycle RàPC assure une fonction d'apprentissage incrémental et paresseux, d'autre part des techniques d'apprentissage, numérique et symbolique, peuvent être employées pour mettre en œuvre les différentes phases du cycle RàPC. Nos travaux dans l'équipe A³ traitent les deux aspects de la relation entre les deux domaines.

Nous proposons une approche RàPC originale, que nous appelons RàPC guidé par les données. Selon cette approche, la mémoire d'un système RàPC est organisée en deux

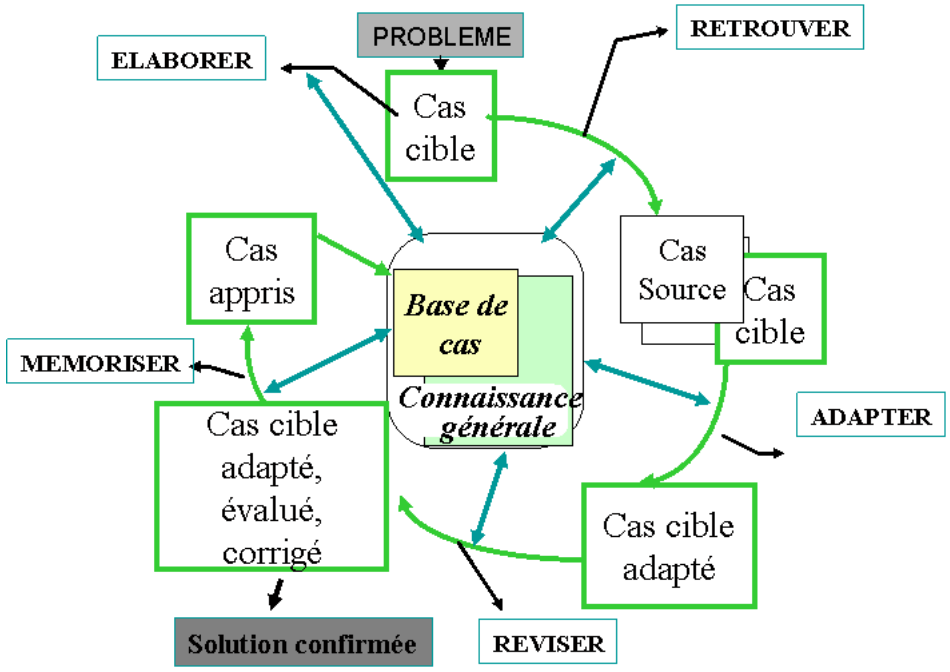


FIG. 1.8: *Le cycle RàPC*

niveaux : la base de cas source et une base de données brutes qui alimente la base de cas. Le cycle RàPC est modifié pour assurer une fonction de maintenance permanante du système (maintenance de la base de cas, des fonctions de similarité et des connsissance d'adaptation). Dans ce cadre, deux axes de recherche sont explorés :

- le dévelopement de systèmes RàPC hybrides combinant des approches symboliques et numériques,
- le dévelopement de systèmes RàPC coopératifs.

1.5.1 Les axes de recherche

Apprentissage par renforcement relationnel Le recherche en apprentissage par renforcement relationnel peut pourtant suivre deux voies complémentaires : d'une part, l'adaptation des algorithmes de programmation logique inductive (PLI) aux contraintes spécifiques des problèmes d'apprentissage par renforcement, largement étudiées à l'heure actuelle dans la communauté de PLI; d'autre part, l'adaptation d'algorithmes de l'état de l'art d'apprentissage par renforcement aux systèmes de PLI employés, pour l'instant plus confidentielle.

Dans un premier temps, nous nous intéressons plus à l'exploitation de la nature relationnelle des données pour améliorer à la fois les techniques d'AR (indirect, hiérarchique etc) de manière à produire des algorithmes efficaces.

Raisonnement à partir de cas A la différence d'un système RàPC classique qui repose pour son fonctionnement sur la disponibilité préalable d'une base de cas sources, un système RàPC guidé par les données vise à alimenter une base de cas, initialement vide, au fur et à mesure de la présentation de nouveaux problèmes à résoudre. La mémoire d'un système RàPC guidée par les données est composée de deux parties : a) une base de cas source initialement vide et b) une base de données brutes contenant des observations sur le système cible. Par exemple, dans le

cadre d'une application de personnalisation automatique de sites Web, les données de trace de navigations (i.e. données de log web) constituent la base de données brutes, et la base de cas est composée d'un ensemble de motifs de navigation qui ont permis de prédire le comportement d'un utilisateur du site cible [CL-27].

Le cycle RàPC classique est modifié afin de prendre en compte la nature de la mémoire du système. D'une manière simplifiée, le nouveau cycle consiste à explorer la base de données brutes afin de proposer des cas sources dits potentiels lorsqu'une des phases de recherche, d'adaptation ou de révision du cycle RàPC classique, échoue. La phase d'apprentissage est aussi modifiée pour ne retenire dans la base de cas que les cas potentiels à partir desquels des bonnes solutions ont été derivées. Ce nouveau cycle permet de constituer une base de cas qui évolue avec la nature des problèmes à résoudre. Des techniques de maintenance asynchrones sont proposées afin d'éliminer de la base les cas à faible contribution [CI-35].

Nous appliquons cette approche dans deux domaines, la *classification et prédiction de séquences* et les *systèmes de recommandations coopératifs*. Le premier axe a fait le sujet de la thèse de F. Zehraoui où nous avons proposé un système hybride de classification et de prédiction de séqeunces de navigation dans un site web. Dans ce travail, les étapes de remémoration et d'adaptation des cas sont implantées sous forme d'une carte auto-organisatrice (la carte M-SOM-ART [CO-33]) conçue pour la classification des séquences et qui possède les deux propriétés de plasticité et de stabilité. Le deuxième axe fait le sujet de la thèse de H. Karoui. Une application pilote de partage et de recommendation de références bibliographiques est développée. Chaque utilisateur est assisté par un agent personel qui observe ses actions afin de détecter ses centres d'intérêt et qui coopère avec les autrs agents afin de recommander de nouvelles références pertinentes pour l'utilisateur associé. Chaque agent implante deux systèmes RàPC guidés par les données afin de 1) calculer les recommandations [CI-104] et 2) apprendre à associer une communauté d'agents à chaque centre d'intérêt détecté [CO-118].

1.5.2 Domaines d'applications

Un domaine d'applications priviligié pour nos travaux en RàPC concerne le dévelopement de système de recommandations [CL-34]. Deux champs d'activités sont particulièrement visés :

- le commerce électronique. Notre approche de traitement et de classification de séquences a été appliquée pour prédire le comportement d'utilisateurs d'un site de E-commerce afin de reconnaître leur intention (i.e. achat ou non achat). Le site peut alors être adapté dynamiquement en fonction de la classe prédite. Cette même approche peut servir aussi pour l'adaptation dynamique d'un site web afin de proposer des raccourcis de navigation.
- l'accès coopératif et intelligent à l'information distribuée. Notre objectif est de permettre à un groupe organisé d'utilisateurs (ex. équipe de recherche) de partager, d'une manière implicite et automatique, leurs expériences de recherche d'information. Trois applications différentes, construites selon la même architecture de type égal-à-égal (P2P), sont proposées : une application de partage de signets web [CL-34], une application de recommandation de références bibliographiques [RE-26] et un méta-moteur de recherche d'information sur le web orienté communautés [CI-1].

1.5.3 Résultats significatifs

Sur le plan théorique les résultats les plus signficatifs sont :

- la définition et la validation de l'approche RàPC guidée par les données [CI-35]
- la proposition et la validation d'un système hybride RàPC connexionniste pour la classification et la prédiction de séquences [CL-27]

Sur le plan applicatif, les réalisations les plus significatives sont :

- Le prototype COBRAS : un système égal à égal pour la recommandation de références bibliographiques. Ce système implante une approche RàPC pour la formation de comité dans les systèmes de type égal à égal [CI-104]
- Le prototype d'un moteur de recherche orienté communauté. Le principe est d'explorer les traces de l'utilisdation du moteur par une communauté d'utilisateurs afin d'améliorer le tri des résultats obtenus par des moteurs classiques de recherche sur le web [CI-13, CI-1].

A³

1.5.4 Perspectives de recherche

L'axe d'apprentissage par renforcement relationnel est un axe qui débute tout juste, et qui est amené à se développer dans le futur. Il s'agit d'un travail novateur et exploratoire qui permet en outre d'initier une collaboration entre le LIPN et l'équipe de Luc de Raedt de l'Université Catholique de Leuven (Belgique).

En ce qui concerne l'axe RàPC, deux axes de recherche reliés sont en cours d'exploration :

- **L'apprentissage coopératif**. Deux problèmes nous semblent particulièrement intéressants : 1) l'identification de commaunautés d'experts dans un système RàPC coopératif et 2) l'étude des différents niveaux de coopération envisageables entre des systèmes RàPC autonomes. En ce qui concerne le premier problème, nous explorons une approche fondée sur l'utilisation des systèmes de réputations et d'analyse de communautés web. En ce qui concerne le deuxième problème, les travaux dans l'état de l'art se concentrent sur des stratégies d'échange de cas entre systèmes RàPC autonomes afin d'améliorer les performances globales du système. Nous voudrons explorer l'utilité d'autres types de coopération notamment la coopération pour la mise en place de chacune des phases du raisonnement, en particulier les phases de la recherche et de l'adaptation. Des algorithmes coopératifs pour le calcul de similarité de structures complexes (i.e. des séquences, des graphes) sont aussi à l'étude.
- **Techniques RàPC pour l'analyse des réseaux sociaux**. L'objectif est d'explorer l'application de la méthodologie RàPC guidé par les données dans des tâches d'identification des réseaux sociaux par l'exploration des traces d'interactions entre acteurs, mais aussi pour des tâches d'analyse des réseaux identifiés notamment pour la prédiction de l'évolution des réseaux et pour la prédiction de flux d'influences dans ce type de réseaux. Les systèmes de recommandation restent des applications cibles favorites, où les réseaux sociaux peuvent être employés à la fois pour enrichir les profils utilisateurs mais aussi pour cibler les utilisateurs de manière à augmenter l'impact des recommendations dans une communauté cible.

1.6 Animation de la recherche

1.6.1 Formation à la recherche

- **Habilitations à Diriger les Recherches**

Francine Krief : Thèse d'HDR en Informatique, Université Paris 6, Les réseaux haut débit : vers la gestion autonome et contextuelle.
Soutenue le 12 décembre 2003.
Coordinateur : G. Pujolle.
Jury : G. Pujolle, N. Agoulmine, D. Gaïti, A. Karmouch, Y. Bennani, P. Dague, O.C.M. Bandeira Durate, S. Zinaty.
Rapporteurs : N. Agoulmine, D. Gaïti, A. Karmouch.
Situation actuelle du diplômé : Professeur à Bordeaux.

- **Thèses soutenues**

Alejandro Guerra-Hernandez : Apprentissage d'Agents Rationnels BDI dans un univers Multi-Agents.
Financement de la thèse : Bourse Mexique
Soutenue le 6 Décembre 2003.
Directeurs : Amal Elfallah-Segrouchni et Henry Soldano.
Situation actuelle du diplômé : Maître de Conférences, Informatique Université de La Cruz (Mexique).

Khalid Benabdeslem : Approches Connexionnistes pour la classification et la visualisation des séquences évolutives : Application aux données issues d'usages de l'Internet.
Financement de la thèse : CDD sur Contrat de recherche.
Soutenue le 17 Décembre 2003.
Situation actuelle du diplômé : Maître de conférences à l'Université Lyon 1.

Farida Zehraoui : Systèmes d'apprentissage connexionnistes et raisonnement à partir de cas pour la classification et le classement de séquences.
Financement de la thèse : CDD sur Contrat de recherche.
Soutenue le 2 octobre 2004 (Très Honorable+Félicitations).
Situation actuelle du diplômé : Maître de conférences à l'Université d'Evry.

Hakima Kadri-Dahmani : Mise à jour incrémentale des bases de données géographiques et maintien de leur cohérence.
Financement de la thèse : Institut Géographique National (IGN france) Soutenue le 22 novembre 2005.
Directeur : Aomar Osmani, Philippe Dague

Nahla El-Zant : Recherche de motifs relationnels dans les séquences : Application aux séquences biologiques.
Financement de la thèse : Allocataire.
Soutenue le 1er Décembre 2005.
Directeurs : Henry Soldano et Daniel Kayser.

Zeina El Jrad : Apport des techniques de l'intelligence artificielle dans la négociation dynamique de niveaux de service : Proposition d'une interface utilisateur pour l'Internet de nouvelle génération.
Financement de la thèse : CDD sur projet RNRT.
Soutenue le 15 mai 2006.
Directeur : Francine Krief.
Situation actuelle du diplômé : Post-doc, INRIA (Rocquencourt).

Alexandre Pauchet : Modélisation cognitive d'interactions humaines dans un cadre de planification multi-agents.
Financement de la thèse : Allocataire.
Soutenue le 22 Septembre 2006. (Très Honorable+Félicitations)
Directeurs : Amal El Fallah-Seghrouchni et Nathalie Chaignaud
Situation actuelle : Post-doc, France Télécom (Lannion).

Hassine Moungla : Evolution vers une Gestion Autonomique des Réseaux de Télécommunications : Extension de l'approche PBM.
Financement : Allocataire
Soutenue le 20 Novembre 2006.
Directeur : Francine Krief

Nejla Amara-Hachmi : Une architecture générique d'agents mobiles sensibles au contexte et auto-adaptables.
Financement : Bourse Tunisie
Soutenue le 6 Décembre 2006.
Directeur : Amal El Fallah
Situation actuelle : Post-doc, France Télécom.

Sébastien Guérif : Réduction de dimension en apprentissage non supervisé.
Financement de la thèse : Bourse CIFRE.
Soutenue le 11 Décembre 2006.
Directeur : Younès Bennani
Situation actuelle : MCF Université Paris 13.

- **Thèses en cours**

Sujeevan Aseervatham : Thèse Paris 13, depuis 10/2004.
Sujet : Apprentissage à noyaux et réduction de dimension.
Directeurs : Emmanuel Viennet et Younès Bennani
Financement : ATER

Guénaël Cabanes : Thèse Paris 13, depuis 10/2007.
Sujet : Classification topographique évolutive à deux niveaux simultanés.
Directeur : Younès Bennani
Financement : Allocataire

Lahcène Dehni : Thèse Paris 13, depuis 10/2004.
Sujet : Anlyse et visualisation des données évolutives.
Directeur : Younès Bennani

Financement : ATER

Mohamed Elati : Thèse Paris 13, depuis 10/2006.
Sujet : Apprentissage symbolique pour la bio-informatique.
Directeur : Céline Rouveirol
Financement : ATER

Hager Karoui : Thèse Paris 13, depuis 11/2003.
Sujet : Un *framework* multi-agents pour l'aide à l'accès coopératif et intelligent à l'information distribuée.
Directeurs : Rushed Kanawati et Laure Petrucci.
Financement : ATER

Nicoleta Rogovschi : Thèse Paris 13, depuis 10/2006.
Sujet : Apprentissage et réseaux sociaux.
Directeurs : Emmanuel Viennet et Younès Bennani
Financement : CDD sur projet ANR

Nistor Grozavu : Thèse Paris 13, depuis 10/2006.
Sujet : Apprentissage non supervisé et pondération local des variables.
Directeur : Younès Bennani
Financement : CDD sur projet InfoMagic

- **Encadrement de stages de recherche**
  - 2004 (Y. Bennani) : F. Ilo (DEA (IA-OC) Université P13), *Outils connexionnistes pour la prévision d'usures*
  - 2004 (Y. Bennani) : B. Laarabi (DEA Statistiques Université P7), *Sélection de variables en classification automatique*
  - 2004 (Y. Bennani et F. Krief) : L. Dehni (DEA (IA-OC) Université P13), *Apprentissage non supervisé et réseaux de capteurs*
  - 2006 (Y. Bennani et S. Guérif) : M. Belgacem (Master Informatique (MICR) Université P13), *Apprentissage connexionniste non supervisé avec pondération et sélection de variables*
  - 2006 (E. Viennet) : B. Berrabah (Master Informatique (MICR) Université P13), *Méthodes locales pour la réduction non-linéaire de dimension : Applications à la fouille de textes*
  - 2006 (D. Bouthinon et H.Soldano) : J. Boussaa (DESS Université CADI AYYAD, Marrakech (Maroc)), *Elaboration d'un système d'apprentissage d'ordre 1 à partir d'exemples ambigus*
  - 2006 (Y. Bennani et C. Recanati) : N. Rogovschi (Master Informatique (MICR) Université P13), *Systèmes d'apprentissage non supervisé connexionnistes et stochastiques pour la fouille de données structurées en séquences*
  - 2006 (R. Kanawati) : M. Lafer (Master Informatique Recherche, Université Paris Sud), *Identification de communautés d'experts dans un système P2P.*
  - 2007 (Y. Bennani) : M. Chelli (Master Informatique (MICR) Université P13), *Apprentissage non supervisé avec pondération locale des variables*
  - 2007 (Y. Bennani) : G. Cabanès (Master Informatique (MICR) Université P13), *Apprentissage non supervisé évolutif*
  - 2007 (P. Gérard) : Ch. Rodrigues (Master Informatique (MICR) Université P13), *Architectures d'Apprentissage par Renforcement Relationel*

- **Enseignements liés à la recherche**

Le travail de recherche d'une grande partie des membres de l'équipe a débouché sur la création de nouveaux enseignements au niveau du Master d'Informatique :

- 2003-2005 (Y. Bennani) : Systèmes d'apprentissage connexionnistes pour la ouille de données, DEA IA-OC Paris 8 - Paris 13
- 2003-2005 (H. Soldano et M. Champesme) : Apprentissage symbolique (option), master 2 Informatique, spécialité EID, Paris 13
- 2005-2006 (Y. Bennani) : Apprentissage numérique (option), master 2 Informatique, spécialité MICR, Paris 13
- 2005-2006 (Y. Bennani) : Apprentissage connexionniste (option), master 2 Informatique, spécialité EID, Paris 13
- 2005-2006 (Y. Bennani) : Reconnaissance statistique des formes (option), master 2 Informatique, spécialité EID, Paris 13
- 2005-2006 (E. Alphonse) : Apprentissage symbolique (option), master 2 Informatique, spécialité EID, Paris 13
- 2005-2007 (A. Osmani et R. Kanawati) : Apprentissage symbolique et RàPC (option), master 2 Informatique, spécialité MICR, Paris 13
- 2006-2007 (Y. Bennani et P. Gérard) : Apprentissage numérique (option), master 2 Informatique, spécialité MICR, Paris 13
- 2006-2007 (C. Rouveirol) : Fouille de données relationnelles (option), master 2 Informatique, spécialité MICR, Paris 13
- 2007-2008 (C. Rouveirol et H. Soldano) : Apprentissage symbolique, master 2 Informatique, spécialité MICR, Paris 13
- 2007-2008 (C. Rouveirol et J.D. Zucker) : Fouille de Données complexes en Bioinformatique, master 2 Informatique, spécialité MICR, Paris 13

1.6.2 Valorisation de la recherche

• Contrats de recherche

Type	Coordinateur	Resp. scient.	Nom	Financeur	Période	Durée (mois)	Budget global (Keuros)	Part LIPN (Keuros)
ANR		Y. Bennani	CADI	MENESR	2008-2010	24		200
ANR		E. Viennet	SEPTIA	MENESR	2007-2009	24	554	100
ANR		Y. Bennani	SILLAGES	MENESR	2005-2008	36	375	113
ARC	•	Y. Bennani	EUROC25	ONERA	2006-2006	6		3
CAP DIGITAL		A. Nazarenko	Infom@gic	DGE	2005-2006	12		48
BQR		Y. Bennani	Faces	Univ. Paris 13	2005-2006	12	30	5
CIFRE		Y. Bennani	AppMark	NUMSIGHT	2004-2006	24		10
ACI-IMPBio		H. Soldano	EVOLREP	MENESR	2003-2005	24		40
CIFRE		Y. Bennani	FouDon	GEOBS	2003-2004	12		8
RNRT		F. Krief	IP-SIG	MENESR	2003-2004	24		120
RNTL		Ph. Dague	RESEDA	MENESR	2001-2003	48		200
RNRT		Ph. Dague	MAGDA2	MENESR	2001-2003	48		180
R&D	•	Y. Bennani	WebMining	NUMSIGHT	2000-2003	48		151
5ième PCRDT	•	Ph. Dague	IDD	UE	2000-2003	36		145

A³

Au niveau international

• Projet européen IDD

« Integrated Design process for on-board Diagnosis » (2000-2003)
Partenaires du projet :

- LIPN-CNRS : Ph. Dague
- Fiat, Daimler-Chrysler, PSA, Renault, Magneti-Marelli, OCC'M et les universités de Munich, de Turin.

Objectifs : L'objectif du projet IDD, « Integrated Design process for on-board Diagnosis », soumis dans le cadre du premier appel d'offres du programme Croissance compétitive et durable du 5ième PCRDT, était d'intégrer l'approche du raisonnement à base de modèles au processus de conception et de développement, afin de prendre en compte dès cette phase les nécessités du diagnostic. Nous sommes convaincus que le raisonnement à base de modèles ne pénétrera réellement l'industrie que s'il accompagne tout le cycle de vie d'un produit, de la conception à la maintenance, en passant par le développement, le test et le diagnostic, avec un partage des modèles dans ces différentes tâches et leur mise à jour en amont au niveau de la conception. Les principales tâches étudiées dans ce projet ont été la diagnosticabilité (détectabilité et discriminabilité) et la génération de systèmes de diagnostic à bord. Le LIPN a été en charge de la conception et la réalisation

d'interfaces entre les outils de CAO ou de simulation numérique (MATLAB/Simulink dans ce cas) et les systèmes de diagnostic à base de modèles (RAZ'R d'OCC'M dans ce cas), ce qui a nécessité des études sur la transformation automatisée de modèles (modèles numériques vers modèles qualitatifs). Il a réalisé le logiciel MOUSE (MOdel StrUcture Extraction) qui extrait les caractéristiques structurelles (composants et connexions) du modèle MATLAB/Simulink et discrétise le modèle numérique de simulation en un modèle qualitatif de diagnostic au niveau des composants, fournissant en sortie, au format xml, un modèle directement exploitable par RAZ' R. L'application retenue par PSA était le système de climatisation du véhicule et MOUSE a exploité le modèle Simulink de ce système.

- **Projet de coopération FS-KDDM**

A^3

Feature Selection for Knowledge Discovery and Data Mining. (1999-2003)
Partenaires du projet :

– LIPN-CNRS : Y. Bennani
– University of Skopje : D. Cakmakov

Objectifs : Ce projet porte sur l'étude de la fonction de base d'une méthode de sélection de variables (caractéristiques) qui est de choisir un sous ensemble de variables pertinentes à partir de la représentation vectorielle des formes observées. Généralement la phase de sélection et de décision (classement) s'effectuent l'une après l'autre dans un système classique de reconnaissance des formes, les méthodes connexionnistes sont très attrayantes car elles sont capables de mener les deux phases (sélection et classement) en parallèle : les variables sélectionnées optimisent le critère d'apprentissage. Y. Bennani et M. Yacoub ont proposé dans le cadre de ce projet quelques mesures de pertinence dont une dépendante des données et du système et une autre dépendante du système. Y. Bennani a aussi proposé une autre approche qui consiste à pondérer les variables en fonction de leurs participation à la discrimination. Ce projet a donné lieu à la publication d'un ouvrage sur le thème : Cakmakov D. & Bennani Y. (2002) , *Feature Selection for Pattern Recognition*, Informa Press, Ed. ISBN 9989-943-02-8.

- **Projet de coopération, COCPRA**

Combining and Optimizing Classifiers for Pattern Recognition Applications. (2004-2007)
Partenaires du projet :

– LIPN-CNRS : Y. Bennani
– University of Skopje : D. Cakmakov, B. Jaksimovski

Objectifs :

Ce projet concerne la conception et la mise en oeuvre de systèmes hybrides qui font coopérer différents rmodèles connexionnistes entre eux et avec des techniques plus classiques dans différents domaines d'application. Ces systèmes constituent une autre génération des modèles connexionnistes et apparaissent comme la seule solution pour la réalisation de tâches complexes et de grande taille. Les études théoriques portent sur les algorithmes d'apprentissage pour ce type de système, la décomposition de tâches, la conception de chaînes de traitement optimales. Y. Bennani a travaillé sur la possibilité de coopération des algorithmes connexionnistes et d'autres méthodes plus classiques comme les Chaînes de Markov Cachées, afin de concevoir des systèmes modulaires hybrides. Ces architectures présentent en théorie l'avantage de combiner les points forts de leurs composants et par exemple d'obtenir des systèmes qui soient à la fois discriminants et

incrémentaux. Y. Bennani et V. Radevski ont proposé aussi des approches pour la fusion de plusieurs types de données (structurelles et statistiques) et pour la fusion de décisions dans des systèmes multi-modulaires.

- **Projet de coopération, DCL&WSN**

Distributed Connectionist Learning for Wireless Sensor Networks " avec Federal University of Ceará, Brazil. (2005-2007)
Partenaires du projet :

– LIPN-CNRS : Y. Bennani
– Federal University of Ceará : J. Neuman de Souza, D. Gomes

Objectifs : Les estimations de la bande passante allouée au trafic vidéo IP sont réalisées par le biais de deux techniques de prédiction complémentaires. Une des contributions a consisté à intégrer un module connexionniste de prédiction dans une architecture de provisionning dynamique de la bande passante entre un fournisseur de VoD et un ISP (Internet Service Provider). Le fonctionnement de cette architecture est géré par un ensemble de politiques à l'instar de l'architecture PBM (Policy Based Management) de l'IETF (Internet Engineering Task Force) et réalisé à travers les technologies du Web. En effet, l'automatisation du provisionning est un concept nouveau qui n'est pas encore bien accepté au niveau des opérateurs qui souhaitent que toutes les procédures soient validées par des contrats papiers. Néanmoins, la complexité des changements associés à la dématérialisation de la société pousse vers des solutions d'automatisation complète de l'installation et de la configuration de nouveaux services. Pour réduire leurs frais opérationnels, les opérateurs vont recourir de plus en plus à l'automatisation des configurations de services et à leur contrôle. Ce domaine en pleine expansion s'appelle : *l'autonomic networking* ou encore les réseaux auto-gérables. L'objectif principal de cette étude consiste à développer des algorithmes d'apprentissage numériques adaptés à un environnement distribué. Ce type d'algorithmes est plus adaptés à notre domaine d'application.

Au niveau national

- **Projet Infom@gic**

[Dans le cadre du pôle de compétitivité Cap Digital (IMVN : Image, Multimédia et Vie Numérique)] (2006-2008)
Pilote industriel : Thales
Partenaires du projet :
BERTIN, EADS, EUROPLACE, INA, INTUILAB,ODILE JACOB, ONERA, PERTIMM , SPIKENET, TEMIS, THALES, VECSYS, XEROX , CEA, ENST, Univ. Paris 6, Univ. Paris 8, Univ. Paris 13, Univ. Paris Dauphine, Univ. Orsay, Univ. Marne la Vallée, CNRS

Objectifs : Ce projet consiste à mettre en place, sur une période de trois ans, un laboratoire industriel de sélection, de test, d'intégration et de validation sur des applications opérationnelles des meilleures technologies franciliennes dans le domaine de l'ingénierie des connaissances. Ce laboratoire s'appuie sur une plate-forme commune d'interopérabilité. Cette plate-forme doit couvrir les grands domaines fonctionnels et techniques de l'analyse d'information que sont la recherche et l'indexation, l'extraction de connaissances, et la fusion d'informations multimédias, sur tous les types de sources : texte, données, images, son. Elle incorpore des applications des technologies développées dans les domaines de la e-Education, et de la gestion des patrimoines culturels numériques.

Participation du LIPN-CNRS :

- Sous-Projet 2 : Extraction d'informations (ST 2.12), A. Nazarenko, T. Poibeau.
- Sous-Projet 3 : Analyse et fusion d'informations (ST 3.12 et ST 3.21), S. Aseervatham, Y. Bennani, N. Grozavu, S. Guérif, M. Lebbah, E. Viennet.

- **Projet (ANR) CADI**

Composants Avancés pour la DIstribution (2008-2010)

Partenaires du projet :

- LIPN-CNRS, Younès Bennani
- LIP6-CNRS, Massih-Reza Amini
- LITIS, Stéphane Canu
- KXEN, Erik Marcade
- Numsight, Eric Janvier
- SPACECODE, Georges Folcke
- SAMSE - La Boîte à Outils, Frédéric Dufau-Joël
- MONDOMIX, Marc Benaïche

Objectifs : Le projet CADI vise à produire un prototype de composants logiciels pour la Distribution, tant brick-and-mortar (magasins) que vente en ligne (site de e-commerce). Dans le nouveau modèle économique de la distribution, dit de la « Longue Traîne », on voit émerger d'un catalogue potentiellement illimité des multiplicités de marchés de micro-niches que le site commerçant peut espérer développer pour dériver des revenus importants. Pour cela, le site doit aider le client à trouver le produit susceptible de l'intéresser dans des micro-niches de produits adaptées à ses goûts et ses attentes et adapter ses pratiques marketing et commerciales traditionnelles pour prendre en compte les communautés qui se créent sur les blogs et les forums autour de ses produits. Visant un marché à l'intersection de quatre domaines en forte croissance (e-commerce, data mining, réseaux sociaux et RFID), représentant un marché potentiel de plusieurs milliards d'euros, le projet CADI produira, dans une approche Web services sur une architecture SOA, les prototypes des composants logiciels suivants :

- Analyse de comportement d'achat
- Moteur de recommandation
- Moteur de rating
- Analyse des communautés

- **Projet (ANR) SEPTIA**

Solution automatique pour le traitement des interactions clients (2007-2009)
Partenaires du projet :

- LIPN-CNRS : E. Viennet
- LIP6-CNRS : P. Gallinari
- KXEN : F. Fogelman
- TEMIS : Charles Huot
- EPTICA : Thierry Gandilhon

Objectifs : Le projet Septia vise à produire un logiciel avancé de traitement des interactions clients dans le cadre d'un centre de contacts. Axé de façon prédominante sur le traitement des emails, le projet étudiera également l'apport des données clients, des

données d'interactions téléphoniques et des données d'interactions inter-clients dans un portail collaboratif. Les caractéristiques principales :

- Extraction automatisée des thèmes des emails et constitution des ontologies associées, en utilisant de stechniques statistiques et sémantiques.
- Aide à la réponse (recherche automatique du "meilleur" interlocuteur.
- Supervision des activités (prévision de séries temporelles)
- Portail collaboratif : étendre les otils à un portail (forum, base client, chats, ...)

Le projet prévoit la construction d'un démonstrateur qui sera testé sur des données réelles de grande taille. Il sera l'occasion de développer des technologies avancées dans le domaine de la fouille de texte, de la fouille de données et de l'étude des "réseaux sociaux".

- **Projet (ANR) SILLAGES**

Suivi Individuel par Label éLectronique Au sein de Groupes complexES. (2006-2008) Partenaires du projet :

- LIPN-CNRS : Y. Bennani
- LEEC-CNRS : D. Fresneau
- ULB (Université Libre de Bruxelles) : J-L. Deneubourg
- SPACECODE : G. Folcke

Objectifs : Les structures sociales animales sont des systèmes dynamiques caractérisées par un grand nombre d'interactions entre individus. Malgré leur diversité, ces structures sont basées sur un nombre limité de règles génériques (amplifications, inhibitions, compétition pour les ressources). Les dynamiques du groupe et ses réponses résultent de la synergie entre ces interactions, des capacités individuelles à traiter l'information et de la diversité de ces réponses individuelles.

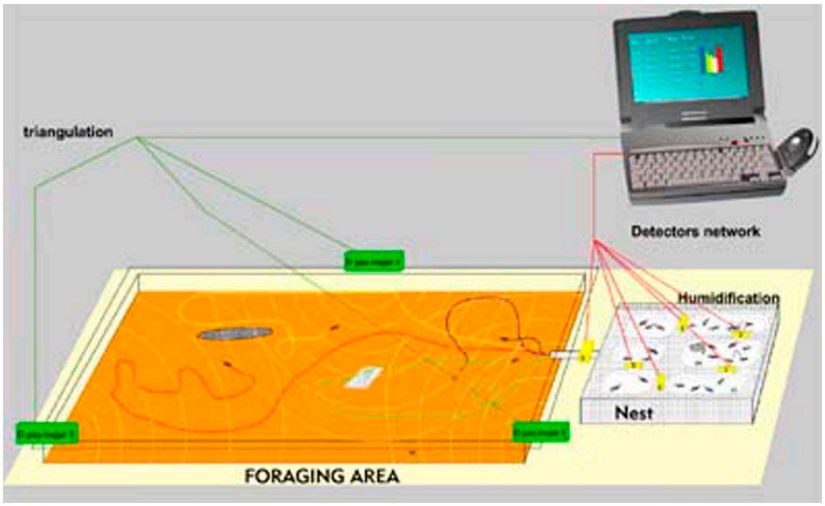


Fourmi avec le capteur RFID

Les objectifs de ce projet sont :

- le développement de nouveaux systèmes autonomes de suivi de l'activité spatio-temporelle des groupes, systèmes donnant accès à des données inaccessibles actuellement ;
- le développement d'outils d'analyse automatique de ces données et de quantification des comportements individuels ou collectifs ;
- l'intégration de ces quantifications aux outils de modélisation qui permettent de valider le lien entre les niveaux individuels et collectifs ;
- ces systèmes autonomes de suivi pourront contrôler des dispositifs et modifier l'environnement en fonction de l'identité des individus. Ceci permettra de tester notamment des hypothèses portant sur la plasticité comportementale.

Ces objectifs font de SILLAGES une action interdisciplinaire à l'intersection des sciences du comportement et du complexe et des sciences informatique et de l'ingénieur.



Dispositif expérimental RFID pour le suivi des fourmis

A³

- **Projet EVOLREP, ACI-IMPBio**

Impact et évolution des répétitions dans les génomes : interactions entre innovation génétique et fonction biochimique (2003-2005).
Porteur : J. Pothier.
Atelier Bioinformatique, UFR 927, Université Pierre et Marie Curie, Paris.

Partenaires du projet :

- Unité Génétique des Génomes Bactériens, URA CNRS 2171, Institut Pasteur
- Equipe Structure et Dynamique du Génome, Institut Jacques Monod, UMR CNRS 7592
- Equipe Helix, INRIA Rhône-Alpes
- LIPN-CNRS : H. Soldano
- Equipe Contrôle de l'Expression Génique Laboratoire Adaptation et Pathogénie des Microorganismes, (CNRS UMR 5163), Université Joseph Fourier, Grenoble.

Objectifs : Une des énigmes majeures de l'évolution des génomes reste le rôle des phénomènes de duplication de l'ADN. Les évènements de duplication sont à l'origine de l'accroissement de la taille des génomes, des familles de gènes et de la création de la nouveauté fonctionnelle. Malgré cela, le rôle et les raisons de leur existence et les mécanismes qui contrôlent leur développement sont encore peu connus. Ce projet avait pour ambition l'étude intégrée des duplications intra-géniques dans les génomes des micro-organismes complètements séquencés (archeae, bactéries et levures) aux trois niveaux de description que sont les séquences nucléiques, protéiques et les structures 3D des protéines. Différentes méthodes de recherche (classiques ou développées dans le cadre du projet) ont été utilisées pour analyser les interactions que ces duplications intra-géniques peuvent avoir aux trois niveaux de description précités. En particulier le LIPN a travaillé sur l'algorthmique de la recherche de motifs structuraux répétés dans les structures 3D de protéines.

Dans le cadre de ce contrat, le CNRS (département STIC) a accordé une bourse de post-doc d'un an au LIPN, attribuée à Nadia Pisanti.

- **Bonus Qualité Recherche (BQR) [Face] (2005-2006)**

Projet de coopération LIPN-LEEC (BQR) : Détection et reconnaissance des mouvements faciaux liés au genre chez l'Homme. (2006-2007)
Partenaires du projet :

- LIPN-CNRS : Y. Bennani, S. Guérif
- LEEC-CNRS : C. Baudoin

Objectifs : Ce projet concerne l'étude de la signalisation et de la reconnaissance du genre à partir des mouvements faciaux. La prise en compte des mouvements faciaux est relativement récente et a été effective dans certains travaux portant sur l'expression des émotions et dans la communication parlée. Curieusement malgré des études comparatives, déjà anciennes sur les communications non verbales, peu de recherches récentes concernent le rôle des mouvements faciaux dans la signalisation d'appartenance à un groupe et dans la reconnaissance sociale chez l'Homme. Les objectifs du projet sont :

- de détecter les mouvements faciaux émis par des hommes et des femmes, appartenant à différents groupes culturels, dans des situations standardisées en situation contrôlée de laboratoire,
- de vérifier les résultats d'une étude pilote indiquant l'existence de mouvements faciaux latéralisés indicateurs du genre,
- de définir si ces mouvements sont perçus par des sujets « juges » des deux sexes, indépendamment d'indices morphologiques du visage,
- d'explorer ces mêmes informations par des réseaux connexionnistes (supervisés et non-supervisés),
- d'interpréter les phénomènes de causalité à partir des résultats de l'exploration de séquences spatio-temporelles des visages au moyen de modèles graphiques de type réseaux bayésiens.

- **Convention Industrielle de Formation par la Recherche, Numsight**

CIFRE N° 654/2002 : Extraction de connaissances à partir de données évolutives : application aux données comportementales. (2003-2006)

Partenaires du projet :

- LIPN-CNRS : Y. Bennani, S. Guérif
- NumSight : E. Janvier

Objectifs : Ce projet de recherche possède deux aspects : un aspect fondamental qui consiste à développer de nouveaux algorithmes d'apprentissage non supervisé pour le traitement et l'exploration de données comportementales, et un aspect applicatif qui rentre dans le cadre des activités de Numsight Consulting France en ingénierie de Marketing. Dans le cadre de cette convention CIFRE, nous avons proposé d'étudier un ensemble de problématiques des modèles connexionnistes et leur mise en oeuvre pour lesquels nous pensons qu'ils offrent des avantages et des spécificités par rapport aux techniques classiques d'analyse des données.

- L'étude est structurée en quatre phases :
- Modélisation globale : cherche un modèle unique de représentation qui résume d'une manière globale l'ensemble des données.
- Étude de l'interstructure : consiste à identifier des proximités ou des différences entre les ensembles de données, sans description.
- Étude de l'intrastructure : permet grâce à une analyse fine des ressemblances ou des différences d'expliquer la structure identifiée.

– Étude de l'évolution des informations : analyse et décrit l'évolution du phénomène d'une manière globale ou étudier les évolutions de chacun des éléments.

- **ARC-ONERA**

Action de Recherche Coopérative avec l'ONERA « Algorithmes d'apprentissage artificiel pour le traitement des profils atmosphériques ». (2005-2006).

Partenaires du projet :

– LIPN-CNRS : Y. Bennani, S. Guérif
– ONERA : T. Huet

Objectifs : Cette étude consiste à définir une méthode de classification automatique permettant une réduction notable des profils atmosphériques constituant les scènes EUROC25 issus de Météo-France. Le domaine EUROC25 comprend 129 points Ouest-Est par 105 points Nord-Sud. Chaque point de grille est caractérisé par 21 niveaux comprenant l'altitude, la pression, la température et l'humidité relative. La situation du profil par rapport à la mer est également précisée dans le fichier. Sous-jacente à ces données, une information fréquentielle est également présente, sous la forme d'un coefficient d'absorption calculé à partir des données précédemment citées. L'étude s'articule autour de l'exploration et l'extraction de connaissances à partir de données EUROC25. L'évolution des données au cours du temps et leur position spatiale nécessitent d'explorer simultanément plusieurs ensembles de données multi-variables. En statistique, des méthodes d'analyse factorielle multiple et de classification ont été développées ces dernières années, mais elles sous-estiment le plus souvent le caractère spatial des données. Par contre, dans le domaine de l'apprentissage artificiel, et en particulier connexionniste, le traitement de ce type de données est un problème qui représente une voie de recherche très prometteuse ouvrant d'excellentes perspectives. Dans le cadre de cette prestation entre l'ONERA et le Laboratoire d'Informatique de Paris Nord (LIPN, UMR 7030 du CNRS), Y. Bennani et F. Carton (étudiant en Master Pro EID) ont proposé d'étudier un ensemble de problématiques des systèmes d'apprentissage connexionnistes et leur mise en oeuvre pour lesquels ils pensent qu'ils offrent des avantages et des spécificités par rapport aux autres techniques. Cela les conduira à réaliser soit des études ponctuelles sur quelques points très ciblés qui doivent pouvoir s'insérer dans d'autres études déjà menées à l'ONERA, soit à développer des nouveaux outils connexionnistes de classification automatique.

- **Projet RESEDA [RNTL]**

RESEDA (Réalisation d'un Environnement Système Evolutif pour le Diagnostic Automobile). (2001-2005)

Partenaires du projet :

– LIPN-CNRS : S. Aseervatham, Y. Bennani, P. Dague, A. Osmani
– Renault : P. Dubois
– Silicomp : V. Joloboff, A. Boutin
– Trialog : P. Robin

Objectifs : Ce projet se propose 3 objectifs :

– Définir une architecture de composants logiciels évolutifs de diagnostic en utilisant des techniques d'objets distribués, et la technologie de code mobile Java.

- Construire un démonstrateur qui permettra de valider l'infrastructure en utilisant l'état de l'art des méthodes de diagnostic chez les constructeurs.
- Elaborer, dans une démarche plus exploratoire, les méthodes et outils qui permettront de bénéficier du retour d'expérience sur le parc de véhicule en circulation pour faire évoluer les composants de diagnostic en tenant compte des problèmes constatés sur le terrain.

- **Projet MAGDA 2 [RNRT]**

MAGDA 2 : Modélisation et Apprentissage pour une Gestion Distribuée des Alarmes De bout En boUt . (2001-2003)

Partenaires du projet :

- LIPN-CNRS : P. Dague, F. Krief, A. Osmani, E. Viennet
- Alcatel : G. Desoblin, H. Papini, A. Aghasaryan, E. Marilly, B. Lapraye
- ILOG : B. Berstel, P. Feillet, E. Mazeran, C. de Sainte Marie
- France Telecom : D.Collobert, C. Dousson
- IRISA : A. Benveniste, Eric Fabre, C. Jard, M-O. Cordier

Objectifs : L'objectif du projet MAGDA 2 est de trouver des solutions avancées aux problématiques de la gestion des réseaux hétérogènes et de la prise en compte d'événements faisant intervenir l'interaction entre la couche réseau et la couche service . A cette fin, nous envisageons d'une part l'étude et le développement d' algorithmes robustes de corrélation et de diagnostic pour les réseaux de télécommunication ainsi que pour l'interaction entre la couche réseau et la couche services. Par ailleurs, nous développerons des techniques d'automatisation (partielle) de l' acquisition des connaissances et modèles nécessaires. Le projet MAGDA2 s'appuie de manière essentielle sur le succès du projet MAGDA. Il vise à étendre les résultats aussi bien techniques qu'industriels de MAGDA en profitant des connaissances communes acquises et de la dynamique que les partenaires ont su créer au sein de ce projet. Plus, Il vise à lever des limitations identifiées par MAGDA et à résoudre les nouveaux problèmes qu'engendre la mise en oeuvre industrielle des techniques développées dans ce projet.

- **Projet IP-SIG [RNRT]**

IP-SIG : Signalisation générique du monde IP. (2003-2004)
Partenaires du projet :

- LIPN-CNRS : F. Krief
- ENST, ALCATEL, UTT, ISEP, LIP6, THALES

Objectifs : Le monde IP a longtemps lutté pour se passer d'une signalisation grâce à l'adresse IP qui est routée et qui se situe dans tous les paquets IP. Cependant, depuis déjà une dizaine d'années, le monde IP se pose la question d'une signalisation permettant de gérer le réseau différemment des solutions existantes. Les évolutions MPLS et GMPLS, la voix sur IP ainsi que les techniques de contrôle par politique ont abouti à de nouvelles signalisations complètement incompatibles. L'IETF a décidé en 2002 de lancer un groupe de travail pour essayer d'unifier toutes ces solutions ou tout au moins de les faire coexister. Le but de ce projet est de participer aux travaux de l'IETF en proposant une norme de signalisation universelle capable de répondre au moins aux besoins de sécurité, de QoS et de mobilité. Notre idée dans ce projet est d'explorer la voie du protocole COPS comme protocole fédérateur. La raison en est simple, COPS est la solution poussée par

de nombreux industriels pour arriver à faire de la qualité de service, COPS est flexible et peut parfaitement s'adapter à la gestion de la sécurité et de la mobilité.
Les objectifs de ce projet sont :

- La définition d'une architecture de signalisation universelle qui est capable de mettre en place une connexion sécurisée tout en tenant compte des besoins de QoS et de mobilité.
- La définition d'un SLS standard.
- Une étude prospective sur la dynamisation du SLS.

• ARC-THALWEG

Projet ARC-THALWEG : Action de Recherche Coopérative avec le projet AIR de l'INRIA, "Modèles de traitement d'image pour les données turbulentes". (2002-2004)
Partenaires du projet :

- LIPN-CNRS : Y. Bennani
- LMD-Ecole Polytechnique : M. Desbois, R. Roca, A. Szantai
- LPS-ENS-Ulm : J.P. Nadal
- INRIA-Rocquencourt : H. Yahia, I. Herlin, J. Grazzini, A. Turiel

Objectifs : Ces travaux sont basés sur l'élaboration de nouvelles méthodes de traitement d'image dans lesquelles on s'intéresse à la caractérisation de la turbulence dans un signal. Il est donc fait appel à des résultats récents utilisant les domaines de la physique et du traitement du signal, de la théorie des ondelèttes, des multifractales, du traitement d'image et de l'apprentissage artificiel. Ces méthodes sont actuellement appliquées à la détermination de certaines zones caractéristiques dans des images météorologiques infrarouges Meteosat. On cherche à segmenter les zones correspondant à une activité pluviométrique dans un nuage, en tenant compte des propriétés de turbulence dans le signal image. Cela conduit à la détermination de zones particulières appartenant à la "MSM ("Most Singular Manifold" ou "variété la plus singulière") contenant l'information des zones de transition maximale dans le signal. Ces ensembles sont de nature multifractale. Dans une première phase, on a donc tenté de définir des attributs reliés à la MSM et conduisant à des opérateurs de traitement d'image permettant la détermination des zones pluviométriques.

• Numsight [Web mining]

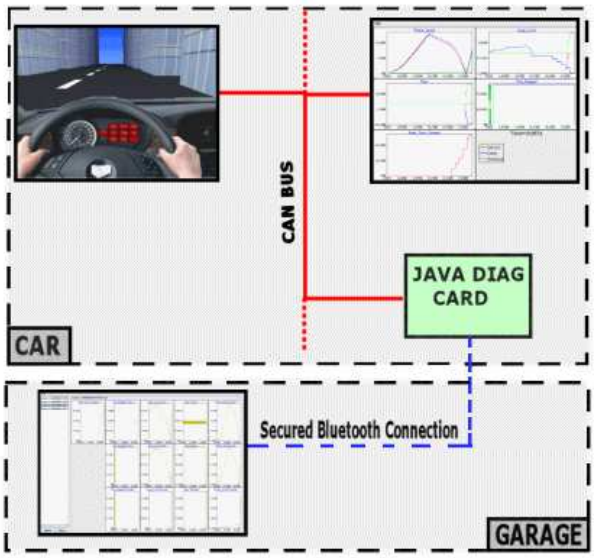
Apprentissage numérique pour la modélisation des Internautes. (2000-2003) Partenaires du projet :

- LIPN-CNRS : K. Benabdeslem, Y. Bennani, R. Kanawati, S. Salotti, F. Zehraoui
- NumSight : E. Janvier, F. Martin

Objectifs : La croissance accélérée du nombre d'utilisateurs d'Internet, pour la plupart novices, motive fortement le développement des approches d'aide adaptative pour faciliter l'accès à l'information.L'objectif de ce projet est le développement de systèmes d'apprentissage connexionnistes pour réaliser d'une part des tâches de classification visant à détecter des profils de comportement et d'autre part de reconnaissance dynamique basée sur l'évolution à court terme des traces de navigations des internautes afin d'anticiper le profil du navigateur ou de détecter l'émergence de comportements caractéristiques d'un profil déjà connu par le système.

- **Réalisations informatiques (logiciels) et Brevets**

Prototype Diagnostic embarqué (RESEDA) : Le démonstrateur de diagnostic embarqué a été réalisé dans le but d'évaluer, en temps réel, les méthodes de diagnostics mises au point par le LIPN dans le cadre du projet RESEDA. Le démonstrateur est constitué d'une unité centrale de diagnostic intégrant un microprocesseur Java et d'un module de communication sans fils "Bluetooth". La carte se connecte au système à diagnostiquer par un bus CAN.



Architecture du système de diagnostic embarqué

Pour la simulation, la carte peut se connecter à un PC, par une liaison CAN, et surveiller le comportement d'un simulateur développé sous Matlab/Simulink ou Java. Les informations de diagnostic peuvent, ensuite, être transmises à un système distant par une liaison Bluetooth sécurisée.

Auteur : S. Aseervatham

http ://www-lipn.univ-paris13.fr/ aseervatham/projects/cardiag.html

Plateforme Galicia : Les travaux sur les treillis de Galois emboîtés sont actuellement implémentés dans le cadre de la «plateforme Galicia»

Auteurs : M. Champesme et H. Soldano

(http ://www.iro.umontreal.ca/ galicia/)

RNA : Un programme de recherche de la structure secondaire commune à un ensemble de séquences d'ARN non alignées.

Auteurs : D. Bouthinon et H. Soldano

http ://www-lipn.univ-paris13.fr/ bouthinon/rnaHomedb/index.html

MACS (Modular Anticipatory Classifier System) : MACS est un Sytème de Classeur qui utilise l'apprentissage de modèles dans une architecture de type «Dyna», tout en offrant des capacités de généralisation.

Auteur : P. Gérard

http ://animatlab.lip6.fr/RechercheProjetXBotMACS

Method for processing Internet Request files : Brevet International, [N° de publication : WO 02/059781, date de publication : 01/08/2002]

Auteurs : K. Benabdeslem, Y. Bennani, E. Janvier

Un moteur de recherche sur le Web orienté communauté : (Système développé en projet de Master). Logiciel développé en Java, MySQL.

Auteur R. Kanawati

Un premier prototype du système CORBAS : Système de gestion et de recommandation de références bibliographiques. Système réalisé dans la cadre de la thèse de H.Karoui. Système dévelopé sur la plate-orme multi-agent Jade.

Auteur R. Kanawati

1.6.3 Diffusion de la recherche

A³

La vie scientifique de l'équipe A^3 est organisée autour de plusieurs manifestations (trimestriel en alternance : chaque mois une manifestation) visant à introduire une dynamique permanente entre les différents chercheurs sur la thématique de l'apprentissage. Citons quelques exemples :

- **Séminaire «Apprentissage de Paris-Nord» (APN)**

Responsables : P. Gérard et M. Lebbah (LIMBio).
La fréquence des interventions est tous les mois, les thèmes abordés portent essentiellement sur l'apprentissage et la fouille de données (http ://www-lipn.univ-paris13.fr/A3/Seminaire/).

- **Tutoriels A^3**

Responsable : A. Osmani
En alternance avec le séminaire, ces exposés sont d'intérêt général (sous forme de cours) et durent de 1h à 1h30 et sont suivis d'une discussion scientifique.

- **Atelier des doctorants A^3**

Responsables : D. Bouthinon et H. Moungla
En complément au « séminaire junior », cet atelier est organisé sous forme d'une table ronde par les doctorants qui s'intéressent de près ou de loin aux techniques d'apprentissage. Le but est de leur permettre de présenter leurs travaux de recherche, les problèmes liés à leur sujet de thèse, ou de partager avec les autres doctorants et les permanents leurs expériences.

- **Journées thématiques «Apprentissage Artificiel et Fouille de Données» (AAFD)**

Responsables : Y. Bennani et E. Viennet
Tous les deux ans, l'équipe A^3 organise les journées (AAFD). L'objectif de ces journées est de faire le point sur l'état de l'art des techniques et des applications utilisant l'apprentissage pour la fouille de données. Ces journées se présentent sous forme d'un forum de rencontres entre chercheurs et industriels oeuvrant dans le domaine de la fouille de données (http ://www-lipn.univ-paris13.fr/A3/AAFD06/).

- **Actions de vulgarisation**

S. Aseervatham, Y. Bennani, L. Dehni, et S. Guérif, ont présenté des démonstrations de plusieurs prototypes, lors de visites du LIPN par des universitaires, des membres du CNRS ou des industriels ou lors de manifestations telles que les journées « Science en

fête - Savante Banlieue » : systèmes de reconnaissance vocale, de reconnaissance visuelle, de détection et d'identification de perturbations sur le réseau téléphonique commuté, de diagnostic autonome et évolutif, économie d'énergie dans les réseaux de capteurs sans fil, systèmes de diagnostics pour automobiles.

1.6.4 Evaluation de la recherche

- **Organisation d'événements scientifiques, Comité d'organisation, de programme et de lecture**

  - Comité de programme de la conférence francophone Extraction et Gestion des Connaissances (EGC'08), Sophia Antipolis, 29 janvier- 1 février 2008 ; Y. Bennani
  - Comité de programme de EPNACS'07 - Emergent Properties in Natural and Artificial Complex Systems, Dresden, Germany October 4-5, 2007 ; Y. Bennani
  - Comité de programme de International Symposium on Signal Processing and its Applications (ISSPA) in conjunction with the International Conference on Information Sciences, Signal Processing and its Applications (ICISSPA) 12 - 15 February 2007, Sharjah, United Arab Emirates (U.A.E.) ; Y. Bennani
  - Comité de programme de Intenational Symposium on Programming and Systems, May 7-9, 2007, Algeria ; Y. Bennani
  - Comité de programme de European Conference on Machine Learning (ECML) 2007, Warsaw, Poland ; C. Rouveirol
  - Comité de programme de ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD) 2007, San José, California (U.S.A) ; C. Rouveirol
  - Comité de programme de IEEE International Conference on Data Mining series (ICDM) 2007, Omaha NE (USA) ; C. Rouveirol
  - Comité de programme de International Conference on Inductive Logic Programming (ILP) 2007, Corvalis, OR (U.S.A.) ; C. Rouveirol
  - Comité de programme du 6th Workshop on Multi-Relational Data Mining (MRDM) 2007, workshop attaché à ECML/PKDD 2007, Warsaw, Poland ; C. Rouveirol
  - Comité de programme du workshop 'Machine Learning in Systems Biology', 2007, Evry ; C. Rouveirol
  - Comité de programme de la conférence d'Apprentissage (CAp) 2007, plateforme AFIA 2007, Grenoble ; Y. Bennani, C. Rouveirol
  - Comité de programme de l'atelier Bases de Données Inductives, plateforme AFIA 2007, Grenoble ; C. Rouveirol
  - Comité de programme de l'atelier 'Réseaux d'Interactions, Analyse, Modélisation, Simulation' (RIAMS) 2007, attaché à la conférence 'Integrative Post Genomics', Lyon ; C. Rouveirol
  - Comité de programme de European Conference on Machine Learning and European Conference on Principles and Practice of Knowledge Discovery (co-location) (ECML/PKDD 2006) ; C. Rouveirol, E. Alphonse
  - Comité de programme de The Second International Workshop on Mining Complex Data - MCD'06 - in Conjunction with IEEE ICDM'06, Hong Kong, 18 - 22 December 2006, The IEEE International Conference on Data Mining (ICDM) ; Y. Bennani
  - Comité de programme de la 7ème édition de la conférence francophone Extraction et Gestion des Connaissances (EGC'07), Namur, Belgique, 23-26 janvier 2007 ; Y. Bennani
  - Comité de programme de l'atelier «Fouille de Données et Algorithmes Biomimétiques» dans le cadre de la conférence Extraction et Gestion de Connaissances (EGC'07) ; Y. Bennani

- Comité de programme de l'atelier «Visualisation de données» dans le cadre de la conférence Extraction et Gestion de Connaissances (EGC'07) ; Y. Bennani
- Comité de programme de SIGMAP is part of ICETE - The International Joint Conference on e-Business and Telecommunications Conference, 7-10 August 2006 Setúbal, Portugal ; Y. Bennani
- Comité de programme de la Conférence francophone sur l'apprentissage automatique CAp 2006, 22-24 mai 2006 - Trégastel ; Y. Bennani
- Comité de programme de FW-EGC06 « Fouille du Web » ; Y. Bennani
- Comité de programme de la conférence internationale WSOM'05, Paris ; Y. Bennani
- Comité de programme The Seventh International Symposium On Programming and Systems, ISPS'05, May 9-11, 2005 Algiers ; Y. Bennani
- Comité de programme de ALAMAS'2005 et ALAMAS'2006 ; H. Soldano
- Comité de programme de GRES 2006 ; Y. Bennani
- Comité de programme de CLA'06 ; H. Soldano
- Comité de programme : PSACE'06, ateliers française sur le RàPC (04-06) ; R. Kanawati
- Comité de programme de CoSSoM'06, Complex Systems and Self-organization Modelling, A workshop for ESM'2006, October 23-25, 2006, LAAS, Toulouse, France ; Y. Bennani et H. Kadri
- Comité de programme des 2es Journées thématiques «Apprentissage Artificiel et Fouille de Données» (AAFD'06), 27-28 avril 2006, Université Paris 13, Institut Galilée ; E. Alphonse, Y. Bennani, D. Bouthinon, M. Champesme, P. Gérard, R. Kanawati A. Osmani, H. Soldano, E. Viennet
- Comité d'organisation CBRSEQ : International Workshop on Applying CBR to Time series prediction). Prmière édition à Trondheim, 2003 et deuxième édition à Madrid Septembre 2004 ; R. Kanawati avec S. Salotti et M. Malek
- Comité d'organisation de l'atelier française sur le raisonnement à partir de cas. Villetaneuse, Mars 2004 ; R. Kanawati avec S. Salotti, F. Zehroui
- Comité d'organisation de l'atelier «Fouille de Données et Algorithmes Biomimétiques» dans le cadre de la conférence Extraction et Gestion de Connaissances (EGC'07) ; H. Azzag
- Comité d'organisation de GRES 2006 ; L. Dehni, Z. El Jrad, H. Moungla
- Comité d'organisation des 2es Journées thématiques «Apprentissage Artificiel et Fouille de Données» (AAFD'06), 27-28 avril 2006, Université Paris 13, Institut Galilée ; Y. Bennani et E. Viennet
- Comité de lecture du Numéro spécial de la Revue des Nouvelles Technologies de l'Information « Fouille du Web » ; Y. Bennani
- Comité de lecture du Numéro spécial de la Revue des Nouvelles Technologies de l'Information (AAFD'06) ; Y. Bennani et E. Viennet

- **Expertises**

- Expertise sur projets ANR (2005-2007) ; Y. Bennani
- Expertise sur projets ACI Masses de Données (2004) ; Y. Bennani
- Expertise sur projets ANR Blanc, Jeunes Chercheurs et Masses de Données (2007) ; C. Rouveirol
- Expertise sur dossiers masters (2007) ; Y. Bennani

- **Jurys de thèses et d'HDR depuis 2003 (hors Paris 13)**

- Jurys de thèse : Y. Bennani (25 dont 13 comme rapporteur), H. Soldano (1), R. Kanawati (2 à l'étranger dont 1 comme rapporteur), C. Rouveirol (2 dont 1 comme

rapporteur).
- Jury d'Habilitation à diriger des Recherches : Y. Bennani (4 dont 2 comme rapporteur).

1.6.5 Collaborations

- **Collaborations internationales**
  - University of Skopje (D. Cakmakov), Feature Selection for Knowledge Discovery and Data Mining ; Y. Bennani.
  - University of Skopje (D. Cakmakov, B. Jaksimovski), Combining and Optimizing Classifiers for Pattern Recognition Applications ; Y. Bennani.
  - Federal University of Ceará (J. Neuman de Souza, D. Gomes), Distributed Connectionist Learning for Wireless Sensor Networks ; Y. Bennani.
  - University of Torino, Machine Learning group, (Attilio Giordana) ; Y. Bennani, H. Azzag.
  - University of Torino, Machine Learning group, (Lorenza Saitta) ; Y. Bennani, C. Rouveirol.
  - Université Libre de Bruxelles, Centre Interdisciplinaire de Phénomènes Non-linéaires et de Systèmes Complexes, (Jean-Louis Deneubourg) ; H. Azzag, Y. Bennani.
  - Université Catholique de Leuven (Kurt Driessens) ; P. Gérard, Céline Rouveirol.
  - TU München - Fakultät für Informatik (Stefan Kramer) ; E. Alphonse et C. Rouveirol.

- **Collaborations industrielles (hors contrat)**
  - Convention Cifre avec la société Géobs (2003-2004, thèse de S. Guérif)
  - Convention Cifre avec la société NumSight (2004-2006, thèse de S. Guérif)
  - Société KXEN (E. Viennet, comité scientifique)

A³

1.7 Publications principales

Livres

[:2006fx] Younès Bennani, éditeur. *Apprentissage Connexionniste*. Hermès, 2006, ISBN 2-7462-1337-0.

Chapitres de livres

[RK07] Rushed Kanawati. *Agents RàPC coopératifs pour la recherche d'information*, Jean Renaud, Brigitte Chebel Morello, Béatrice Fuchs et Jean Lieber, éditeurs, *Raisonnement à partir de cas 2, surveillance, diagnostic et maintenance*, volume 2 de *Informatique et système d'information*, Chapitre 6, pages 177–195. Hermès, 2007.

[zehraoui07] Farida Zehraoui, Rushed Kanawati et Sylvie Salotti. *CASEP2 : Système Hybride pour le traitement de séquences de navigation sur le Web*, Jean Renaud, Brigitte Cheble-Morello, Bétarice Fuchs et Jean Lieber, éditeurs, *Raisonnement à Partir de Cas 2 : surveillance, diagnostic et maintenance*, volume 2 de *Informatique et système d'information*, pages 151–176. Hermès, 2007.

[Pisanti :2006li] Nadia Pisanti, Henry Soldano, Mathilde Carpentier et Joël Pothier. Implicit and explicit representation of approximated motifs. C. S. Iliopoulos, K. Park et K. Steinhoefel, éditeurs, *Algorithms in Bioinformatics*, pages 1–15. College Press, 2006.

Articles dans des revues internationales avec comité de lecture sélectif

[RI-2*] Erick Alphonse et Aomar Osmani. On the connection between the phase transition of the covering test and the learning success rate in ilp. *Machine Learning Journal*, 2007.

[Zehraoui :2005wy] Farida Zehraoui et Younès Bennani. New self-organising maps for multivariate sequences processing. *International Journal of Computational Intelligence and Applications*, 5(4) :439–456, 2005.

Articles dans des revues nationales ou internationales

[HK-RK-LP-SGAI-06] H. Karoui, R. Kanawati et L. Petrucci. An intelligent peer-to-peer multi-agent system for collaborative management of bibliographic databases. *SGAI's expert update magazine*, 8(3) :22–31, printemps 2006.

[Ventos :2005vp] Véronique Ventos et Henry Soldano. Treillis de Galois alpha. *Revue d'Intelligence Artificielle RSTI série RIA*, 19(4-5) :799–227, 2005.

Communications dans des conférences internationales avec comité de lecture

A³

[Azzag2007j] Hanane Azzag, Younès Bennani et Mustapha Lebbah. A stochastic algorithm for clustering inspired from cockroaches behavior. *V International Conference on Microelectronics and Computer Science*, 2007.

[cabanes :2007] Guénaël Cabanes et Younès Bennani. A simultaneous two-level clustering algorithm for automatic model selection. *(ICMLA'07), International Conference on Machine Learning and Applications*, Cincinnati, Ohio, USA, décembre 2007.

[Azzag2007a] Julien Lavergne, Hanane Azzag, Christiane Guinot et Gilles Venturini. Construction incrémentale et visualisation de graphes de voisinage par des fourmis artificielles. Monique Noirhomme-Fraiture et Gilles Venturini, éditeurs, *Extraction et gestion des connaissances (EGC'2007)*, volume I de *Revue des Nouvelles Technologies de l'Information (RNTI-E-9)*, pages 135–146, Namur, Belgique, Janvier 2007. Cépadués Éditions.

[Alphonse2006b] Erick Alphonse et Céline Rouveirol. Extension of the top-down data-driven strategy to ILP. *16th International Conference on Inductive Logic Programming (ILP 2006)*, /LNAI 4455/, pages /49–63/, 2006.

[Guerif :2006xz] Sébastien Guérif et Younès Bennani. Selection of clusters number and features subset during a two-levels clustering task. *Proceeding of the 10th International Conference Artificial intelligence and Soft Computing 2006*, pages 28–33, Palma de Mallorca, Spain, août 2006.

[karoui06b] Hager Karoui, Rushed Kanawati et Laure Petrucci. COBRAS : Cooperative CBR System for Bibliographical Reference Recommendation. *8th European Conference on Case-Based Reasonning (ECCBR'06)*, volume 4106 de *LNCS*, pages 76–90, Fethiye, Turkey, 2006. springer.

[MMaRK05] Maria Malek et Rushed Kanawati. A Data Driven CBR Approach : A Contineous Maintenance Strategy. *International Conference on Advances in Intelligent Systems : Theory and Applications*, page 6. IEEE, Novembre 2004.

[Zehraoui :2004qb] Farida Zehraoui et Younès Bennani. M-SOM-ART : Growing self organizing map for sequence clustering and classification. *ECAI'04, European Conference on Artificial Intelligence*, 2004.

[Bennaceur :2003mw] Hachemi Bennaceur et Aomar Osmani. Computing lower bound for MAX-CSP problems. *International Conference on Industrial, Engineering Applications of Artificial Intelligence and Expert Systems*, Lecture Notes in Artificial Intelligence, pages 614–624. Springer, 2003.

[kanawati03b] Rushed Kanawati. A Multi-agent System for Improving Result Ranking Service of Web Search Engines. *10th ISPE International conference on concurrent engineering (CE 2003)*, volume 1, pages 91–99, Funchal, Madeira, June 2003. Balkema.

[Zeboulon :2003kv] Arnaud Zeboulon, Younès Bennani et Khalid Benabdeslem. Hybrid connectionist approach for knowledge discovery from web navigation patterns. *ACS/IEEE International Conference on Computer Systems and Applications*, juillet 2003.

[kanawati07b] Rushed Kanawati et Maria malek. Computing Social Networks for Information Sharing : A Case-based Approach, volume = 4564, year = 2007. Schuler Douglas, éditeur, *Second international conference on Online Communities and Social Computing*, LNCS, pages 86–95, Beijing, China, July. Springer.

Workshops et autres conférences avec comité de lecture

[karoui06d] Hager Karoui, Rushed Kanawati et Laure Petrucci. Cooperative CBR system for peer agent committee formation. *Proc. 5th Workshop on Agents and Peer-to-Peer Computing (AP2PC'06)*, pages 76–90, Hakodate, Japan, mai 2006.

[Guerif :2005ya] Sébastien Guérif, Younès Bennani et Eric Janvier. μ-SOM : Weighting features during clustering. *Proceedings of the 5th Workshop On Self-Organizing Maps (WSOM'05)*, pages 397–404, Paris, France, septembre 2005.

[Guerra-Hernandez :2004xi] Alejandro Guerra-Hernandez, Amal El Fallah-Seghrouchni et Henry Soldano. Learning in BDI multi-agent systems. *Fourth International Workshop on Computational Logic in Multi-Agent Systems (CLIMA IV)*, volume 3259 de *Lecture Notes in Computer Science*, pages 218–233. Springer, 2004.

[Zehraoui :2004xc] Farida Zehraoui et Younès Bennani. SOM-ART : Incorporation des propriétés de plasticité et de stabilité dans une carte auto-organisatrice. *Atelier FDC : Fouille de Données Complexes dans un processus d'extraction de connaissances (EGC 2004)*, pages 169–180, janvier 2004.

A³

Chapitre 2
LCR (Logique, Calcul et Raisonnement)

Sommaire

2.1 Membres de l'équipe 70
2.1.1 Evolution de l'équipe (permanents) 73
2.2 Combinatoire algébrique 74
2.2.1 Thèmes de recherche 74
2.2.2 Pespectives de recherche 80
2.2.3 Diffusion de la recherche 80
2.2.4 Activités doctorales 82
2.2.5 Publications principales 83
2.3 Logique et Programmation 85
2.3.1 Thèmes de recherche 85
2.3.2 Perspectives de recherche 90
2.3.3 Diffusion de la recherche 91
2.3.4 Activités doctorales 95
2.3.5 Publications principales 96
2.4 Spécification et Vérification 98
2.4.1 Thèmes de recherche 98

2.4.2 Perspectives de recherche 102
2.4.3 Diffusion de la recherche 104
2.4.4 Activités doctorales 107
2.4.5 Publications principales 108
2.5 Bases et Entrepôts de Données 111
2.5.1 Thèmes de recherche 111
2.5.2 Perspectives de recherche 112
2.5.3 Séjours Scientifiques, Invitations, Tutoriels 112
2.5.4 Autres 112
2.6 Séminaires d'équipe 113

2.1 Membres de l'équipe

(*membres ou, en italique, anciens membres, au 01/01/2007*)

Permanents

LCR

	Nom	*Prénom*	*Situation*	*Institution*
	Baillot	Patrick	CR	CNRS
	Boudes	Pierre	MCF	IG
	Boufares	Faouzi	MCF	IUTV
	*Castaing*	*Jacqueline*	MCF	IG
	Choppy	Christine	PU	IG
	Coupey	Pascal	PU	IUTV
	Duchamp	Gérard	PU	IG
	Fouqueré	Christophe	PU	IG
	Jacobé de Naurois	Paulin	CR	CNRS
	Klai	Kaïs	MCF	IG
	*Lellahi*	*Kazem*	MCF	IG
	Loddo	Jean-Vincent	MCF	IUTV
	Mayero	Micaela	MCF	IUTV
	Mogbil	Virgile	MCF	IG

	Moyen	Jean-Yves	MCF	IG
	Petrucci	Laure	PU	IUTV
	Schwer	Sylviane	PU	IUTV
	Tollu	Christophe	MCF	IG
	Toumazet	Frédéric	MCF	IUTV
	Vauzeilles	Jacqueline	PU	IG

Doctorants

	Nom	*Prénom*	*Situation*	*Institution*
	Atassi	Vincent	AM	Université Paris 13
	Bertrand	Olivier	CDI	ONERA
	Falempe	Lionel	AM	Université Paris 13
	*Hamdoun*	*Sana*		Université Paris 13
	*Jaff*	*Luai*	B	Université Paris 13
	Kenniche	Hichem	AM	Université Paris 13

LCR

personnel temporaire du 1/9/2002 au 1/9/2007

Nom	*Prénom*	*Situation*	*Début*	*Fin*	*Financement*
Jacobé de Naurois	Paulin	ATER	09/2005	09/2006	ATER
Moyen	Jean-Yves	Post-doc	09/2005	09/2006	CNRS (STIC)
Dal Lago	Ugo	Post-doc	05/2006	11/2006	Université Paris 13
Popov	Todor	poste rouge	10/2004	01/2005	CNRS (STIC)
Pulcini	Gabriele	Post-doc	04/2007	03/2008	Région IdF
Mazza	Damiano	Post-doc	09/2006	08/2007	ANR

personnels invités du 1/9/2002 au 1/9/2006 (au moins 1 semaine)[1]

Prénom	*Nom*	*Origine*	*Financement*
2007			
Stefano	Guerrini	Univ. di Roma 1 (Italie)	professseur invité
Gianna	Reggio	Univ. di Genova (Italie)	professseur invité
Max	Kanovich	Queen Mary, Univ. of London (UK)	professseur invité
Kurt	Jensen	Univ. of Aarhus (Danemark)	professseur invité
Michele	Pagani	Univ. di Roma 3 (Italie)	ANR NOCOST
Martin	Hofman	Ludwig-Maximilians-Universität München (Allemagne)	professseur invité
2006			
Kazushige	Terui	NII, Tokyo (Japan)	ANR NOCOST
Cosimo	Laneve	Univ. di Bologna (Italie)	professseur invité
Gianna	Reggio	Univ. di Genova (Italie)	professseur invité

[1] professeur invité : 1 mois d'invitation par l'université Paris 13

Max	Kanovich	Queen Mary, Univ. of London (UK)	professseur invité
Paolo	Coppola	Univ. d'Udine (Italie)	MCF invité
Gleb	Koshevoy	Labo. Poncelet, Moscou (Russie)	professeur invité
Anne	Haxthausen	Technical Univ. of Denmark	professseur invité
Harry	Mairson	Brandeis University (USA)	Projet ANR No-Cost
Jonathan	Billington	Univ. of South Australia (Australie)	PAI FAST
Charles	Lakos	Univ. of Adelaide (Australie)	PAI FAST
			professeur invité
Ronald	King	Univ. of Southampton (UK)	professseur invité
2005			
Jonathan	Billington	Univ. of South Australia (Australie)	professseur invité
Charles	Lakos	Univ. of Adelaide (Australie)	PAI FAST
Kazushige	Terui	NII, Tokyo (Japan)	ACI CRISS
Harry	Mairson	Brandeis University (USA)	
Max	Kanovich	Queen Mary, Univ. of London (UK)	professseur invité
Ugo	Dal Lago	Univ. di Bologna (Italie)	ACI CRISS
Marco	Pedicini	CNR IAC Roma (Italie)	projet franco-italien (CNRS)
Gleb	Koshevoy	Labo. Poncelet, Moscou (Russie)	professeur invité
Gianna	Reggio	Univ. di Genova (Italie)	professseur invité
Anne	Haxthausen	Technical Univ. of Denmark	professseur invité
Maritta	Heisel	Univ. Duisburg-Essen (Allemagne)	
2004			
Simone	Martini	Univ. di Bologna (Italie)	professseur invité
Kazushige	Terui	NII, Tokyo (Japan)	ACI CRISS
Marco	Pedicini	CNR IAC, Rome (Italie)	IAC CNR-CNRS
Max	Kanovich	Univ. of Pennsylvania (USA)	professseur invité
Glynn	Winskel	Cambridge Univ. (UK)	professseur invité
Gianna	Reggio	Univ. di Genova (Italie)	professseur invité
Charles	Lakos	Univ. of Adelaide (Australie)	PAI FAST
Jonathan	Billington	Univ. of South Australia (Australie)	Projet ARC
Charles	Lakos	Univ. of Adelaide (Australie)	Projet ARC
Ronald	King	Univ. of Southampton (UK)	professseur invité
2003			
Marco	Pedicini	CNR IAC Roma (Italie)	Projet VINCI
Max	Kanovich	Univ. of Pennsylvania (USA)	professseur invité
Kazushige	Terui	NII, Tokyo (Japan)	AS Mobilité
Alexandre	Zamulin	Acad. des Sciences, Novosibirsk (Russie)	professseur invité
Maritta	Heisel	Univ. de Magdeburg (Allemagne)	professseur invité
Gianna	Reggio	Univ. di Genova (Italie)	professseur invité
Lorenzo	Tortora de Falco	Univ. Rome III (Italie)	Projet VINCI

2.1.1 Evolution de l'équipe (permanents)

Départs

- D. Béchet, MCF (thème logique), mutation à Nantes (2004)
- M. Masseron, MCF (thème logique), départ en retraite (2004)
- K. Lellahi, MCF (thème bases de données), départ en retraite (2006)
- J. Castaing, MCF (thème logique), départ en retraite (2007)

Arrivées

- J.-V. Loddo, MCF IUTV (thème logique), 2003 (thèse Paris 7)
- L. Petrucci, PU IUTV (thème spécification), 2003 (auparavant MCF Evry, LSV Cachan)
- M. Mayero, MCF IUTV (thème spécification), 2004 (thèse Paris 6)
- G. Duchamp, PU IG (thème combinatoire), 2004 (auparavant PU Rouen, LIFAR)
- P. Boudes, MCF IG (thème logique), 2005 (thèse Aix-Marseille 2)
- V. Jacobé de Naurois, CR CNRS (thème logique), 2006 (thèse INPL-Hong Kong)
- K. Klai, MCF IG (thème spécification), 2006 (thèse Paris 6)
- J.-Y. Moyen, MCF IG (thème logique), 2006 (thèse INPL)
- *L. Poinsot, MCF IG (thème combinatoire), 2007 (thèse Toulon)*

Après des départs pour convenances personnelles, l'équipe a conservé son potentiel de recherche dans la thématique logique en élargissant ses compétences vers la programmation fonctionnelle (J.-V. Loddo) et l'analyse de la complexité en septembre 2006 (P. Jacobé de Naurois et J.-Y. Moyen). Dans le même temps, elle a véritablement constitué une thématique spécification autour de C. Choppy par les arrivées de L. Petrucci comme professeur et M. Mayero et K. Klai comme MCF. Elle a enfin donné à la thématique combinatoire un professeur (G. Duchamp) qui travaille avec les deux MCF déjà présents, auquel s'ajoutera à la rentrée 2007 un MCF.

2.2 Composante Combinatoire

Participants : G. Duchamp, C. Tollu, F. Toumazet

2.2.1 Thèmes de recherche

L'activité du groupe de combinatoire au cours de la période 2003-2006 peut être répartie en deux pôles qui semblent disjoints et dont les multiples liens n'ont pas encore été exploités : un pôle «Combinatoire et Physique» (G. Duchamp, C. Tollu, F. Toumazet) et un pôle d'«Informatique Théorique» (G. Duchamp).

Les interactions entre combinatoire, informatique et physique ont été, dès la création de notre groupe de travail en 2001, au cœur de nos recherches. Les membres du groupe sont convaincus qu'un important effort en combinatoire se révèlera extrêmement profitable pour une approche algorithmique en physique.

Il est en effet devenu courant de constater que la solution de nombreux problèmes en physique et chimie (classique et quantique) est étroitement liée à la théorie des représentations des groupes et des algèbres, par exemple des algèbres de Lie, c'est-à-dire à la façon dont ces derniers agissent sur des espaces de dimension finie ou infinie. Encore faut-il, dans des applications concrètes, être en mesure de calculer des grandeurs pertinentes caractérisant ces représentations : éléments matriciels, décompositions en irréductibles de produits tensoriels de représentations, traces de représentations irréductibles (*i.e.* caractères), cœfficients de Clebsch-Gordan et de Racah.

LCR

C'est à ce niveau que la combinatoire joue un rôle clé, en permettant de décrire explicitement («combinatoirement») des actions de groupes et d'algèbres (de dimension finie ou infinie) qui satisfont des contraintes précises, physiquement pertinentes, ou en fournissant le paramétrage des objets de l'algèbre par des familles de structures combinatoires (diagrammes ou tableaux de Young, graphes, etc.) sujettes aux manipulations élémentaires, donc au calcul effectif. La théorie des fonctions symétriques (*i.e.* invariantes par permutation des variables) en offre une brillante et classique illustration : on travaille dans des bases orthogonales explicites[2] (par exemple les fonctions de Schur) paramétrées par des diagrammes de Young, et la combinatoire des tableaux de Young permet de mener des calculs sans de coûteux développements en monômes. C'est d'ailleurs sur l'utilisation de telles familles de polynômes (symétriques) orthogonaux que reposent les algorithmes calculant les objets (facteurs irréductibles, caractères, etc.) mentionnés à la fin du paragraphe précédent, quand ces algorithmes existent (essentiellement, dans le cas des groupes finis et des groupes de Lie compacts).

- **Mécanique des matrices infinies**

Cette activité, initiée par la combinatoire de l'ordre normal bosonique, est fortement inspirée par la «mécanique des matrices» de Werner Heisenberg. En effet, il est connu qu'il n'existe aucune représentation (fidèle) de l'algèbre de Weyl (même à deux dimensions) par des opérateurs bornés (donc aucune représentation de dimension finie). La première telle représentation est dans un espace de Fock de dimension infinie. C'est ce même espace de Fock qui est utilisé pour trouver l'ordre normal de puissances d'opérateurs homogènes, i.e. du type

$$\Omega = \sum_{w\in W_2,\ |w|_{a^+}-|w|_a=e} \alpha_w w \tag{2.1}$$

[2]entre lesquelles des formules effectives de changement de base existent.

où w est un produit d'opérateurs $a^+, a \in W$ et W est l'algèbre de Weyl de dimension 2.

On voit facilement que, selon que l'excès (le degré diraient les mathématiciens) $e = r - s$ est positif ou négatif, on a pour forme normale de Ω^n,

$$(a^+)^{ne}\left(\sum_{k=0}^{\infty} S_\Omega(n,k)(a^+)^k a^k\right) \text{ ou } \left(\sum_{k=0}^{\infty} S_\Omega(n,k)(a^+)^k a^k\right)(a)^{n|e|}. \tag{2.2}$$

L'étude des matrices infinies $S_\Omega(n,k)$ et de leurs implications combinatoires a occupé G. Duchamp et ses collaborateurs physiciens dans les travaux [RI-10, RI-39, CI-48, CI-97].

L'article [CI-48] fait le lien entre les représentations de Bargmann-Fock d'opérateurs homogènes à une annihilation et les champs de vecteurs sur la droite. On a en fait l'action d'un groupe de Lie fréchetique sur les fonctions développables en série de Taylor. Les sous-groupes à un paramètre de ce groupe étant déterminés par leur vecteur tangent (qui est une «fonction»), on obtient le calcul des cœfficients des matrices associées par une intégration élémentaire. On voit alors que ces matrices sont stables par substitution et multiplication (propriété dite «de Sheffer» par les combinatoristes [RI-40, CI-97]), ce qui donne un éclairage analytique à la composition des graphes (de type «Grafting» [RI-40, CI-93]).

LCR

- **Algèbres de Hopf, déformations**

Dans une publication relativement récente, Bender, Brody et Meister introduisent une théorie des champs spéciale décrite par une formule du produit (un produit de Hadamard sur les fonctions génératrices exponentielles) pour montrer que certaines suites de nombres peuvent être décrites par des règles appliquées à des types de diagrammes de Feynman. Ces diagrammes étiquettent des monômes en deux alphabets, que l'on peut nommer potentiel et action, lorsque les fonctions génératrices couplées ont pour terme constant 1. Les structures discrètes mises au point par l'informatique théorique (combinatoire des mots) qui ont été employées pour développer la théorie des Fonctions Symétriques Non-Commutatives (NCSF III et VI) réinterviennent ici et la théorie des déformations d'algèbres de Hopf met en correspondance, comme l'ont montré G. Duchamp et ses collaborateurs physiciens, l'algèbre de Hopf déformée des diagrammes de type Feynman étiquetés ($LDIAG(q_c, q_s, t)$, dans laquelle on garde la trace des superpositions q_s et des croisements q_c) avec les travaux récents de Connes et Kreimer par l'intermédiaire de l'algèbre des arbres planaires de Foissy.

D'autre part, le fait que ces algèbres sont libres a deux conséquences. La première est que l'on peut constuire une véritable déformation de Hopf (le paramètre t de $\mathbf{LDIAG}(q_c, q_s, t)$) entre **LDIAG** et **MQSym** (fonctions symétriques non commutatives matricielles). La deuxième est que leurs duaux de Sweedler (souvent considérés comme trop gros par les spécialistes et malcommodes à manipuler) sont calculés à l'aide d'un théorème de Kleene-Schützenberger modifié. Ce théorème donne la structure des fonctions reconnues par des automates à multiplicités [RI-14, RI-30, CI-80, CI-81] lorsque l'alphabet de commande est infini. Ceci est le véritable lien entre ces structures de pure combinatoire algébrique et la théorie des automates. Un article sur ce sujet, "Sweedler's dual and Schützenberger's calculus", par G. Duchamp et C. Tollu, est prêt à être soumis à l'automne 2007.

L'algèbre **LDIAG** et sa déformation $\mathbf{LDIAG}(q_c, q_s)$ peuvent aussi, comme toute une famille d'algèbres de Hopf dont les bases sont indexées aux matrices tassées d'entiers ou d'ensembles d'entiers, être réalisées sur des bimots. Cette réalisation combinatoire permet de mettre en évidence une structure de bigèbre bidendriforme sur **MQSym**, de laquelle découlent nombre de propriétés de structure importantes : liberté, co-liberté, structure de Lie des éléments primitifs [CI-154].

- **Multiplicités et combinatoire des polytopes**

C. Tollu et F. Toumazet étudient, en collaboration étroite avec R.C. King (université de Southampton, Royaume-Uni), le comportement, sous l'action de transformations combinatoires, de nombres qui apparaissent à de multiples reprises en algèbre (dimension de l'espace des vecteurs invariants du produit de trois représentations) en géométrie (calcul de Schubert) et en physique (états d'énergie résultant de l'interaction de plusieurs oscillateurs harmoniques).

En physique, les représentations irréductibles (polynômiales) de dimension finie du groupe général linéaire $GL_n(\mathbb{C})$ ou du groupe unitaire $U(n)$ jouent un rôle de premier plan. Il est connu de longue date qu'elles sont paramétrées par les partitions contenant au plus n parts, *i.e.* les suites d'entiers $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$. La combinatoire de ces représentations est intimement liée à des nombres entiers positifs ou nuls, appelés cœfficients de Littlewood-Richardson. Ces nombres correspondent à des multiplicités dans la décomposition en représentations irréductibles du produit tensoriel de deux représentations polynômiales ; partant, ils apparaissent aussi comme constantes structurelles de l'algèbre des fonctions symétriques si la base choisie est celle des fonctions de Schur, qui «sont" les caractères irréductibles de $GL_n(\mathbb{C})$. Les nombres de Littlewood-Richardson, notés $c^{\nu}_{\lambda,\mu}$, sont paramétrés par trois partitions λ, μ et ν ; dans de nombreux contextes, il est primordial, sinon de calculer exactement un (ou plusieurs) $c^{\nu}_{\lambda,\mu}$, du moins d'avoir des informations précises (nullité du cœfficient, encadrement précis de sa valeur). La règle de Littlewood-Richardson permet de calculer $c^{\nu}_{\lambda,\mu}$ pour tout triplet de partitions fixé, mais le coût de cet algorithme devient prohibitif dès que la taille et le poids des partitions deviennent grands.

LCR

Récemment, Knutson, Tao (et Woodward) ont fait des avancées spectaculaires[3] grâce à deux nouveaux modèles combinatoires pour les cœfficients de Littlewood-Richardson, celui des ruches (*hives*) et celui des nids d'abeilles (*honeycombs*), qui font clairement apparaître que les cœfficients de Littlewood-Richardson comptent les points à coordonnées entières dans des polytopes (ensembles convexes bornés de l'espace euclidien). Ils ont montré d'une part, que $c^{\nu}_{\lambda,\mu} = 0$ si et seulement si la dilatation $c^{N\nu}_{N\lambda,N\mu}$ (ici, chaque part de λ, μ et ν est multipliée par l'entier $N > 0$) est nulle (conjecture de saturation), d'autre part, que $c^{\nu}_{\lambda,\mu} = 1$ si et seulement si $c^{N\nu}_{N\lambda,N\mu} = 1$ (conjecture de Fulton-Kirillov). Dans le prolongement de ces travaux, King, Tollu et Toumazet ont étudié les relations entre $c^{\nu}_{\lambda,\mu}$ et $c^{N\nu}_{N\lambda,N\mu}$ quand le cœfficient de base $c^{\nu}_{\lambda,\mu} > 1$. Ils ont accumulé assez de données expérimentales[4] pour proposer et étayer une nouvelle conjecture, qui généralise tous les résultats des travaux décrits ci-dessus.

Conjecture 1 *[CI-39] Le nombre $c^{N\nu}_{N\lambda,N\mu}$ est un polynôme en N à cœfficients rationnels positifs dont le terme constant vaut 1 et dont le degré est donné par la dimension d'un polytope associé à $c^{\nu}_{\lambda,\mu}$.*

Dans [CI-39], cette conjecture est demontrée dans deux cas particuliers, (i) celui où la longueur des partitions spécifiant $c^{\nu}_{\lambda,\mu}$ est inférieure à 5, et (ii) celui où l'une des deux partitions λ, μ est une équerre.[5]

[3] Voir A. Knutson and T. Tao, «The honeycomb model of $GL_n(\mathbb{C})$ tensor products. I : Proof of the saturation conjecture", *J. Amer. Math. Soc.* **12** (1999) 1055-1090, et A. Knutson, T. Tao and C. Woodward, «The honeycomb model of $GL_n(\mathbb{C})$ tensor products. II : Puzzles determine facets of the Littlewood-Richardson cone", *J. Amer. Math. Soc.* **17** (2004) 19-48.

[4] Des calculs MAPLE et surtout SCHUR ont été effectués pour des valeurs de $N|\nu|$ de l'ordre de 500, ce qui représente plusieurs centaines d'heures de calcul. Auparavant, les calculs n'avaient été menés que pour des données à peu près dix fois plus petites.

[5] La polynômialité des $c^{N\nu}_{N\lambda,N\mu}$ a été depuis démontrée par E. Rassart.

Cette conjecture en entraîne une similaire portant sur d'autres nombres, les nombres de Kostka, $K_{\lambda\beta}$, qui donnent la multiplicité de l'espace de poids β dans la représentation irréductible paramétrée par la partition λ.

Conjecture 2 *[CI-39] Le nombre $K_{N\lambda,N\beta}$ est un polynôme en N à cœfficients rationnels positifs dont le terme constant vaut 1 et dont le degré est donné par la dimension d'un polytope associé à $K_{\lambda,\beta}$.*

En adaptant le modèle des ruches aux nombres de Kostka, King, Tollu et Toumazet ont aussi montré comment décomposer les polynômes $P_{\lambda\beta}(t) = K_{t\lambda,t\beta}$ en facteurs «primitifs» et déterminer une suite de facteurs linéaires $(t+m)$. Quant au degré de ces polynômes, ils ont obtenu une formule close et aisément calculable, qui donne un nouveau majorant, nettement plus bas que le majorant évident donné par la dimension de l'espace ambiant dans lequel le polytope est plongé [CI-71, RI-53]. Dans le cas où $P_{\lambda\beta}(t)$ est un polynôme primitif, ce majorant est toujours atteint[6].

Certains de ces résultats ont été généralisés aux cœfficients de Littlewood-Richardson (critère combinatoire pour la factorisation des polynômes $P^{\nu}_{\lambda\mu}(t) = c^{t\nu}_{t\lambda,t\mu}$, détermination d'une suite de facteurs linéaires), même si d'autres généralisations semblent aujourd'hui hors de portée [CI-39, RI-52]. Une publication sur ce sujet, "The factorisation of Littlewood-Richardson coefficients" par R.C. King, C. Tollu et F. Toumazet, est soumise à *J. Comb. Th. A.*

Tous ces résultats sont actuellement exploités pour le développement d'algorithmes distribués permettant de calculer les nombres $K_{\lambda\beta}$ et $c^{\nu}_{\lambda,\mu}$. Les performances des protocoles mis au point au LIPN se comparent favorablement à celles d'un sytème comme `LaTTe` (UC Davis, États-Unis) [CI-98]. Les possibilités de parallélisation sont prometteuses et méritent d'être mieux prises en compte, car il est vain d'espérer des algorithmes séquentiels efficaces pour des nombres $K_{\lambda\beta}$ et $c^{\nu}_{\lambda,\mu}$ associés à des partitions de grands poids : le calcul de ces nombres est $\#P$-complet[7].

LCR

- **Invariants et information quantique**

F. Toumazet, en collaboration avec J.-G. Luque et J.-Y. Thibon (Laboratoire d'informatique de l'Institut Gaspard Monge, université de Marne-la-Vallée) est engagé dans un projet de classification des états intriqués *via* le calcul d'invariants pour les hyperdéterminants.

L'intrication, propriété des systèmes quantiques, a été étudiée depuis longtemps, et remonte aux travaux de Schrödinger. Elle s'est récemment distinguée avec l'avènement de travaux théoriques sur le calcul quantique et d'autres phénomènes purement quantiques. Aujourd'hui, des mesures et des critères théoriques ne sont connus que pour les systèmes les plus simples, bien que des progrès aient été récemment accomplis dans cette direction.

En fait, il n'existe même aucune définition universellement acceptée pour les systèmes de plus de deux particules, et la recherche d'une bonne définition constitue un préalable à la détermination de mesures utilisables. En effet, l'intrication est l'une des caractéristiques fondamentales qui distinguent les systèmes quantiques du monde classique, et par là même une propriété physique importante à comprendre.

C'est aussi une ressource importante dans la théorie de l'information quantique, jouant un rôle vital dans la cryptographie quantique et la communication quantique, par exemple.

[6] Cela a été confirmé par T. McAllister, «Degrees of stretched Kostka coefficients", ArXiv :`math.RT/0603173`.

[7] Ce résultat a été publié dans H. Narayanan, «The computation of Kostka numbers and Littlewood-Richardson coefficients is $\#P$-complete», ArXiv :`math.CO/0501176`.

Dans le calcul quantique par états purs, il a été montré qu'une accélération exponentielle pour certains problèmes n'est possible qu'avec une intrication suffisante.

Toutefois, il reste de nombreuses questions ouvertes, concernant en particulier les systèmes à états mixtes.

Des travaux récents sur l'intrication d'unités quantiques de dimensions supérieures (qudits), particulièrement les qutrits, ont de plus montré que de tels systèmes peuvent exhiber de nouvelles propriétés telles que la non-localité quantique sans intrication, et de plus fortes violations des inégalités de Bell. Les implications de ces travaux ne sont pas encore complètement comprises.

Des méthodes basées sur la théorie des groupes peuvent être avantageusement utilisées dans ce contexte, et c'est dans cet esprit que ce travail est mené.

Ainsi, depuis sa découverte théorique, en 1935 par Einstein, Podolsky et Rosen l'intrication a fait l'objet de nombreuses investigations. La connaissance d'une classification des états intriqués devrait permettre de tirer profit des propriétés des systèmes quantiques de particules afin de construire des applications nouvelles. L'idée de recourir à la théorie de Mumford dans ce contexte a été proposée en Juin 2002 par Alexander Klyachko. Le but de l'opération est d'identifier les invariants pertinents susceptibles d'une définition uniforme, et d'évaluer l'utilité des critères accessibles à la théorie géométrique des invariants de Hilbert et Mumford, dont le but est de remédier à l'impossibilité des calculs explicites.

LCR

Luque, Thibon et Toumazet ont déjà obtenu des résultats sur les invariants des 4-qubits, 3-qutrits et 5-qubits ainsi que sur les invariants unitaires des systèmes purs de k-qubits (les invariants unitaires interviennent dans la construction de mesures d'intrication) [RI-49*].

- **Calcul formel pour la physique**

L'étude systématique des propriétés des groupes et algèbres de Lie, leurs applications à la chimie, la physique, les mathématiques, a conduit à l'élaboration de nombreux outils de calculs symboliques dans différents types de langage de calcul formel (bibliothèques de John Stembridge sous MAPLE, bibliothèque `combinat` sous MUPAD etc). Une autre alternative fut le développement d'environnements spécifiques tels LIE ou encore GAP. Ces outils, très efficaces, sont aussi très lourds à mettre en œuvre et tout en se spécialisant dans leurs domaines mathématiques, ils n'offrent pas toujours les fonctions de base liées à l'utilisation des fonctions symétriques, très utiles dans l'étude systématique des représentations polynômiales des groupes et algèbres de Lie. Dans ce cadre, F. Butelle (LIPN, équipe OCAD) et F. Toumazet ont décidé (avec R.C. King, Université de Southampton) de poursuivre le développement du programme SCHUR, important outil de calcul permettant la manipulation des objets combinatoires les plus usuels. Ce programme, initialement développé à la fin des années 70 par B. G. Wybourne et ses étudiants du département de physique de l'université de Christchurch (Nouvelle-Zélande), a été commercialisé dans les années 80. Devenu confidentiel et à usage privé dans les années 90 (un très petit cercle autour de son auteur), il a subi d'importantes phases de changement comme sa traduction en C et l'addition de nombreuses fonctionnalités, ajoutées année par année suivant les derniers développements scientifiques dans le domaine de la combinatoire algébrique.

Après le décès de B.G. Wybourne, ses amis physiciens et combinatoristes ont souhaité continuer à faire vivre le programme SCHUR. Le logiciel SCHUR vient d'être diffusé en GPL sur le site Sourceforge (http ://schur.sourceforge.net). Avant cette publication, le code C a été naturellement complètement remanié, rendu plus lisible et amendable par la communauté du logiciel libre. En outre, il est livré avec une importante documentation (environ 230 pages) et un ensemble d'exemples permettant de valider et de comprendre le fonctionnement du logiciel. Actuellement compilé pour Linux et Windows, une version

pour MacOS X est en cours de préparation. Ce programme, de par sa taille modeste, peut s'exécuter sur de nombreuses machines et ne nécessite pas de configuration très importante pour fonctionner (du moins pour les fonctionnalités de base). Avec plus de 200 fonctions, il permet outre les manipulations standards sur les fonctions symétriques, de nombreux calculs sur les représentations des groupes et algèbres de Lie de tout type ainsi que sur les représentation du groupe symétrique et dans une moindre mesure des algèbres de Hecke.

Ainsi, SCHUR est un outil puissant, et malgré tout léger, spécialement efficace pour examiner de nouvelles conjectures en théorie des représentation des groupes (compacts comme non compacts). Il est aussi utile pour chercher de nouvelles identités sur les fonctions symétriques et valider ou invalider des conjectures. Pour citer quelques exemples, mentionnons que SCHUR a permis les premiers calculs pour chercher à établir de nouvelles récurrences permettant le calcul des puissance des q déterminants de Vandermonde [RI-12], d'obtenir les critères précis concernant les produits génériques pour les groupes symplectiques réels (produits de représentations de dimension finie par des représentations de dimension infinie) [CI-92] ou encore de trouver les premiers polynômes associés à la dilatation des cœfficients de Littlewood-Richardson ou des nombres de Kostka [CI-98, CI-39, CI-71].

A ce sujet et pour terminer, à chaque prise en compte de nouveaux résultats, les algorithme sont amendés et il est prévu d'implémenter très prochainement le modèle des ruches de Knutson-Tao dans SCHUR, jusqu'à présent uniquement utilisé dans des programmes MAPLE, et de l'utiliser pour faire du calcul distribué.

LCR

- **Automates (à multiplicités)**

La théorie des automates à multiplicités telle qu'elle est exposée dans le livre de S. Eilenberg «Automata, Languages and Machines» est due essentiellement à Schützenberger et à son école. Il s'agit d'une véritable généralisation de la théorie classique puisqu'en variant les scalaires on peut obtenir les automates booléens, stochastiques (chaînes de Markov), plus courts chemins et transducteurs. Les scalaires les plus généraux doivent être pris dans un semi-anneau (généralisation des anneaux où on ne demande pas qu'il y ait une opération de «passage à l'opposé»). La structure de donnée adaptée au calcul dans les algèbres à cœoefficients dans les semi-anneaux est celle de table (avec les réécritures adéquates) [CI-80]. D'autre part, la structure de données adaptée au calcul rationnel est celle de classe d'expressions rationnelles selon des congruences assez grossières pour permettre des résultats de décidabilité [RI-14]. Enfin, les phénomènes de chemins infinis (cycliques) sont concentrés dans les problèmes d'étoile qui se formulent naturellement dans les automates quand on a rajouté une transition dont le symbole (ϵ) doit être supprimé dans l'étiquetage des chemins [RI-30]. Enfin, lorsque le semi-anneau des scalaires est un corps (de caractéristique nulle ou positive), on peut minimiser les automates grâce à des arbres couvrant leur graphe. Ces arbres sont indexés par des codes suffixes. Ceci permet de casser les fonctions booléennes de la cryptographie de façon compatible avec la permutation de leurs variables [CI-81].

- **Systèmes Complexes**

Sans pouvoir donner une définition formelle des Systèmes Complexes, on peut essayer de donner une idée de cette théorie fascinante grâce à l'étymologie, comme le fait, par exemple, l'article «Systèmes Complexes» de Wikipédia.

`http ://fr.wikipedia.org/wiki/Système_complexe`

«Étymologiquement, compliqué (du latin cum pliare, empiler avec) signifie qu'il faut du temps et du talent pour comprendre l'objet d'étude, complexe (du latin cum plexus,

attaché avec) signifie qu'il y a beaucoup d'intrications, que "tout est lié" et que l'on ne peut étudier une petite partie du système de façon isolée.»

La théorie des systèmes complexes a de multiples aspects. Le cadre de certains pans de celle-ci, comme la théorie du chaos et celle des systèmes dissipatifs en physique (et encore quand la dissipation est une fonction linéaire) sont relativement stabilisés. D'autre fragments de théorie restent à construire ; il s'agit ici de *science descriptive* allant de la phénoménologie pure ([CO-135], en collaboration avec l'Inserm) à la description de dynamiques, d'espaces de phases et de métriques sur ceux-ci [RI-54*, CO-1*] ou à la proposition de changements de point de vue [CI-74, CL-20].

Ici, G. Duchamp s'est employé dans trois directions : les algorithmes génétiques sur des populations d'automates à multiplicités [CI-112], les structures de données dynamiques [RI-54*] et les bases de donnes géographiques [CI-74, CL-20].

2.2.2 Pespectives de recherche

G. Duchamp intensifie sa collaboration avec les physiciens et poursuit ses travaux sur les déformations des algèbres de Hopf de diagrammes de Feynman. On espère en particulier des résultats nouveaux sur les liens qui existent entre algèbres de Hopf combinatoires qui apparaissent comme des spécialisations de **LDIAG**(q_c, q_s, t), leur combinatoire et leur algorithmique. Il espère aussi approfondir ses travaux sur les systèmes complexes et les connecter à ses autres axes de recherche.

LCR

C. Tollu et F. Toumazet poursuivent leur collaboration avec R.C. King sur le thème «Multiplicités et combinatoire des polytopes". Les méthodes qui ont permis les récents progrès sur la factorisation des cœfficients de Littlewood-Richardson vont être étendues à des multiplicités apparaissant dans la combinatoire des tours d'algèbres dont le diagramme de Bratteli (ou graphe de branchement) est proche du graphe de Young (diagramme de Hasse des partitions d'entiers ordonnées par l'inclusion). La combinatoire et l'analyse spectrale de ces graphes de branchement sera aussi un des thèmes de recherche d'Hayat Cheballah, qui commence une thèses de doctorat au LIPN, sous la direction de G. Duchamp et C. Tollu, en septembre 2007.

F. Toumazet se propose de continuer les investigations avec J.-G. Luque et J.-Y. Thibon sur le thème «Calcul d'invariants pour l'information quantique". En particulier, ils s'intéresseront aux liens existant entre la théorie des invariants des petits systèmes et la notion d'algèbre de Lie graduée. En effet, des groupes de réflexions complexes apparaissent dans l'étude de certains petits systèmes quantiques. Ce phénomène peut s'expliquer via la notion de groupe de Weyl des algèbres de Lie graduées découverte par Vinberg en 1975. Il est opportun d'appliquer cette théorie aux systèmes de 4-qubits.

Enfin, le développement du logiciel SCHUR occupera F. Toumazet (en collaboration avec F. Butelle).

2.2.3 Diffusion de la recherche

- **Evaluation de la recherche**

comités de programmes

- *Jordan International Conference on Computer Sciences and Engineering*, octobre 2004, Al-Balqa Applied University, Al-Salt, Jordanie ; G. Duchamp (International Advisory Committee)
- *EPNADS'2005, Emergent Properties in Natural and Artificial Dynamical Systems, workshop of the European Conference on Complex Systems*, http://complexsystems.lri.fr/, novembre 2005, Paris, France ; G. Duchamp (comité scientifique)

- *12th IEEE ICECS'2005 : International Conference on Electronics Circuits and Systems*, Gammarth, Tunisie, décembre 2005 ; G. Duchamp (comité de programme)
- *CoSSoM'06 Workshop, Complex Systems and Self-organization Modelling within ESM'2006*, octobre 2006, Toulouse, France ; G. Duchamp (comité de programme)
- *EPNACS'2005, Emergent Properties for Natural and Artificial Complex Systems, workshop of the European Conference on Complex Systems*, octobre 2007, Dresden, Allemagne ; G. Duchamp (comité d'organisation)
- *CoSSoM'07 Workshop, Complex Systems and Self-organization Modelling within ESM'2007*, octobre 2007, St Julian's, Malte ; G. Duchamp (comité d'organisation)
- *SSPCM'2007, Symmetry and Structural Properties of Condensed Matter*, Myczkowce, Pologne ; G. Duchamp (International Advisory Committee)

- **Organisation de la recherche**

organisation d'événements scientifiques et conférences **Conférences et workshops internationaux**

- (G. Duchamp et F. Toumazet) : *Complex Systems and Self-organization Modelling (ESM)*, LAAS, Toulouse, France, October 23-25th, 2006.
- (G. Duchamp) : *Modeling, Computation and Systems (IEEE-ICECS)*, Gammarth, Tunisie, December 11-14th, 2005.
- (G. Duchamp) : *EPNADS'05 Workshop (ECCS'O5)*, Paris, France, November 17th, 2005
- (G. Duchamp) : *Simulation, Evaluation and Graphics (JICCSE)*, Al-Balqa'Applied University, Al-Salt, Jordan, October 4-10th, 2004.

Conférences et workshops nationaux

- (G. Duchamp, C. Tollu, F. Toumazet) : IXe Journées du LIPN, "Combinatoire, informatique et physique", 26-27 mai 2003, Villetaneuse, 50 participants

LCR

Groupes de travail À la rentrée 2001, a été lancé un groupe de travail «Combinatoire, Informatique et Physique» tant il semblait important de fédérer les collaborations que nous pouvions avoir sur ces trois thèmes (qui s'interpénètrent) au Havre (Informatique), Marne-la-Vallée (Combinatoire), Jussieu (Physique). Le groupe de travail est devenu en 2002 le séminaire CIP (Combinatoire, Informatique & Physique, avec D. Barsky, C. Tollu et F. Toumazet) complémentaire des rencontres du CSCA (à Rouen), du Havre et le groupe de travail de Combinatoire Algébrique (à Marne-la-Vallée). Ce séminaire nous permet de tisser également des liens avec le LPTL (Laboratoire de Physique théorique des Liquides, Paris VI).
Notre but est de proposer un suivi des sujets de recherche en combinatoire (programmes, dynamique et interfaces), mais aussi d'améliorer la synergie de recherche "en réseau" qui existe déjà entre les Universités mentionnées plus haut. Nous tenons, comme ci-dessus à attirer l'attention du lecteur sur le fait que la division en deux paragraphes "Informatique" et "Combinatoire et Physique" ne doit pas faire perdre de vue l'interpénétration profonde de ces trois disciplines.
En outre, il s'avère depuis peu que l'Analyse Fonctionnelle (et donc sa Combinatoire : fonctions spéciales et représentations) peut tisser des liens avec la Logique par le biais de la *Géométrie de l'Interaction* (programme de recherche lancé par J.-Y. Girard, dont un des axes est l'interprétation de la logique dans des algèbres d'opérateurs). Nous avons mis en place, à cet effet, un groupe de travail, le GTH (Groupe de Travail "Hilbert") dont les ambitions sont de mélanger les niveaux de collaboration (exercices informels, exposés et rédaction de pans de théorie).

- **Dissémination**

G. Duchamp a donné une mini-conférence sur le thème «Combinatoire, informatique et physique : des liens anciens et étroits» pour la fête de la science en Ile-de-France (Savante banlieue, 13-14 octobre 2005).

- **Collaborations et Invitations**

Invitations à l'étranger

- G. Duchamp a été invité à faire des séjours de recherche à New-York, à Moscou, en Jordanie, à Cracovie (Institut de Physique Nucléaire).
- G. Duchamp et C. Tollu ont été invités a faire un séjour de recherche à l'université indépendante de Moscou (laboratoire Poncelet, UMI CNRS) en novembre 2006.
- G. Duchamp et C. Tollu ont été conférenciers invités au colloque "Combinatorics and Physics", organisé à l'Institut Max-Planck de Bonn en mars 2007.

Autres séjours à l'étranger

- F. Toumazet a séjourné plusieurs au Cameroun pour le lancement de collaborations de recherche avec différentes universités - Dschang, Douala et l'Ecole polytechnique de Yaoundé -, et la préparation d'un colloque informatique à Douala en mars/avril 2007 ; il a aussi fait un séjour de recherche à l'université de Southampton (Royaume-Uni) et un à l'université Nicolas Copernic de Torun (Pologne) ;
- C. Tollu a fait plusieurs séjours de recherche à l'université de Southampton.

LCR

2.2.4 Activités doctorales

Cours de DEA / Master 2 recherche

- 2005-2007 (G. Duchamp, C. Tollu) : "Calcul symbolique et fonctions spéciales" (24h), Master 2 recherche Mathématiques-Informatique de l'Institut Galilée.
- 2005-2007 (G. Duchamp) : "Calcul Formel et Systèmes Complexes", Master 2 Recherche MIASC (Mathématiques et Informatique Appliquées aux Systèmes Complexes) de l'Université du Havre.

Thèses en cours

Nom	Prénom	1ère insc	Situation		
		Titre		*Directeur*	*Co-encadrant*
NZEUTCHAP	Janvier	2005	ATER Rouen		
		Dualité de graphes gradués et correspondance de Robinson-Schensted généralisée		*F. Hivert (Rouen)*	*F. Toumazet*
CHEBALLAH	Hayat	2007	Allocataire Ministère		
		Combinatoire des graphes de transition		*G. Duchamp*	*C. Tollu*
KENNICHE	Hichem	2007	Allocataire Ministère		
		Combinatoire des mots circulaires		*C. Lavault*	*G. H. E. Duchamp*
GHNEMAT	Rawan	2007			
		Complex Systems Modeling for Geographic Information Systems		*C. Bertelle*	*G. H. E. Duchamp*

Jurys de thèses et d'HdR

- G. Duchamp a été examinateur de 3 thèses et de 1 Habilitation à Diriger des Recherches.

2.2.5 Publications principales

Chapitres de livres

[BD3] H. Kadri-Dahmani, C. Bertelle, G.H.E. Duchamp et A. Osmani. *Consistent Updating of Geographical DataBase as Emergent Property over Influence System*, M. A. Aziz-Alaoui et C. Bertelle, éditeurs, *Emergent Properties in Natural and Artificial Dynamical Systems*. Understanding Complex Systems. Springer, 2006, ISBN 978-3-540-34822-1.

Articles dans des revues internationales avec comité de lecture sélectif

[RI-10*] L. Jaff, G.H.E. Duchamp et C. Bertelle. Shift operators and complex systems. *Int. Journal of Modeling, Identification and Control*, 2006.

[KTT-Kostka-SLC] R. C. King, C. Tollu et F. Toumazet. The hive model and the factorisation of kostka coefficients. *Séminaire Lotharingien de Combinatoire*, 54A :22 pages, 2006.

[KTW-LR-SLC] R. C. King, C. Tollu et F. Toumazet. The hive model and the polynomial nature of stretched Littlewood-Richardson coefficients. *Séminaire Lotharingien de Combinatoire*, 54A :19 pages, 2006.

[RI-7*] J.-G. Luque, J.-Y. Thibon et F. Toumazet. Unitary invariants of qubit systems. *Mathematical Structures in Computer Science*, page 19 pages, 2006. À paraître.

[GOF8] P. Blasiak, A. Horzela, K. A. Penson, G. H. E. Duchamp et A.I. Solomon. Normal ordering via substitutions and Sheffer-type polynomials. *Phys. Lett. A*, 338 :108–, 2005.

[GOF7] P. Blasiak, K.A. Penson, A.I. Solomon, A. Horzela et G.H.E. Duchamp. Some useful combinatorial formulas for bosonic operators. *J. Math. Phys.*, 46 :052110–, 2005.

LCR

[DHKL] G. Duchamp, Hatem Hadj Kacem et Eric Laugerotte. Algebraic elimination of ϵ-transitions. *Discrete Mathematics and Theoretical Computer Science*, 7(1) :51–70, 2005.

[DC1] G. Duchamp et J.-M. Champarnaud. Derivatives of rational expressions and related theorems. *Theoretical Computer Science*, 313(1) :31–44, 2004.

[KTW-vdM] R. C. King, F. Toumazet et B. G. Wybourne. The square of the Vandermonde determinant and its q-generalisation. *J. Phys. A : Math. Gen.*, 37 :735–767, 2004.

[GOF1] K.A. Penson, P. Blasiak, G. Duchamp, A. Horzela et A.I. Solomon. Hierarchical Dobiński-type relations via substitution and the moment problem. *J. Phys. A : Math. Gen.*, 37 :3457–, 2004.

Communications dans des conférences internationales avec comité de lecture

[DLNTT-fpsac] G.H.E.D. Duchamp, J.-G. Luque, J.-C. Novelli, C. Tollu et F. Toumazet. Hopf algebras of diagrams. *Proceedings of FPSAC'07*, Tianjin, China, 2007.

[BD1] R. Ghnemat, K. Khatatneh, S. Oqeili, C. Bertelle et G.H.E. Duchamp. Automata-based adaptive behavior for economic modeling using game theory. M.A. Aziz-Alaoui et C. Bertelle, éditeurs, *Emergent Properties in Natural and Artificial Dynamical Systems*, Understanding Complex Systems series, pages 173–185, Gammarth, Tunisia, 2006. Springer.

[NTB-Torun] J. Nzeutchap, F. Toumazet et F. Butelle. Kostka numbers and Littlewood-Richardson coefficients : distributed computation. R.C. King, M. Bylicki et J. Karwowski, éditeurs, *Symmetry, Spectroscopy and SCHUR, Proc. Prof. Brian G. Wybourne Commemorative Meeting*, Torun, Poland, 2006. Nicolaus Copernicus University Press.

[GOF9] K. A. Penson, P. Blasiak, G. Dattoli, G. H. E. Duchamp, A. Horzela et A. I. Solomon. Monomiality principle, Sheffer-type polynomials and the normal ordering principle. *Journal of Physics : Conference Series, SSPCM'05*, volume 30, pages 86–97, Myczkowce, Poland, 2006. IoP.

[GOF2] A.I. Solomon, P. Blasiak, G. Duchamp, A. Horzela et K.A. Penson. Combinatorial physics, normal order and model Feynman graphs. B. Gruber, G. Marmo et N. Yoshinaga, éditeurs, *Proceedings of the Symposium Symmetries in Science XIII*, pages 150–166, Bregenz, Austria, 2006. Kluwer Academic Publishers.

[Tou-Torun] F. Toumazet. Characterization of generic products in $Sp(2n, R)$ and $SO^*(2n)$. R.C. King, M. Bylicki et J. Karwowski, éditeurs, *Symmetry, Spectroscopy and SCHUR, Proc. Prof. Brian G. Wybourne Commemorative Meeting*, Torun, Poland, 2006. Nicolaus Copernicus University Press.

[BFCA05] G. H. E. Duchamp, H. Hadj-Kacem et E. Laugerotte. On the decomposition of boolean functions. J.-F. Michon, P. Valarcher et J.-B. Yunès, éditeurs, *Proceedings of BFCA'05 Conference, March 7–8, 2005 Rouen, France*, pages 115–128. Universités de Rouen et du Havre, 2005.

[DKB] G. H.E. Duchamp, K. Khatatneh et C. Bertelle. Rewriting tables and memorized semirings. Noureddine Boudriga et Mohammad S. Obaidat, éditeurs, *Proc. of the 12th IEEE International Conference on Electronics, Circuits and Systems*, Gammarth, Tunisia, 2005. IoP.

[KBDO] H. Kadri-Dahmani, C. Bertelle, G. H.E. Duchamp et A. Osmani. Evolutive complex scheduling in interaction networks for quality improvment in geographical data base updating. Noureddine Boudriga et Mohammad S. Obaidat, éditeurs, *Proc. of the 12th IEEE International Conference on Electronics, Circuits and Systems*, Gammarth, Tunisia, 2005. IoP.

[KTT-fpsac] R. C. King, C. Tollu et F. Toumazet. The hive model and the factorisation of Kostka coefficients. *Proceedings of FPSAC'05*, pages 721–732, Taormina, Italy, 2005.

[GOF4] G. Duchamp, K.A. Penson, A.I. Solomon, A. Horzela et P. Blasiak. One-parameter groups and combinatorial physics. J. Govaerts, M. N. Hounkonnou et A. Z. Msezane, éditeurs, *Proceedings of the Third International Workshop on Contemporary Problems in Mathematical Physics (COPROMAPH3)*, pages 436–, Porto-Novo, Benin, 2004. World Scientific Publishing.

[KTT-crm] R.C. King, C. Tollu et F. Toumazet. Stretched Littlewood-Richardson and Kostka coefficients. P. Winternitz, J. Harnad, C.S. Lam et J. Patera, éditeurs, *Symmetry in physics : in Memory of Robert T. Sharp*, volume 34 de *CRM Proceedings and Lecture Notes*, pages 99–112. Amer. Math. Soc., 2004.

LCR

Workshops et autres conférences avec comité de lecture

[BaD] S. Baudic et G. H. E. Duchamp. Modelling and reasoning on emotions : theories and clinical implications. *Complex Systems and Self-organization Modelling, CoSSOM'06 (Toulouse)*, 2006.

[CO-1*] L. Jaff, Hatem Hadj-Kacem, G. H.E. Duchamp et C. Bertelle. Moderate growth time series for dynamic combinatorics modelisation. submitted to the *International Journal of Electronic Commerce*.

2.3 Composante Logique et Programmation

Participants : P. Baillot, P. Boudes, J. Castaing, P. Coupey, C. Fouqueré, J.-V. Loddo, V. Mogbil, J. Vauzeilles, V. Atassi, L. Falempe, P. Jacobé de Naurois, J.-Y. Moyen, U. Dal Lago.

2.3.1 Thèmes de recherche

La logique linéaire (LL), introduite par Jean-Yves Girard en 1985, est un raffinement à la fois des logiques classique et intuitionniste, qui a permis en particulier de donner un statut logique aux opérations de duplication et d'effacement d'hypothèses. Depuis son introduction, la logique linéaire a eu un rôle sans cesse croissant dans l'étude et la compréhension de divers problèmes d'informatique : aspects opérationnels des langages de programmation et de leurs stratégies, sémantique des langages séquentiels et concurrents, complexité calculatoire.

Nos travaux portent sur l'utilisation de la LL dans les deux principaux paradigmes de calcul logique : la *correspondance preuves-programmes* d'une part, avec ses applications à la programmation fonctionnelle ; le calcul par *recherche de preuve*, ou *programmation logique*, d'autre part. Dans le premier cas une preuve formelle peut être vue comme une dérivation de type d'un programme ; dans le second elle correspond à l'exécution de celui-ci.

LCR

La LL a permis la définition d'outils puissants pour ces deux approches, comme les *réseaux de preuves*, une représentation des démonstrations sous forme de graphes. Elle permet aussi de distinguer plusieurs familles de connecteurs logiques : *additifs*, *multiplicatifs* et *exponentiels*, avec des comportements spécifiques en termes d'usage des ressources.

Dans le contexte des deux paradigmes de programmation mentionnés, trois domaines applicatifs ont été particulièrement étudiés dans le cadre de l'ACI GEOCAL (2003-06) : la LL comme système de types permettant de garantir des bornes de complexité sur les programmes ;les structures de preuves comme solution de problèmes de planification d'Intelligence Artificielle ; la LL comme base de la programmation logique où les clauses sont définies comme réseaux de preuve partiels.

Nous avons aussi débuté un projet de spécification et développement d'un langage de programmation adaptée pour le web. Un site web est vu comme composé de parties à contenu évolutif à typage statique, permettant ainsi au programmeur de concevoir des sites Web hautement interactifs proposant une plus grande souplesse et simplicité, ainsi qu'une plus grande fiabilité.

- **Logique linéaire et complexité implicite**

Contexte. L'analyse logique de la réutilisation de ressources dans la LL a permis de définir des variantes de ce système correspondant à différentes classes de complexité calculatoire : les LL Light (LLL, Girard 1995) et Soft (SLL, Lafont 2002) pour le temps polynomial (Ptime) et ELL pour le temps élémentaire. Ces travaux s'inscrivent dans le domaine de la Complexité Implicite, qui vise à definir des calculs ou des langages où tous les programmes ont par construction une borne de complexité donnée (par exemple Ptime).

Objectifs. Nous cherchons par le biais de la Logique linéaire et de la Complexité implicite à mieux comprendre le calcul à ressources bornées, et en particulier en temps polynomial. D'une part nous souhaitons en dériver des méthodes pour vérifier statiquement des bornes de complexité sur les programmes fonctionnels. Il est en particulier nécessaire de déterminer l'efficacité de ces méthodes et leur couverture en termes d'expressivité intensionnelle (algorithmique). D'autre part, plus fondamentalement nous étudions par ce

biais les classes de complexité de problèmes (notamment grâce aux circuits booléens) et la sémantique des programmes à complexité bornée.

- **Typage et garantie de complexité en lambda-calcul :**
  Nous avons poursuivi l'étude de l'utilisation des logiques LLL et ELL comme systèmes de types pour le lambda-calcul, garantissant respectivement des bornes polynomiales et élémentaires sur l'exécution des programmes (via les réseaux de preuves). Nous avons ainsi montré que l'inférence de type en LLL pour le lambda-calcul était décidable [RI-21]. En collaboration avec Terui (NII, Tokyo) nous avons ensuite défini un sous-sytème de cette logique, appelé DLAL, offrant de meilleures propriétés vis-à-vis du lambda-calcul : les programmes typés sont en effet Ptime pour la réduction du lambda-calcul, et non seulement pour la réduction des réseaux [CI-59]. Par ailleurs nous avons donné des algorithmes d'inférence de type efficaces (Ptime) dans ELL et DLAL pour le lambda-calcul [CI-88, CI-131, RI-81]. Enfin, nous avons analysé la *réduction optimale* (au sens des travaux de Lévy, Lamping) des programmes du lambda calcul typé dans LLL (et ELL, DLAL ...). Celle-ci définit des stratégies d'exécution des programmes minimisant en un certain sens le travail d'évaluation et décrites par des règles de réécriture de graphes. Nous avons démontré que pour le lambda-calcul typé dans LLL, la réduction optimale est de complexité en temps polynomiale [CI-169]. Ce résultat a été obtenu en utilisant des outils de *géométrie de l'interaction* venant des réseaux de preuves.

LCR

- **Complexité implicite et expressivité :** Comme les systèmes LLL et SLL permettent de représenter toutes les *fonctions* Ptime mais pas tous les *algorithmes*, il est intéressant, d'un point de vue pratique en particulier, de considérer des extensions de ces logiques. Nous avons ainsi étudié l'extension de SLL avec des types récursifs et défini un lambda-calcul étendu spécifique pour ce langage ([CI-58]). D'autre part avec Dal Lago (Bologne) nous avons montré que LLL sans second-ordre mais avec types récursifs était complet pour la classe Ptime [RI-57].
  Par ailleurs nous nous sommes intéressés au modèle de Blum-Shub-Smale (BSS) qui permet de représenter le calcul sur les réels ous sur d'autres anneaux. Nous avons donné une nouvelle mesure d'espace dans ce modèle, permettant de définir des classes LOGSPACE et PSPACE ([CI-119]). Avec Pedicini (Rome) nous avons montré qu'une extension du lambda-calcul typé dans LLL permettait de caractériser la classe BSS Ptime sur un anneau arbitraire [CO-137].
  Nous avons aussi considéré une autre approche de la complexité implicite, basée sur les systèmes de réécriture, les *quasi-interprétations* de Bonfante-Marion-Moyen, et étudié son expressivité algorithmique en donnant des conditions nécessaires sur les algorithmes admettant des quasi-interprétations [CO-138]. Enfin, nous avons introduit la notion de *graphe de contrôle de re ssources* comme un cadre général permettant de regrouper différentes analyses d e terminaison de programme et de complexité.
- **Complexité et réseaux de preuves :** Initiée par Terui et Mairson (2003-04), une autre approche logique de la complexité consiste à représenter une fonction non pas par une seule preuve (comme dans LLL) mais par une famille uniforme de preuves, où chacune correspond à une taille de données fixée, à la manière des circuits booléens. Ces derniers sont un modèle de calcul parallèle usuel en théorie de la complexité. Dans cette perspective, il est suffisant de considérer un petit fragment de la LL, sans duplication : la logique linéaire multiplicative (MLL). En passant à la syntaxe des réseaux de preuves qui est intrinsèquement parallèle, Terui a donné une traduction des circuits booléens dans les réseaux de preuves et vice-versa.
  En y ajoutant une notion d'uniformité pour les réseaux de preuves préservée par cette approche nouvelle, nous caractérisons dans les réseaux de preuves, la classe de complexité NC : ce sont les fonctions effectivement parallélisables [CI-141]. Nous

avons aussi étendu l'approche initiale à d'autres familles de circuits et à d'autres fragments usuels de la LL préservant l'accélération due au parallélisme du modèle de calcul. Notamment la classe de circuits NNC, ajoutant du non déterminisme à NC, est montrée équivalente à la classe de réseaux de preuves booléens de la logique linéaire multiplicative et additive (MALL). Nous caractérisons ainsi la classe NP usuelle dans les réseaux de preuves MALL [CI-141].

– **Sémantique en complexité implicite :**
Enfin, nous étudions les invariants dénotationnels des logiques linéaires à complexité implicite afin d'éclairer leur structure mathématique. À cette fin, nous avons introduit les espaces cohérents stratifiés [RI-22] et nous considérons actuellement les contraintes capturant la complexité implicite au sein de sémantiques de la logique linéaire.

Ces travaux ont été menés dans le cadre de l'ACI CRISS (2003-06) et depuis décembre 2005 du projet No-Cost (projet ANR Jeunes chercheuses et jeunes chercheurs).

- **Logique linéaire et planification**

LCR

Dans plusieurs articles précédents nous avons montré que la logique linéaire permettait de modéliser de façon adéquate le problème typique d'Intelligence Artificielle de réaliser un plan d'actions à accomplir par un robot pour atteindre un état final, en partant d'un état initial donné. A cette fin, nous avons introduit la logique linaire de Horn et montré que contrairement à l'indécidabilité de la logique linéaire propositionnelle de Horn, le problème de planification est décidable pour une large classe de systèmes de robots.

Nous avons en particulier prouvé que le problème de planification était EXPTIME-complet pour les systèmes de robots permettant des actions non déterministes et que cette complexité pouvait être réduite à PSPACE pour les systèmes de robots dont les actions sont uniquement déterministes.

Au cours de la période 2003-2006, nous avons poursuivi cette étude de la planification dans deux directions :

– *La planification concernant des domaines contenant un grand nombre d'objets identiques.*
Nous avons montré que dans ce cas, la logique linéaire permettait une réduction drastique de l'espace de recherche (polynomial au lieu d'exponentiel). Pour cela, nous remplaçons le nombre non borné de noms spécifiques d'objets par un nom générique et nous contractons ainsi l'espace de recherche exponentiel sur les objets 'réels' en un espace de recherche polynomial sur un objet 'générique' en exhibant une formulation plus abstraite et en prouvant que la solution peut être traduite directement en des solutions polynomiales (et optimales) du problème initial [CI-12, RI-73].
– *La planification temporelle.*
L'hypothèse faite dans les précédents travaux était que les effets des actions étaient instantanés. Dans ce cadre, les plans et les stratégies associées sont finies, si on suppose qu'il n'y a qu'un nombre fini d'états. Nous avons ensuite étudié la planification pour des actions avec effets différés : même dans le cas d'un nombre fini d'états, la difficulté rencontrée est qu'alors les stratégies sont de branchement infini car le nombre de délais possibles est infini, puisque le temps est supposé être le temps réel. Nous avons montré que, même dans ce cadre plus complexe, la correspondance entre les preuves en Logique Linéaire et les plans était maintenue [Ra-20].

- **Programmation par modules**

La programmation logique a connu de nouveaux développements depuis quinze ans, tant dans la compréhension profonde du mécanisme de calcul par le biais de l'étude de la normalisation des preuves, que dans l'intégration de la concurrence dans le calcul par une formalisation en théorie des graphes. Ainsi, Andreoli a pu montré que les propriétés d'associativité et de commutativité des connecteurs alliées à la propriété de *focalisation* induisent une polarisation des formules, une stratification des formules, des réseaux de preuves (cette notion fondamentale est également considérée dans les travaux de Girard sur la ludique ou de Laurent sur la logique linéaire polarisée). Le fait de pouvoir normaliser les preuves par *polarisation* en LL permet de considérer celle-ci comme un langage de programmation (la décomposition d'une formule polarisée correspond à l'application d'une clause en programmation en logique classique). En notant que la présentation par réseaux de preuves de la LL est une déséquentialisation de celles-ci, il est possible de rendre concurrente la programmation sous réserve de vérification d'un critère de correction garantissant la validité. En d'autres termes, la programmation logique usuelle est basée sur une réduction étape-par-étape de buts à prouver au moyen d'une résolution ou d'une règle de progression, la correction d'un calcul est alors réduite à un appariement entre une certaine partie du but courant et la tête d'une clause choisie (appariement structurel dans le cas où la LL complète est considérée). A contrario, l'approche par réseaux de preuves initiée par Andreoli est un cadre naturel de modélisation de la programmation concurrente.

LCR

Nous avons étudié cette approche d'abord dans le cadre simple multiplicatif puis dans un cadre avec exponentiels en adoptant un point de vue de théorie des graphes. Clauses de notre système, les *modules*, sont des réseaux de preuve partiels. Tenant compte du fait que la LL est polarisée, nos objets de base sont alors des structures de preuve avec deux strates, appelés *structures bipolaires* ou *modules élémentaires*. Un programme logique est alors donné par un ensemble de telles structures bipolaires et son exécution résulte de la composition de ces structures, moyennant la satisfaction d'un critère de correction.

Dans le cadre multiplicatif, nous avons étudié les modules fermés. Un module fermé correspond à un calcul terminé. Cette étude jouet nou s a permis de préciser les mécanismes de correction que nous avons généralisés ensuite. Le mécanisme de correction que nous utilisons est une forme polarisée du critère de Danos-Regnier. Le critère de Danos-Regnier est basé sur les propriétés graphiques des réseaux de preuve : les structures de preuve correctes sont dans un certain sens connexes et acycliques. Nous en avons déduit un critère de correction par réécriture pour les modules fermés. Nous avons montré que ce critère peut être prolongé aux modules bipolaires ouverts (correspondant à des exécutions partielles de programmes logiques) en remplaçant le critère de connexité par un critère de *connectabilité*. Nous avons enfin étudié comment il était possible de vérifier incrémentalement ces critères[Ra-12].

Nous avons récemment entrepris l'extension de cette étude au cas exponentiel, le cas le plus simple étant de considérer que les modules bipolaires sont simplement réutilisables. Enfin nous avons généralisé le langage de module au cas général (exponentiels autorisés dans les modules) en donnant le critère de correction correspondant[So-2].

- **Programmation et web**

Les sites Web étaient il y a quelques années composés essentiellement de pages statiques, mais de plus en plus, ces pages sont générées automatiquement par le serveur en fonction des requêtes du client et du contenu de bases de données. Récemment est apparue une troisième génération de sites, proposant une interactivité encore plus grande en utilisant des modifications dynamiques calculées côté client (Javascript , AJAX). Les langages traditionnellement utilisés pour programmer un site Web, par exemple PHP,

Zope ou JSP, ne sont pas vraiment adaptés à ces nouvelles exigences des développeurs Web. Même lorsque le comportement dynamique est limité au côté serveur, des sites complexes deviennent vite très difficiles à développer et maintenir dans ces langages. Nous pensons que les concepts issus de la recherche théorique sur les langages de programmation peuvent permettre de résoudre ces problèmes.

Nous nous sommes fixés comme objectif de définir une façon radicalement nouvelle de programmer des sites Web hautement interactifs proposant une plus grande souplesse et simplicité pour le programmeur, ainsi qu'une plus grande fiabilité. Notre première idée consiste à se départir du modèle de programmation traditionnel des sites page après page pour considérer au contraire un site dans sa globalité et rapprocher sa conception de celle d'une application traditionnelle. Il faut cependant prendre en compte des spécificités telles que la possibilité de revenir en arrière dans l'exécution du programme (bouton « back » des navigateurs) ou d'avoir plusieurs fils d'exécution simultanés (duplication d'une fenêtre). Ce sont des problèmes difficiles, souvent mal résolus dans les sites actuels. Le concept de « continuation » utilisé en programmation fonctionnelle permet de résoudre ce problème de manière très élégante et puissante. Le typage nous permettra en outre d'avoir un contrôle précis des liens et formulaires garantissant l'absence de liens morts et l'adéquation entre un formulaire et la page cible. Enfin ce style de programmation nous permettra d'avoir une gestion totalement transparente des sessions, sans que le programmeur n'ait à se préoccuper du passage et de la récupération des paramètres de session à chaque page. Le typage nous permettra également de résoudre un point important qui fait défaut dans les outils actuels : la génération de pages conformes aux standards. Le Web repose sur des standards édictés par le W3C, qui garantissent l'interopérabilité et l'accès universel aux informations. Aujourd'hui il est impossible de garantir la validité d'une page générée dynamiquement puisque les outils de validation ne peuvent prendre en compte que des pages statiques. Nous proposons de nous appuyer sur des travaux théoriques sur le typage de documents XML, qui nous permettrons de garantir la validité des pages grâce à une analyse statique du programme.

Cette étude se fait en coordination avec une entreprise italienne (Net7, Pise) spécialisée dans la conception de sites web avec laquelle un projet Marie Curie - Transfer of Knowledge a débuté en 2006, en collaboration avec une équipe de PPS (projet ANR jeune chercheur débutant fin 2006). Ce projet fait suite au projet HyperLearning (http://www.hypernietsche.org) intégrant des équipes en sciences humaines et en informatique pour lequel nous collaborons dans le cadre d'un GDRE.

- **De la représentation et du Raisonnement Temporel**

S. Schwer a montré que les relations temporelles qualitatives sont en correspondance naturelle avec une variante d'un problème d'analyse combinatoire standard : celui des combinaisons. Une façon élégante de décrire ces objets est celle des mots écrits sur un alphabet de second ordre, *i.e.* ensemble de parties d'un alphabet de base. L'alphabet de base décrit les identités des objets temporels manipulés, ses parties représentant des *types d'instants* au sens de points de synchronisation. C'est ainsi qu'à partir d'une représentation mathématique non temporelle, nous avons restitué une certaine temporalité à ce problème : la temporalité événementielle. Cette représentation permet de se dégager du cadre binaire dans lequel les relations temporelles sont traitées par les différentes logiques utilisées, de dénombrer et d'énumérer les relations qualitatives temporelles n-aires, et de fournir une représentation spatiale de chacune de ces relations dans un espace de dimension n, c'est-à-dire le nombre d'éléments en relation et non en fonction de la taille des éléments à l'intérieur d'une relation binaire (comme le fait le modèle de Ligozat). De plus, cette spatialisation conserve un sens temporel, puisqu'elle associe à une chronologie, une promenade sans perdre aucune des propriétés mises en lumière par les modèles précé-

dents. J.-M. Autebert et S. Schwer ont prouvé que l'ensemble de ces relations forme une classe d'équivalence d'un système de réécriture et que cette classe, pour l'ordre induit par la dérivation, possède une structure de treillis distributif dans le cas binaire, non modulaire (donc non distributif) dès que $n \geq 3$. Dans le cadre d'un nouveau groupe de travail, en collaboration avec une spécialiste du permutoidre (treillis des permutations) et des groupes de Coxeter de l'EHESS (Claude Barbut), S. Schwer a continué, dans le cadre de sa délégation au Lallic, à étudier les propriétés des treillis associés aux S-langages. Ce travail dépasse largement le cadre temporel. Il a permis en particulier de fournir une application non triviale des chemins de Delannoy, bien connus des spécialistes de parcours dans les treillis discrets et d'enrichir l'encyclopédie des séquences entières de Sloane d'une nouvelle suite et de plusieurs références directes aux articles de Delannoy. Les applications abondent, en particulier dans le domaine des séquences audiovisuelles, en bases de données temporelles, dans les systèmes décisionnels, en planification, en diagnostic, dans l'étude des séquences génétiques ...

Le raisonnement temporel est une classe particulière de problèmes de sa tisfaction de contraintes. Il se traduit en termes d'opérations algébriques classiques sur les S-langages (intersection, shuffle). Un problème de satisfaction de contraintes temporelles qualitatives admet une solution si et seulement si le langage temporel associé est non vide. Chacune de ces solutions est donnée par un mot qui est intrinsèquement une chronologie. Les calculs sont de type fortement paresseux afin de juguler, autant que faire se peut, l'explosion combinatoire inhérente à ce genre de problème. Nous avons montré, en particulier, que dans le cas de contraintes "convexes", nous avons la même complexité de calcul que par celui des réseaux.

LCR

Si le formalisme des S-mots permet de représenter les modèles temporels discrets comme ceux de Allen, Ladkin, Ligozat ou des modèles linguistiques ponctuels (de Bauzée-/Reichenbach, de Marc Wilmet et Co Vet), intervallaires (de Moeschler et Laurent Gosselin), ... nous nous proposons de l'appliquer dans le cadre des interactions géniques. Une S-lettre est un état du cycle de vie dans le développement d'un organisme. Cette représentation permettra de reconstituer le réseau d'interaction.

- **Histoire des Sciences : Henri-Auguste Delannoy**

S. Schwer s'est attachée depuis quelques années à l'étude de l'oeuvre de Delannoy[RI-7]. Alors que nous connaissons les personnes associées aux suites portant leur nom, comme Eugène Catalan (1814–1894), Axel Thue (1863–1922), Ernst Schröder (1841–1902), Walther von Dyck (1856–1934), Jan Łukasiewicz (1878–1956), Eric Temple Bell (1883–1960) ou bien Theodore Motzkin (1908–1970), aucun article de ces trente dernières années ne cite un article de cet auteur. Les seules références renvoient soit au livre de combinatoire de Louis Comtet soit à un article de Errera, qui eux mêmes renvoient au livre de théorie des nombres de Lucas. C'est à partir de ces observations que s'est créé un groupe de recherche sur Delannoy, qui s'est d'abord intéressé à la vie propre de Delannoy, puis à la communauté des combinatoriciens et théoriciens des nombres entre 1870 et 1914.

2.3.2 Perspectives de recherche

Nous comptons poursuivre nos efforts de recherche dans les directions actuelles : logique linéaire et applications (complexité implicite, planification, programmation logique), programmation et web, S-mots. En particulier, le programme de recherche portant sur la complexité implicite reste encore largement ouvert. Les problèmes d'expressivité ne sont pas totalement résolus et il conviendra de mieux savoir faire collaborer des approches différentes (approche par les quasi-interprétations par exemple) ou des cadres différents (circuits, complexité sur les réels). Le domaine de la programmation adaptée au web

ouvre à des problématiques importantes (calcul concurrent avec données réparties, programmation avec données semi-structurées, programmation par partie).

2.3.3 Diffusion de la recherche

- **Evaluation de la recherche**

comités de programmes

- Workshop *Lambda Calculus, Type Theory and Natural Language*, septembre 2005, King's College London ; J. Vauzeilles
- Workshop *Logic and Linguistics* de la session *Geometry of Computation*, Marseille, France ; J. Vauzeilles
- *Journées internationales organisées pour le 60ème anniversaire de Jean-Yves Girard*, septembre 2007, Institut Henri Poincaré, Paris ; J. Vauzeilles
- GEOCAL, *Geometry of Computation*, janvier - février 2006, Marseille, France ; P. Baillot, C. Fouqueré, J. Vauzeilles
- SETIT2005, *Sciences of Electronic, Technologies of Information and Telecommunications*, mars 2005, Sousse, Tunisie ; C. Fouqueré
- ACFAS, *colloque "Description linguistique pour le traitement automatique du français"*, mai 2006, Montréal, Canada ; C. Fouqueré
- FG-MOL2005, *10th conference on Formal Grammar and 9th Meeting on Mathematics of Language*, août 2005, Edimbourg, Ecosse ; C. Fouqueré
- MOL8, *8th Meeting on Mathematics of Language*, juin 2003, Bloomington, Indiana, USA ; C. Fouqueré

LCR

comités d'évaluation

- (membre nommé) ANR Programme *Masses de données - connaissances ambiantes*, 2005 et 2006 ; C. Fouqueré
- (expert) *Evaluation nationale de la recherche (VTR)* organisée par le comité italien pour l'évaluation de la recherche (CIVR), Ministère des Universités et de la Recherche Italien, depuis décembre 2004 ; J. Vauzeilles
- (membre élue) *Conseil National des Universités* (CNU), 2003 à 2007 ; J. Vauzeilles
- (expert) *Evaluation de dossiers de master, PAI, ACI*, Mission Scientifique Technique et Pédagogique du ministère de l'éducation nationale ; J. Vauzeilles
- (expert) ANR, 2005 et 2006 ; J. Vauzeilles
- (membre extérieur) *Jury d'admissibilité du recrutement CR2 de l'Inria Lorraine*, 2003 et 2005 ; J. Vauzeilles

autres

- Jurys de thèse : P. Baillot (1), C. Fouqueré (4 dont 1 Italie), J. Vauzeilles (4 dont 1 Italie).
- Jury d'Habilitation à diriger des Recherches : J. Vauzeilles (2)

- **Organisation de la recherche**

organisation d'événements scientifiques et conférences

- *Implicit Computational Complexity and Logic*, organisé par P. Baillot et V. Mogbil, septembre 2004, université Paris 13 (ACI GEOCAL), 30 participants, dont 8 étrangers.
- *Workshop on Implicit Computational Complexity*, organisé par P. Baillot et J.-Y. Marion, 13-17/2/2006, (GEOCAL, *Geometry of Computation*, 2006), Marseille. 20 exposés et 50 participants (de 7 pays).

gestion de groupes de travail

– *Complexité implicite*, depuis sept. 2005 (environ 18 séances par an), P. Baillot.

- **Contrats**

Type	Coordinateur	Resp. scient.	Nom	Financeur	Période	Durée (mois)	Budget global (Keuros)	Part LIPN (Keuros)
MC-TOK	•	C. Fouqueré	WEBSICOLA	UE	05/2006-04/2010	48	809	404
ANR		J.-V. Loddo	Prog. fonc. et web	MENESR	09/2006-09/2010	48	153	53
COST		J.-V. Loddo	Hyperl	UE	03/2006-03/2010	48	200	_
GDRE		C. Fouqueré	HyperLearning	CNRS et UE	06/2005-06/2010	48	340	_
ACI		P. Baillot	GEOCAL	MENESR	09/2003-09/2006	36	100	12
ACI		P. Baillot	CRISS	MENESR	07/2003-07/2006	36	110	27
ANR	•	P. Baillot	No-Cost	MENESR	09/2005-09/2008	36	150	140
AS		P. Baillot	Mobilité	CNRS	09/2002-09/2003	12	35	7

LCR

WEBSICOLA *Autre partenaire : Net7 (Pise, Italie)*

WEBSICOLA - *WEB SIte COnception LAnguage* - est un projet européen du 6ème PCRDT obtenu dans le cadre des actions Marie Curie - Transfer of Knowledge. L'objectif majeur consiste à développer un langage de programmation fonctionnel permettant d'assurer par le typage statique une fiabilité des sites web. La recherche effectuée au LIPN porte sur la spécification de ce langage, la mise en oeuvre du développement étant effectuée par l'entreprise italienne Net7, spécialisée dans la programmation et la programmation web. En parallèle, ce projet permet un transfert de compétences entre les deux partenaires.

HyperLearning *Autres partenaires : 30 laboratoires européens*

HyperLearning est un projet ayant pour cadre un GDRE avec un financement du CNRS et de l'union européenne (projet COST) dont l'objectif principal est de créer un système de e-learning avancé en Sciences Humaines. Le projet a quatre composants principaux : 1) recherche en programmation fonctionnelle pour le développement de sites web interactifs, 2) développement d'une plate-forme web distribuée, 3) mise en place de communautés d'apprentissage collaboratif virtuel, 4) création d'un cadre pédagogique approprié. Notre recherche s'effectue dans le premier de ces objectifs : développement d'un langage de programmation adapté aux services web nécessaires pour le cadre particulier des Sciences Humaines (tâche Functional Programming for Complex Interactive Web Systems).

GEOCAL *Autres partenaires : PPS (Univ. Paris 7), IML (Univ. Aix-Marseille II, coord.), LIFM (Univ. Aix-Marseille I), LIP (ENS Lyon), projet* MIMOSA *(Inria-Sophia), projet* CALLIGRAMME *(Inria-Loria), Inria-Futurs, I3M (Univ. Montpellier II), LSV (ENS Cachan), IRMA (Univ. Louis Pasteur Strasbourg)*

GEOCAL - *Géométrie du calcul* (ACI Nouvelles Interfaces des Mathématiques)

Le but du projet GEOCAL est de favoriser la communication entre des équipes de mathématiciens (logiciens et topologues algébristes) et d'informaticiens, partageant des intérêts

convergents dans le domaine de l'étude théorique du calcul. Il vise à promouvoir de nouvelles interfaces entre logique et topologie d'une part, et théorie de la concurrence, de la complexité algorithmique et de la programmation logique de l'autre. Il comprend cinq volets : Sémantique ; Calcul de processus et logique ; Parallélisme et topologie algébrique ; Complexité algorithmique et élimination des coupures ; Programmation logique.

CRISS *Autres partenaires : LIFM (Univ. Aix-Marseille I), projet* MIMOSA *(Inria-Sophia), projet* CALLIGRAMME *(Inria-Loria)*

CRISS - *Contrôle de ressources et d'interférence pour systèmes synchrones* (ACI Nouvelles Interfaces des Mathématiques)

Le projet CRISS se focalise sur les questions de sécurité soulevées par le concept de code mobile que l'on envisage aujourd'hui d'utiliser dans des domaines aussi variés que les réseaux programmables, les jeux en réseau et les cartes à puces. Les premières propositions – les "applets" de Java – ont visé à assurer l'intégrité de l'environnement d'exécution. Nous nous intéressons à deux autres classes de propriétés qui apparaîssent naturellement dans le contexte du code mobile :

- La dérivation de bornes sur les ressources nécessaires à l'exécution du code afin d'éviter des attaques de type déni de service.
- Le contrôle du flux d'information afin d'éviter la fuite de données confidentielles (non-interférence).

L'objectif du projet est de développer des méthodes automatiques pour assurer et certifier ces propriétés et ceci dans le cadre d'un langage de coordination de modules séquentiels qui permet une exécution synchrone et déterministe. Les recherches menées au LIPN portent plus particulièrement sur l'inférence automatique de bornes de complexité pour les programmes fonctionnels et l'utilisation de la logique linéaire.

LCR

No-Cost *Autre partenaire : PPS (Univ. Paris 7)*

No-Cost : Nouveaux Outils pour la Complexité : Sémantique et Types (ANR programme Jeunes chercheuses et jeunes chercheurs 2005)

La complexité implicite étudie des calculs et des langages dont les programmes ont une complexité bornée, en terme de temps de calcul ou de mémoire requise. Plusieurs approches ont été utilisées pour garantir de telles bornes, notamment la Logique linéaire et ses variantes, les systèmes de réécriture avec quasi-interprétations, les types linéaires avec ressources pour les langages fonctionnels. Cette théorie nécessite désormais que l'on trouve des éléments unificateurs, soit par des systèmes plus généraux soit par des approches plus sémantiques. Le projet NO-CoST cherche dans cette perspective à (i) développer une approche sémantique unificatrice pour les divers systèmes de complexité implicite à travers la sémantique des jeux ; (ii) étendre les résultats concernant les systèmes connus, notamment la logique linéaire ; (iii) enfin établir de nouveaux systèmes plus généraux.

Mobilité *Autre partenaire : LIENS (ENS Paris), LIF (Marseille), LIP (ENS Lyon), PPS (Paris 7)*

Il s'agit d'une AS Action Spécifique CNRS "Méthodes formelles pour la mobilité") (dans le cadre du Réseau Thématique Prioritaire RTP23 ; action spécifique 92 de responsable R. Amadio). Cette action spécifique a porté sur le développement et l'expérimentation de méthodes formelles pour la programmation et la validation d'applications basées sur la notion de code mobile et sur des standards d'échange de données semi-structurées. L'AS s'est articulée en trois volets : Fondements logiques (sémantiques de jeux ; typage, logique linéaire et complexité) ; Code mobile ; Données semi-structurées.

Programmation fonctionnelle pour le web *Autre partenaire : PPS (Univ. Paris 7)*

Le projet *Programmation fonctionnelle pour le web* a été accepté par l'appel d'offres 2006 de l'ANR pour le programme Jeunes chercheuses et jeunes chercheurs. Ce projet complète le projet Marie Curie (Websicola) en cherchant à développer des techniques de compilation vers Javascript qui permettront de programmer un site entier sans sortir d'un langage fonctionnel typé. Des techniques inspirées des langages à plusieurs niveaux pourront être utilisées pour spécifier les parties de code devant être exécutées sur le client. Un intérêt majeur de cette approche est de permettre un contrôle statique du code client qui permettra de s'assurer qu'il interagit correctement avec le serveur. Afin de gérer au mieux les aspects concurrents, tout particulièrement au niveau du client (interactions avec l'utilisateur et communication asynchrone avec le serveur), nous pensons fournir des primitives avancées de concurrence et des mécanismes de communications de haut-niveau et dont la sûreté est garantie par typage. Ce projet s'appuie sur des travaux antérieurs ou en cours : XDuce, un langage de manipulation de données arborescentes XML, Ocsigen, un serveur Web qui permet de déléguer la génération des pages à des fonctions Objective Caml, et Websicola, le langage expérimental destiné à la programmation de sites et de services Web, non implémenté mais doté d'une sémantique proche du pi-calcul, où les interactions browser-serveur sont modélisées par un échange de noms privés entre processus concurrents.

LCR

- **Dissémination**

actions de vulgarisation

- P. Baillot, V. Mogbil : Fête de la Science/ Savante Banlieue, université Paris 13, octobre 2003 et 2004 : Poster ("Raisonner en informatique").

invitations, tutoriels, école d'été

- P. Baillot : exposé invité au '6th International Workshop on Logic and Computational Complexity (LCC'04) (satellite de LICS'04), "Type Inference for Light Affine Logic Via Constraints on Words", 12/7/04, Turku, Finlande.
- P. Baillot : tutoriel (*Linear logic, lambda-calculus and polynomial time complexity*) dans le cadre d'un cours "Complexity and Logic" de l'Ecole d'hivers de la session GEOCAL06 (Geometry of Computation), CIRM Marseille, février 2006.
- P. Baillot : tutoriel (1 séance) dans le cadre du cours *Implicit Computational Complexity* (responsable : Simone Martini), Bertinoro International Spring School for Graduate Studies in Computer Science, Italie, mars 2006.
- P. Baillot : tutoriel "On the Linear Logic Approach to Implicit Computational Complexity" (avec U. Dal Lago), Workshop on Linear Logic deidcated to J.-Y. Girard on his 60th Birthday, Siena, Italie, 17-20 mai 2007.
- P. Baillot : exposé invité à *8th international Conference on Typed Lambda Calculi and Applications* (TLCA'07), "From Proof-Nets to Linear Logic Type Systems for Polynomial Time Computing", 26-28 juin 2007, Paris.
- C. Fouqueré a été invité à donner un tutoriel sur les calculs de processus et la mobilité à la 2ème école franco-mexicaine sur les systèmes répartis coopératifs, Oct 2003
- C. Fouqueré : "Logic programming revisited", GEOCAL06 session, Marseille, 1 mars 2006 ; et Workshop Geometria, Roma 3, 28-30 avril 2006.
- J. Vauzeilles : Final Meeting of the Project PROTOCOLLO, 14-17 Septembre 2004, Turin
- J. Vauzeilles : Workshop La geometria del calcolo, 28-29 avril 2006, Rome
- J. Vauzeilles : First International Symmetry Conference, 14-17 janvier 2007, Edimburgh.

- **Collaborations et Invitations**

Collaborations universitaires

- 2002 : programme Vinci France-Italie (Université Franco-Italienne) d'échange de chercheurs. Partenaire : Université Rome 3, Italie.
- 2004-2007 : Accord bilatéral Socrates/Erasmus avec l'université Rome 3, Italie.
- 2004-2007 : coopération CNRS-CNR France-Italie (Direction des relations internationales du CNRS), échange de chercheurs. Titre : *Interaction et complexité*. Partenaire : IAC CNR Rome, Italie.
- 2006- : GDR "Informatique Mathématique" : groupe de travail "Géométrie du calcul"
- 2006-2008 : Projet franco-italien : "Réseau italo-français, logique et géométrie de la computation", Ministères de la recherche français et Italien

Collaborations industrielles

- projet Marie Curie - Transfer of Knowledge avec Net7 (Pise, Italie), 2006-2010.

Invitations à l'étranger

- P. Baillot a été invité à l'Université de Keio, Tokyo, Japon, 14-21 mars 2006 (1 semaine), par M. Okada, et au NII, Tokyo, Japon, 14-21 avril 2005 (1 semaine), par K. Terui.

Autres séjours à l'étranger

- 2006 P. Baillot : Université de Padoue, Italie (5 semaines)
- 2005 P. Baillot : National Institute of Informatics, Tokyo, Japon (2 semaines).
- 2005 P. Baillot : Université de Padoue, Italie (8 semaines).
- 2005 V. Mogbil : IAC CNR Rome, Italie (1 semaine)
- 2004 V. Mogbil : Université de Rome 3, Italie (1 semaine)
- 2004 P. Baillot : Université de Padoue, Italie (6 semaines).
- 2004 P. Baillot : IAC CNR Rome, Italie (1 semaine).
- 2003 P. Baillot : Université de Cambridge (UK), Computer Lab (3 semaines).

2.3.4 Activités doctorales

- **cours de DEA et stages / Master 2 recherche**

Cours de DEA / Master 2 recherche

- 2002-2005 (C. Fouqueré) : introduction aux logiques de description, DEA IA-OC Paris 8 - Paris 13
- 2001-2003 (P. Baillot, C. Fouqueré, J. Vauzeilles) : cours de théorie de la démonstration (option), DEA logique et fondements de l'informatique, Paris 7
- 2005-2006 (C. Fouqueré, V. Mogbil, J. Vauzeilles) : cours de Programmes et Preuves (option), master 2 recherche Informatique, Paris 13
- 2003-04 (P. Baillot, avec O. Laurent) : cours de théorie de la démonstration (option), DEA logique et fondements de l'informatique, Paris 7
- 2004-2007 (P. Baillot) : participation au cours *Logique linéaire et paradigmes logiques du calcul* de R. Di Cosmo et D. Miller (3h/an en 2004-05, 9h/an depuis 2005), Master Parisien de Recherche en Informatique (P7-ENS Paris-ENS Cachan-X).
- 2005-2007 (C. Fouqueré) : cours de théorie de la démonstration, DEA Mathématique - Informatique, Antananarivo (Madagascar)

Stages de DEA / Master 2 recherche

- 2003 (P. Baillot) : D. de Carvalho (DEA de logique université P7), *Logique linéaire, complexité polynomiale et théorie des ensembles naive.*
- 2004 (C. Fouqueré) : L. Falempe (DEA de logique université P7), *Langage de Programmation Quantique : un prototype simple*
- 2005 (P. Baillot, J. Vauzeilles) : V. Atassi (MPRI, Master STL université P6), *Inférence de type en logique linéaire élémentaire.*
- 2006 (V. Mogbil) : V. Rahli (MPRI), *Réseaux de preuves et circuits booléens*

2.3.5 Publications principales

Livres

[MeFoSyLoMa-06] S. Haddad, F. Kordon et L. Petrucci, éditeurs. *Méthodes formelles pour les systèmes répartis et coopératifs*. Hermès, octobre 2006, ISBN 2-7462-1447-4.

Articles dans des revues internationales avec comité de lecture sélectif

[ABT07] V. Atassi, P. Baillot et K. Terui. Verification of ptime reducibility for system f terms : Type inference in dual light affine logic. *Logical Methods in Computer Science*, 2007. à paraître. 32 pages.

[KanVau-TCS] Max Kanovich et Jacqueline Vauzeilles. Strong planning under uncertainty in domains with numerous but identical elements (a generic approach). *Theoretical Computer Science*, 379(1–2) :84–119, 2007.

[DLB06] U. Dal Lago et P. Baillot. On Light Logics, Uniform Encodings and Polynomial Time. *Mathematical Structures in Computer Science*, 16(4) :713–733, 2006. (une version préliminaire avait été présentée au *Workshop on Logics for Resources, Processes, and Programs* LRPP'04).

[B04a] P. Baillot. Stratified coherence spaces : a denotational semantics for Light Linear Logic. *Theoretical Computer Science*, 318(1-2) :29–55, 2004.

[B04b] P. Baillot. Type inference for Light Affine Logic via constraints on words. *Theoretical Computer Science*, 328(3) :289–323, 2004.

[AutebertSchwer03a] Jean-Michel Autebert et Sylviane R. Schwer. On generalized Delannoy paths. *SIAM Journal on Discrete Mathematics*, 16(2) :208–223, 2003.

Communications dans des conférences internationales avec comité de lecture

[BaillotCoppolaDalLago07] P. Baillot, P. Coppola et U. Dal Lago. Light Logics and Optimal Reduction : Completeness and Complexity. *Proceedings of Symposium on Logic in Computer Science (LICS'07)*, pages 421–430. IEEE Computer Society, 2007.

[MR07] V. Mogbil et V. Rahli. Uniform circuits, & boolean proof nets. Sergei Artemov et Anil Nerode, éditeurs, *Logical Foundations of Computer Science*, volume 4514 de *Lecture Notes in Computer Science*, pages 401–421. Springer-Verlag Heidelberg, 2007. International Symposium, LFCS 2007, New York, NY, USA, June 4-7.

[ABT05] V. Atassi, P. Baillot et K. Terui. Verification of Ptime reducibility for system F terms via Dual Light Affine Logic. *Proceedings of Computer Science Logic (CSL'06)*, volume 4207 de *Lecture Notes in Computer Science*, pages 150–166. Springer, 2006.

[deNaurois06] Paulin Jacobé de Naurois. A measure of space for computing over the reals. Arnold Beckmann, Ulrich Berger, Benedikt Löwe et John V. Tucker, éditeurs, *Logical Approaches to Computational Barriers, Second Conference on Computability in Europe, CiE 2006, Swansea, UK, June 30-July 5, 2006, Proceedings*, volume 3988 de *Lecture Notes in Computer Science*, pages 231–240. Springer, 2006.

[BT05] P. Baillot et K. Terui. A Feasible Algorithm for Typing in Elementary Affine Logic. *Proceedings of 7th International Conference on Typed Lambda Calculi and Applications (TLCA 2005)*, volume 3461 de *Lecture Notes in Computer Science*, pages 55–70. Springer, 2005.

[BT04] P. Baillot et K. Terui. Light types for polynomial time computation in lambda-calculus. *Proceedings of 19th IEEE Symposium on Logic in Computer Science (LICS 2004)*, pages 266–275. IEEE Computer Society, 2004.

[BM04] Patrick Baillot et Virgile Mogbil. Soft lambda-calculus : a language for polynomial time computation. *Proceedings of the 7th International Conference on Foundations of Software Science and Computation Structures (FOSSACS'04)*, volume 2987 de *Lecture Notes in Computer Science*, pages 27–41. Springer, 2004.

[DBLP :conf/csl/KanovichV03] Max I. Kanovich et Jacqueline Vauzeilles. Coping polynomially with numerous but identical elements within planning problems. *Computer Science Logic, 17th International Workshop, CSL 2003, 12th Annual Conference of the EACSL, and 8th Kurt Güdel Colloquium, KGC 2003, Vienna, Austria, August 25-30, 2003, Proceedings*, volume 2803 de *Lecture Notes in Computer Science*, pages 285–298. Springer, 2003.

Workshops et autres conférences avec comité de lecture

[BDLM06] P. Baillot, U. Dal Lago et J.-Y. Moyen. On Quasi-interpretations, Blind Abstractions and Implicit Complexity. 8th International Workshop on Logic and Computational Complexity (LCC'06) (Satellite de LICS'06), Seattle, Prepublication HAL ccsd-00023668, août 2006.

[BP04] P. Baillot et M. Pedicini. An Embedding of the BSS Model of Computation in Light Affine Lambda-Calculus. 8th International Workshop on Logic and Computational Complexity (LCC'06) (Satellite de LICS'06), Seattle, Prepublication HAL ccsd-00085547, août 2006.

LCR

Rapports

[KanVau-FSTTCS] Max Kanovich et Jacqueline Vauzeilles. Linear logic in temporal planning for actions with delayed effects. Rapport Technique, LIPN, 2006.

[cfvm05SD] Christophe Fouqueré et Virgile Mogbil. Rewritings in polarized (partial) proof structures. Francois Lamarche Paola Bruscoli et James Stewart, éditeurs, *1st Workshop on Structures and Deductions*, volume ISSN 1430-211X de *Technical Report*, pages 95–109. Technische Universität Dresden, juillet 2005.

Travaux en cours de soumission

[cfvm06-TCS] Christophe Fouqueré et Virgile Mogbil. Rewritings in polarized proof structures. *Mathematical Structures in Computer Science*.

2.4 Composante Spécification et Vérification

Participants : C. Choppy, K. Klai, M. Mayero, L. Petrucci

2.4.1 Thèmes de recherche

Les systèmes développés de nos jours sont de plus en plus complexes et leur fonctionnement peut avoir des conséquences importantes sur leur environnement, voire irréversibles : systèmes avioniques, médicaux ...

Il est par conséquent indispensable d'obtenir des systèmes sûrs, dont le fonctionnement a pu être vérifié avant leur mise en œuvre.

La spécification d'un système présente tout d'abord des avantages quant à la description et la compréhension du système considéré :

- L'écriture de la spécification permet au concepteur de mieux comprendre le fonctionnement attendu de son système ainsi que les interactions entre les différents composants. L'écriture du modèle conduit à remettre en cause des choix conceptuels d'une part et à une meilleure compréhension du système développé d'autre part.
- Le modèle constitue une documentation précise lors de la maintenance.

LCR

De plus, les modèles formels permettent de prouver formellement (i.e. mathématiquement) le bon fonctionnement du système développé. On est alors complètement sûr que quelle que soit l'évolution du système, son comportement sera celui attendu.

Outre les avantages précédents, propres à la spécification, les méthodes formelles offrent des outils d'analyse du système modélisé :

- La simulation permet de se convaincre du bon fonctionnement du système, ou éventuellement de découvrir quelques erreurs. Leur correction à ce stade est peu coûteuse, comparée à celle d'un système déjà mis en œuvre (que ce soit avec des moyens logiciels ou matériels).
- La vérification exhaustive du comportement du système. On applique dans un premier temps des techniques de simulation pour effectuer un débogage grossier, puis le modèle est affiné lors de la vérification des propriétés souhaitées.

Les activités de la thématique *spécification et vérification* s'articulent autour de plusieurs axes de recherche :

- La méthodologie de développement d'une spécification : quel modèle/langage choisir, à quel niveau d'abstraction se placer, comment considérer le problème, comment commencer à écrire et développer la spécification.
- Les systèmes à modéliser étant en général de taille assez conséquente, ils sont spécifiés de manière structurée, par exemple avec une architecture modulaire. L'analyse modulaire vise à tirer parti de cette structuration pour repousser les limites des techniques d'analyse de propriétés.
- Lorsqu'une spécification a été validée à un niveau abstrait, des détails peuvent être introduits par un processus de raffinement. La vérification du modèle raffiné doit alors s'appuyer sur celle du modèle abstrait.
- Enfin, l'utilisation de prouveurs de théorèmes permet de valider les étapes de raffinement de modèles. L'utilisation de ces prouveurs pour des problèmes numériques devra à terme permettre d'aborder des propriétés de systèmes en temps continu.

- **Méthodologie de spécification**

Si l'utilité et la nécessité des spécifications formelles est bien démontrée dès qu'il s'agit de développer ou de manipuler des logiciels complexes, écrire les spécifications d'un système demeure une tâche ardue et délicate. Nous explorons ici comment rendre ce travail plus aisé en fournissant des lignes guide, soit pour écrire directement la spécification, soit en utilisant un langage ou un formalisme intermédiaire plus accessible, ou encore en s'appuyant sur des schémas identifiés (pour cette dernière approche, le lien avec les étapes ultérieures de développement est également étudié).

Ce travail de méthodologie a tout d'abord eu pour but de fournir des spécifications écrites dans le langage de spécification algébrique CASL, ou dans une extension CASL-LTL [Ra-3] pour les systèmes dynamiques. Mais les idées alors développées sont assez générales pour trouver leur application dans l'écriture de spécifications dans d'autres langages (Z, LOTOS, réseaux de Petri, RAISE).

Dans [RI-58] des lignes guide sont proposées pour l'écriture de spécifications de systèmes dynamiques (simples ou structurés) ainsi que de types de données. Différentes caractéristiques constitutives d'un système sont mises en évidence et les propriétés à rechercher sont détaillées.

Le lien de cette méthode avec les cas d'utilisation (qui sont munis d'une description structurée) est décrit dans [CI-54], ce qui permet de fournir une graphique familière aux utilisateurs de la notation UML.

LCR

Il existe des « patterns » pour décrire les problèmes, ce sont les « Problem Frames » de Michael Jackson. Afin de bénéficier des concepts structurants des problem frames (et aussi de leur graphique) nous avons travaillé pour les intégrer dans une démarche de spécification, ce qui nous a conduit à les équiper d'une description plus détaillée proposée en UML à laquelle nous savons associer une spécification formelle [CO-53, CI-53, RI-34]. Ces travaux nous ont aussi conduit à proposer des schémas de problème plus complexes adaptées toutefois à des larges classes d'application [RI-34, CI-121].

Les « patterns » ont d'abord été proposés pour des étapes ultérieures du développement, aussi avons nous exploré comment faire le lien entre une démarche de spécification utilisant les problem frames et un développement à l'aide de styles d'architecture [CI-18], et nous avons proposé dans [CO-55] une approche pour les systèmes d'information, puis dans [RI-35, CI-122] une approche plus générale, dans les deux cas produisant une architecture par composants.

Nous souhaitions mettre en œuvre pour d'autres formalismes les idées proposées dans [RI-58] pour l'écriture de spécification, et nous avons donc exploré dans [CO-54] l'écriture de réseaux de Petri. Ces travaux ont été poursuivis dans le cadre d'un stage de Master Recherche en 2006.

Dans [CL-21, CL-16, TU-2], nous présentons des idées générales sur les critères qui interviennent dans l'écriture des spécifications, et sur certains des travaux présentés ci-dessus. Il est important de noter que chaque méthode de spécification est plus ou moins adaptée à un problème spécifique. C'est pourquoi, il faut tirer parti au mieux des différentes techniques existantes [RE-25].

L'utilisation conjointe de différents outils est nécessaire. Pour atteindre un tel objectif, nous sommes impliqués dans le processus de normalisation des réseaux de Petri [Ra-6, RE-19, CI-107]. Les réseaux de Petri sont normalisés ISO/IEC 15909-1 (première partie d'une norme en 3 volets). Les aspects pratiques concernant un format d'échange (ISO/IEC 15909-2) sont en cours. Ils impliquent la définition d'un métamodèle. Ceci a permis d'automatiser la lecture/écriture d'un fichier écrit en PNML (Petri Net Markup Language) en utilisant une approche MDA (*Model-Driven Architecture*). Le langage PNML a également évolué au cours de ces travaux. De plus, nous avons soumis un addendum à la partie 1 de la norme pour définir les réseaux symétriques, qui sont des réseaux de haut niveau

simples, et servent de base aux travaux sur la partie 2. Enfin, les travaux de la partie 3, portant sur les extensions de réseaux de Petri, viennent de démarrer. Nous sommes éditeurs de ce volet de la norme.

- **Vérification Modulaire**

Nous nous intéressons ici tout particulièrement à des systèmes complexes, parallèles, asynchrones, modélisés à l'aide de réseaux de Petri.

La principale méthode de vérification de propriétés à partir d'un modèle spécifié par un réseau de Petri consiste en l'exploration du *graphe d'états accessibles*. Toutefois, pour des systèmes de taille importante, on se heurte très rapidement au « problème de l'explosion combinatoire de l'espace d'états ». Le graphe d'états est alors trop gros pour être stocké en mémoire.

De nombreuses techniques ont été proposées pour pallier ce problème (ordres partiels, bit-state, ...). Nos travaux se concentrent sur l'analyse modulaire qui tire parti de la structure modulaire du modèle à analyser. L'activité des modules est explorée de manière indépendante plutôt qu'entrelacée. La technique de graphes d'états modulaires que nous développons consiste alors à construire un graphe d'états « local » par module plus une structure appelée « graphe de synchronisation » décrivant les interactions entre les différents modules [IN-4].

LCR

Nous avons proposé, dans [CI-36], des algorithmes effectuant la construction du graphe d'états modulaire d'un réseau de Petri, ainsi que la vérification de propriétés (bornes, blocage, vivacité, ...). L'implémentation d'un prototype a permis de réaliser des expérimentations sur un problème classique d'automates industriels (les véhicules autoguidés). Les résultats ainsi obtenus ont montré des gains considérables — passage de 31.000.000 à 900 états. Toutefois, des études de cas ont été menées dans [RE-18], qui montrent que la technique est particulièrement adaptée à des systèmes composés de sous-systèmes autonomes faiblement couplés tels que ceux considérés en automatique. Dans le pire des cas, lorsque les sous-systèmes sont totalement synchronisés, on aboutit à un léger surcoût.

Nous avons adapté cette approche à la synthèse de réseaux de Petri, en collaboration avec Philiipe Darondeau de l'IRISA [Ra-27].

Suite aux résultats encourageants de la technique de graphes d'états modulaires, nous l'avons étendue à des modèles de réseaux de Petri comportant une modélisation du temps. Nous nous sommes intéressés à la fois aux réseaux temporisés, dans lesquels on considère une horloge globale et les jetons sont estampillés par la date à laquelle ils sont prêts, et aux réseaux temporels, où le temps est compté localement, un intervalle est associé à chaque transition et indique la période dans laquelle la transition doit être franchie, à partir du moment où toutes ses préconditions sont satisfaites.

Nous avons donné, dans [CO-75, RI-72], des algorithmes de construction de graphe d'états modulaires pour les réseaux temporisés. L'inconvénient majeur réside dans l'utilisation de l'horloge globale. En effet les états du graphe sont des couples (marquage, date). Lorsque le système peut évoluer indéfiniment, le graphe obtenu est alors infini.

Dans [CO-112], une technique similaire est présentée pour les réseaux temporels. Les états du graphe sont composés d'un marquage et d'une vecteur de valuations temporelles associées aux transitions. Le graphe obtenu est alors fini (si le réseau de Petri est borné).

Ces différents algorithmes ont été implémentés, et des expérimentations [RI-72] ont montré les apports de l'approche.

une autre approche est l'analyse compositionnelle, qui consiste à vérifer des propriétés d'un sous-système pour valider celles du système complet. Nous avons développé un tel

algorithme pour le vérification de formules LTL\X [CI-147].

- **Raffinement de spécifications de haut niveau**

De nos jours, les systèmes à modéliser et à analyser sont de taille de plus en plus grande. Pour s'affranchir des problèmes inhérents à l'explosion du nombre d'états, une spécification est souvent développée pas à pas (ce qui correspond à une démarche classique de développement de génie logiciel).

Tout d'abord, un modèle abstrait est défini. Une fois que ses propriétés sont vérifiées, une étape de raffinement, qui introduit un niveau de détail supplémentaire, peut être effectuée. Un tel ajout peut préciser la description de l'actuel fonctionnement d'une partie du système, ou introduire une partie supplémentaire. Le modèle ainsi raffiné est ensuite vérifié. Une nouvelle étape de raffinement peut alors être appliquée, et ce jusqu'à atteindre un niveau suffisant de description.

Trois types de raffinement ont été introduits par Lakos et Lewis, pour les réseaux de Petri de haut niveau : le raffinement de sous-réseau, le raffinement de nœud et le raffinement de type. Ces raffinements ont été utilisés dans le cadre d'une approche incrémentale à la spécification de systèmes complexes. Outre les avantages du raffinement pour la spécification de systèmes, ils permettent de faciliter la vérification de propriétés. En effet, d'une part certaines propriétés sont préservées par le raffinement, et d'autre part, l'analyse du modèle abstrait permet de guider celle du modèle concret, économisant ainsi en temps de calcul et espace mémoire lors de la construction de graphes d'états.

LCR

Jusqu'à présent, le processus de raffinement de réseaux est complètement manuel. La vérification des contraintes peut être assez complexe. Par conséquent, nous souhaitons prouver automatiquement qu'un réseau est effectivement un raffinement d'un autre réseau (abstrait). Pour cela, nous utilisons des techniques de preuve de théorèmes, en particulier, l'assistant d'aide à la preuve Coq, développé à l'Inria. Nous proposons donc une formalisation en Coq des définitions de raffinements ainsi que des réseaux auxquels ils s'appliquent. Ceci a tout d'abord été développé pour les raffinements de sous-réseaux et de nœuds des réseaux places/transitions, et appliqué à quelques exemples. Ce travail est étendu, dans le cadre d'un stage de master, aux réseaux colorés.

Le raffinement de types pour les réseaux de Petri de haut niveau a été peu étudié. Pour développer cet aspect, nous nous sommes intéressés dans un premier temps à la vérification de systèmes paramétrés. Le bon fonctionnement de la famille de protocoles « stop-and-wait » a ainsi pu être vérifié [CO-94, RI-41], et ce, de manière paramétrée.

- **Spécifications et preuves formelles**

Dans le cadre du nouveau projet CerPAN [8] (Certification de Programmes d'Analyse Numérique, ANR décembre 2005), nous nous intéressons au développement et à la mise en application des méthodes permettant de démontrer formellement la correction de programmes issus du domaine de l'analyse numérique. Le projet concerne plus particulièrement des programmes apparaissant de manière récurrente dans la résolution de problèmes critiques. Beaucoup de programmes critiques sont issus de ce domaine, mais les travaux traitant directement des applications des méthodes formelles aux programmes d'analyse numérique sont rares. La principale raison est l'utilisation intensive des nombres réels (à virgule flottante) dans les programmes numériques, alors que les méthodes formelles manipulent plutôt des nombres entiers, ou plus généralement des structures discrètes. Deux méthodes sont explorées dans ce projet, toutes deux étant vouées à être mises en pratique sur une étude de cas (gradient analytique).

[8] http://www-lipn.univ-paris13.fr/CerPAN/

1. L'une consiste à utiliser des outils de vérification de programmes tels que Why et Caduceus (dédié aux programmes C) développés au LRI. Il s'agit de travailler sur un programme déjà existant, qui va être annoté suivant les propriétés à préserver. Ces annotations engendrent alors des obligations de preuve pour divers systèmes de preuve formelle (Coq, PVS, Isabelle/HOL, HOL 4, HOL Light, Mizar,...). Ces outils de vérification de programmes ne traitaient pas les nombres flottants. Nous avons donc participé aux phases de décision concernant l'extension de ces outils aux nombres flottants sous forme d'un langage de spécification pour les propriétés relatives aux nombres à virgule flottante. Ces ajouts sont distribués avec le système.
2. L'autre méthode consiste à effectuer la formalisation et la preuve du problème correspondant au programme ou à l'algorithme étudié, puis à utiliser le mécanisme d'extraction du système Coq afin d'obtenir automatiquement un programme prouvé correct. Ce principe d'extraction ne traite pas non plus la famille des nombres réels. Nous avons commencé par décider d'extraire vers des réels exacts et non vers des nombres flottants et nous sommes actuellement en train de rechercher un modèle adéquat de réels exacts (réaliste d'un point de vue implantatoire, efficace, ...) pour étendre l'extraction de Coq.

LCR

Nous essayons de tenir compte des aspects d'efficacité qui peuvent jouer un rôle déterminant dans le domaine de l'analyse numérique en particulier. Les méthodes développées dans ce projet auront un large champ d'application dans les environnements destinés à produire du code numérique garanti correct par rapport à des spécifications formelles et devra à terme permettre d'aborder des propriétés de systèmes en temps continu lors d'étapes de raffinement.

Nous avons développé une procédure de décision dans Coq pour les corps algébriquement clos [CI-118]. Cette procédure de décision est basée sur la méthode d'élimination des quantificateurs et utilise l'interface [RI-31] que nous avons développée entre le prouveur Coq et l'outil de calcul formel Maple.

Nous avons effectué en Coq la preuve du $20^{\text{ème}}$ problème de Diophante (l'aire d'un triangle rectangle dont les côtés sont mesurés par des entiers peut-elle être un carré parfait ?). Ce problème fut résolu négativement par Fermat au moyen de la méthode de descente infinie. La méthode de descente infinie permit à Fermat de prouver le théorème pour $n = 4$. Nous avons formalisé cette preuve en Coq.

Nous proposons une méthodologie pour détecter efficacement les erreurs sur des composants matériel génériques (SoC), utilisant à la fois des techniques de model-checking, de test et des prouveurs de théorèmes.

2.4.2 Perspectives de recherche

- **Méthodologie de spécification**

Les idées développées dans [RI-58] peuvent être reprises et adaptées en fonction des langages de spécification formelle cibles. Des travaux ont été entamés notamment en direction :

- du langage Raise/RSL, en collaboration avec Anne Haxthausen qui est experte de ce langage et qui, à l'occasion du développement d'études de cas, a mené une réflexion sur les aspects méthodologie (sans l'avoir formalisée en tant que telle).
- des réseaux de Petri, pour lesquels, à la suite de [CO-54], des travaux ont été poursuivis en collaboration avec Gianna Reggio, pour les réseaux places/transitions [Ra-22], pour les réseaux colorés [Ra-28], et seront étendus aux réseaux hiérarchiques, puis temporisés.

Ces travaux vont nous permettre d'identifier une partie "générique" indépendante du formalisme cible, et une partie spécifique induite par le choix du type de modèle.

- **Vérification modulaire et distribuée**
  - Vérification modulaire temporisée : les travaux menés jusqu'à présent permettent la construction d'un graphe d'états modulaire pour des modèles de réseaux de Petri avec contraintes de temps. Toutefois, l'intérêt de l'approche est de pouvoir vérifier les propriétés du réseau de Petri sans avoir à déplier le graphe modulaire.
  - Etudes de cas : les techniques développées seront appliquées à des exemples complexes. Nous avons entamé cette démarche en utilisant le « fieldbus protocol ».
  - Vérification distribuée : dans [CI-37], nous avons présenté une implémentation distribuée du calcul de graphe d'états. Nous avons également montré comment les propriétés simples d'un réseau de Petri peuvent être vérifiées. L'analyse de propriétés plus élaborées telles que la vivacité est plus complexe. De plus, nous souhaitons utiliser un cluster de machines mis en place récemment au LIPN pour pouvoir implémenter une plateforme distribuée et effectuer des tests. Cela nous permettra en particulier d'évaluer la répartition des nœuds du graphe sur les différents processeurs. À terme, l'architecture logicielle devra également permettre la construction distribuée du graphe modulaire.
  - Heuristiques pour la construction de graphes d'états : nous avons soumis un projet de « discovery grant » à l'ARC (Australian Research Council), HAVOCS (Heuristic Approaches to the Verification Of Complex concurrent Systems »). L'objectif est d'utiliser des approches heuristiques (telles que les algorithmes fourmis) pour guider la génération du graphe d'états. Cette politique de construction devra rechercher en priorité des chemins pouvant conduire à des états invalidant une propriété souhaitée. Ce travail sera effectué en collaboration avec Hanene Azzag de l'équipe A³.

LCR

- **Raffinement**
  - Les premières expériences effectuées avec le prouveur de théorèmes Coq pour montrer qu'un réseau de Petri places/transitions est un raffinement d'un autre ayant été probantes, nous souhaitons étendre ces travaux aux différents types de raffinements de réseaux de Petri colorés (stage de master recherche en cours). De plus, pour rendre ces travaux utilisables, il est nécessaire d'implémenter un outil faisant la traduction automatique d'un réseau de Petri en une spécification Coq.
  - Parmi les raffinements proposés par Lakos et Lewis, on en distingue deux (raffinements de nœud et de sous-réseau) qui portent sur la structure du réseau de Petri, et un (raffinement de type) qui s'intéresse aux données manipulées dans un réseau de haut niveau. La définition du raffinement de type est pour l'instant très limitée. Nous envisageons donc de l'étendre en nous inspirant des techniques d'abstraction de données et des travaux existant sur l'indépendance de données.
  - Ces travaux seront appliqués à des études de cas, mettant ainsi en évidence les problèmes spécifiques que l'on peut rencontrer lors de tests à grande échelle.

- **Autres activités**
  - Collaboration avec l'ONERA (coencadrement de thèse depuis février 2006) : modélisation de la reconnaissance de chroniques dans le cadre de simulations de systèmes distribués [CO-186, CO-185]
  - Collaboration avec l'équipe A³ : modélisation et vérification du protocole de communication du systèmes multi-agents développé par Hager Karoui, à l'aide de réseaux de Petri de haut niveau.

– Standardisation : les travaux sur la norme ISO/IEC 15909 partie 2 concernant le format d'échange de réseaux de Petri seront poursuivis. Un des points cruciaux est la prise en compte d'extensions qui seront normalisées par la suite dans la partie 3 du standard, dont nous sommes porteurs.

2.4.3 Diffusion de la recherche

- **Evaluation de la recherche**
- **présidence de comités de programme**
  - ACSD, *International Conference on Application of Concurrency to System Design*, juin 2006, Turku, Finlande ; L. Petrucci
  - PNS, *International Workshop on Petri Nets Standards*, juin 2007, Siedlce, Pologne ; L. Petrucci
  - PNTAp, *International Workshop on Petri Nets Tools and APPlications*, mars 2008, Marseille, France ; L. Petrucci
- **comités de programme**

LCR

  - AASAC, *The 23 rd Annual ACM Symposium on Applied Computing (Requirements Engineering)*, mars 2008, Fortaleza, Brazil ; C. Choppy
  - ACSD, *International Conference on Application of Concurrency to System Design*, juin 2004, Hamilton, Canada ; juin 2005, St Malo, France ; juin 2006, Turku, Finlande ; juillet 2007, Bratislava, Slovaquie ; L. Petrucci
  - AICCSA, *ACS/IEEE International Conference on Computer Systems and Applications*, mars 2006, Dubai, UAE ; C. Choppy ; avril 2008, Doha, Qatar ; C. Choppy, L. Petrucci
  - AMAST, *International Conference on Algebraic Methodology and Software Technology*, juillet 2004, Stirling, UK ; juillet 2006, Kuressaare, Estonie ; C. Choppy
  - ATPN, *International Conference on Application and Theory of Petri Nets and Other Models of Concurrency*, juin 2006, Turku, Finlande ; juin 2007, Siedlce, Pologne ; L. Petrucci
  - COOPN, *Workshop on Coordination and Petri Nets (ATPN)*, juin 2005, Miami, USA ; L. Petrucci
  - CPN, *Workshop on Practical use of Coloured Petri Nets*, octobre 2004, 2005, 2006 et 2007, Århus, Danemark ; L. Petrucci
  - DCACSD, *Doctoral Consortium (ACSD)*, juin 2006, Turku, Finlande ; L. Petrucci
  - IEE, *Informatics Education Europe*, novembre 2007, Thessalonike, Grèce ; novembre 2006, Montpellier, France ; special issue in : ITALICS, Innovation in Teaching And Learning in Information and Computer Sciences (2007) ; C. Choppy
  - ISPS, *International Symposium on Programming and Systems*, mai 2005 et 2007, Alger, Algérie ; L. Petrucci
  - IWAAPF, *International Workshop on Applications and Advances in Problem Frames*, mai 2006, Shangai, Chine ; C. Choppy
  - MW, *14th Monterey Workshop, Workshop on Innovations for Requirements Analysis : from stakeholders needs to formal designs* septembre 2007, Monterey, CA, USA ; C. Choppy
  - RST, *International Conference on Reliable Software Technologies (ADA-Europe)*, juin 2007, Genève, Suisse ; juin 2008, Venise, Italie ; L. Petrucci
  - PNML, *Petri Net Markup Language and standardisation*, juin 2004, Bologne, Italie ; mai 2005, Helsinki, Finlande ; juin 2006, Turku, Finlande ; L. Petrucci
  - PNS, *Workshop on Petri Nets Standards*, juin 2007, Siedlce, Pologne ; L. Petrucci

- PNTAp, *Workshop on Petri Nets Tools and APPlications*, mars 2008, Marseille, France ; L. Petrucci
- TeaConc, *Workshop on Teaching Concurrency (ATPN)*, juin 2006, Turku, Finlande ; juin 2007, Siedlce, Pologne ; L. Petrucci

- **comités éditoriaux**

  - dso, ieee Distributed Systems Online, section « Software Engineering », 2006, 2007 ; L. Petrucci
  - ToPNoC, Transactions on Petri Nets and other models of Concurrency, journal publié dans la série lncs, 2007 ; L. Petrucci
  - édition d'un numéro spécial de la revue Fundamenta Informaticae, 2007 ; L. Petrucci
  - édition d'un numéro spécial de la revue TSI, 2008 ; L. Petrucci

- **organismes de normalisation**

  - iso, groupe de travail SC7/WG19 ; L. Petrucci

LCR

- **comités d'évaluation**

  - dfg, *Deutsche Forschungsgemeinschaft*, 2006 ; C. Choppy
  - anr 2006 ; expertises : M. Mayero, L. Petrucci
  - crsngc, *Conseil de recherches en sciences naturelles et en génie du Canada*, 2004, 2005 et 2007 ; C. Choppy — 2006 ; L. Petrucci
  - fqrnt, *Fonds Québécois de la Recherche sur la nature et les technologies*, 2004 ; L. Petrucci
  - nosr, *Netherlands Organisation for Scientific Research*, 2005 ; L. Petrucci

- **autres**

  - Jurys de thèse : L. Petrucci (13)
  - Rapporteurs de thèse étrangère : C. Choppy (1 Québec), L. Petrucci (1 co-tutelle Allemagne, 3 Australie, 1 Cameroun)

- **Organisation de la recherche**

organisation d'événements scientifiques et conférences

  - EuroTICS, *European Computer Science Summit*, octobre 2006, Zurich, Suisse ; C. Choppy
  - forte, *International Conference on Formal Methods for Networked and Distributed Systems*, septembre 2006, Paris, France ; C. Choppy
  - mompes, *MOdel-based Methodologies for Pervasive and Embedded Software (ACSD)*, juin 2005, Rennes, France ; L. Petrucci

- **Contrats**

Type	Coordinateur	Resp. scient.	Nom	Financeur	Période	Durée (mois)	Budget global (Keuros)	Part LIPN (Keuros)
ANR	•	M. Mayero	CerPAN	MENESR	12/2005-12/2008	36	150	40
Linkage Int.		L. Petrucci		ARC	01/2004-12/2004	12	6	3
PAI	•	L. Petrucci	FAST	MAE	07/2004-06/2006	24	23	12
Small grant		L. Petrucci		UniSA	09/2005-12/2005	4	6	0

CerPAN *Autres partenaires : Inria-Rocquencourt, CEDRIC (CNAM), LRI (Univ. Paris 11)*

CerPAN : Certification de programmes d'analyse numérique (ANR programme blanc 2005)

LCR

Dans ce projet nous nous intéressons au développement et à la mise en application des méthodes permettant de démontrer formellement la correction de programmes issus du domaine de l'analyse numérique. Beaucoup de programmes critiques sont issus de ce domaine, mais les travaux traitant directement des applications des méthodes formelles aux programmes d'analyse numérique sont rares. La principale raison est l'utilisation intensive des nombres réels (à virgule flottante) dans les programmes numériques, alors que les méthodes formelles manipulent plutôt des nombres entiers, ou plus généralement des structures discrètes. Deux méthodes sont explorées dans ce projet, toutes deux étant vouées à être mises en pratique sur une étude de cas (gradient analytique). La première consiste à utiliser des outils de vérification de programmes (déjà existants) tels que Caduceus. La seconde consiste à effectuer la formalisation et la preuve du problème correspondant au programme ou à l'algorithme étudié, puis à en extraire (principe d'extraction de Coq) un programme qui sera, de fait, prouvé correct.

PAI FAST & ARC Linkage international & UniSA small divisional grant *Autres partenaires : UniSA (University of South Australia), University of Adelaide*

Ce projet, démarré avec un programme Linkage International de l'ARC en 2004, poursuivi jusqu'en juin 2006 par un PAI FAST (French-Australian Science and Technology programme) et complété par une bourse de l'University of South Australia finançant quelques mois de travail d'un étudiant en thèse, s'est intéressé aux techniques modulaires d'analyse de réseaux de Petri. Il comportait quatre volets : l'analyse de réseaux communiquant par partage de données (places), l'étude de systèmes temporisés, la prise en compte des données et la mise en œuvre de ces techniques dans des outils ainsi que des études de cas. Il a conduit à plusieurs publications : [RI-72, RI-41, CO-75, CO-94, CI-36].

- **Dissémination**

invitations, tutoriels, école d'été, tables rondes

EUROTICS, *European Computer Science Summit*, octobre 2005, Zurich, Suisse ; C. Choppy (table ronde)

IASSE, *ISCA International Conference on Intelligent and Adaptative Systems and Software Engineering*, juillet 2004, Nice, France ; C. Choppy (exposé invité)

ICTAC, *International Colloquium on Theoretical Aspects of Computing*, novembre 2006, Tunis, Tunisie ; C. Choppy et L. Petrucci (tutoriel)

ISPS, *International Symposium on Programming and Systems*, mai 2005, Alger, Algérie ; L. Petrucci (exposé invité)
PNS, *Workshop on Petri Nets Standards (ATPN)*, juin 2007, Siedlce, Pologne ; L. Petrucci (exposé invité, synthèse et groupe de travail)
TeaConc, *Workshop on Teaching Concurrency (ATPN)*, juin 2006, Turku, Finlande ; L. Petrucci (table ronde)
VECOS, *Workshop on Verification and Evaluation of Computer and Communication Systems (ISPS)*, mai 2007, Alger, Algérie ; L. Petrucci (exposé invité)

- **Collaborations et Invitations**
- **Collaborations universitaires**
  - ONERA, co-encadrement par C. Choppy et P. Carle de la thèse de O. Bertrand (depuis février 2006)
  - University of South Australia et University of Adelaide (Australie), PAI FAST 2004–2006
  - University of South Australia et University of Adelaide (Australie), ARC Linkage International grant, 2004
  - University of South Australia small grant (Australie), 2005
  - Århus Universitet (Danemark), BRI Paris 13, 2004
  - USTHB (Algérie), encadrement de 2 thésards (C. Boukala et S. Mazouz) par L. Petrucci
  - LSV, ENS Cachan. Co-encadrement par L. Petrucci et A. Finkel de la thèse de S. Bardin, soutenue le 20 octobre 2005
  - Paris 7, encadrement par C. Choppy de la thèse de E. Sendroiu, soutenue le 16 décembre 2004
  - groupe de travail MeFoSyLoMa (Méthodes Formelles pour les Systèmes Logiciels et Matériels) : LIP6, Lamsade, LIPN, CEDRIC et LTCI

- **Séjours scientifiques**

Christine Choppy :
  - Université de Gênes, Italie, 1 semaine en août 2003.
  - Université de Duisburg-Essen, Allemagne, 1 semaine en Décembre 2003 et 1 semaine en Août 2005.
  - University of South Australia et University of Adelaide, Australie, 1 mois en Avril/Mai 2005 et 1 mois en Avril 2006.

Micaela Mayero :
  - Universidad de la Coruña, Espagne, 1 semaine en septembre 2006.

Laure Petrucci :
  - Université d'Århus, Danemark, 2 semaines en octobre 2004, et 1 semaine en octobre 2006.
  - University of South Australia et University of Adelaide, Australie, 1 mois en Août 2004 et 1 mois en novembre/décembre 2005.
  - Université de Yaoundé, Cameroun, 1 semaine en juillet 2004.
  - USTHB, Alger, Algérie, 1 semaine en mai 2005, 10 jours en mai 2007.

2.4.4 Activités doctorales

- **cours de DEA et stages / Master 2 recherche**

Cours de DEA / Master 2 recherche

- 2005-2007 (C. Choppy et L. Petrucci) : cours de systèmes parallèles asynchrones (option), master 2 recherche informatique, Paris 13.
- 2005-2007 (C. Choppy et L. Petrucci) : cours de spécification et vérification de systèmes (tronc commun), master 2 recherche informatique, Paris 13.

Stages de DEA / Master 2 recherche

- 2004 (C. Choppy, L. Petrucci) : M. Bounif (DEA Paris 13), stage abandonné.
- 2005 (F. Kordon (LIP6), L. Petrucci, N. Trèves (CEDRIC)) : L. Hillah (M2 SAR Paris 6), *Construction d'un standard d'échange de réseaux de Petri dans le cadre de la norme ISO/IEC-15909.*
- 2006 (C. Choppy, L. Petrucci) : W. Luo (M2R MICR Paris 13), *Une méthode d'analyse et spécification pour les système complexes avec des réseaux de Petri de haut niveau et des spécifications algébriques.*
- 2007 (L. Petrucci) : Steven Gay (MPRI, ENS Cachan), *Raffinement de réseaux de Petri.*
- 2007 (C. Choppy, M. Mayero, L. Petrucci) : N. Ghozlane (M2R MICR Paris 13), *Raffinement de réseaux de Petri de haut niveau.*
- 2007 (K. Klai, L. Petrucci) : Rabih Chahine (M2R SLCHP, INP Toulouse), *Analyse compositionnelle de réseaux de Petri.*

LCR

2.4.5 Publications principales

Livres

[MeFoSyLoMa-06] S. Haddad, F. Kordon et L. Petrucci, éditeurs. *Méthodes formelles pour les systèmes répartis et coopératifs*. Hermès, octobre 2006, ISBN 2-7462-1447-4.

Chapitres de livres

[CC-LP-MeFo-06] C. Choppy et L. Petrucci. *Démarches de spécification*, Chapitre 2. Haddad et al. [LI-6], octobre 2006, ISBN 2-7462-1447-4.

[LP-MeFo-06] L. Petrucci. *Panorama des modèles et langages de spécification*, Chapitre 1. Haddad et al. [LI-6], octobre 2006, ISBN 2-7462-1447-4.

Articles dans des revues internationales avec comité de lecture sélectif

[CL-LP-STTT-06] C. Lakos et L. Petrucci. Modular state space exploration for timed Petri nets. *Journal of Software Tools for Technology Transfer*, 9(3–4) :393–411, juin 2007.

[ChoppyReggioJLAP06] C. Choppy et G. Reggio. A formally grounded software specification method. *Journal of Logic and Algebraic Programming*, 67(1-2) :52–86, 2006.

[JB-GG-LP-NJC-05] J. Billington, G. E. Gallasch et L. Petrucci. FAST verification of the class of stop-and-wait protocols modelled by coloured Petri nets. *Nordic Journal of Computing*, 12(3) :275–307, automne 2005.

[ChoppyHatHeisIE05] C. Choppy, D. Hatebur et M. Heisel. Architectural patterns for problem frames. *IEE Proceedings - Software, Special Issue on Relating Software Requirements and Architectures*, 152(4) :198–208, 2005.

[ChoppyReggioIST05] C. Choppy et G. Reggio. A UML-Based Approach for Problem Frame Oriented Software Development. *Journal of Information and Software Technology*, 47 :929–954, 2005.

[DD-MM-JSC] David Delahaye et Micaela Mayero. Dealing with Algebraic Expressions over a Field in Coq using Maple. 39(5) :569–592, 2005.

Articles dans des revues nationales ou internationales

[FK-LP-DSO-06] F. Kordon et L. Petrucci. Toward formal methods oecumenism? *IEEE Distributed Systems Online*, 7(7), juillet 2006.

[LH-FK-LP-NT-PNML-05] L. Hillah, F. Kordon, L. Petrucci et N. Trèves. Building an API for ISO/IEC 15909, based on model engineering techniques. *Petri Net Newsletter*, 69, octobre 2005.

[LP-PNNL-05] L. Petrucci. Cover picture story : Experiments with modular state spaces. *Petri Net Newsletter*, 68 :Cover page and 5–10, avril 2005.

Communications invitées dans une conférence internationale

[invLP-ISPS-05] L. Petrucci. Modularity and Petri nets. *Proc. 7th International Symposium on Programming and Systems (ISPS'2005), Algiers, Algeria*, pages 7–8, mai 2005.

Tutoriaux dans des conférences internationales

[CC-SH-HK-FK-LP-YTM-ICTAC-06] C. Choppy, S. Haddad, H. Klaudel, F. Kordon, L. Petrucci et Y. Thierry-Mieg. Tutorial on formal methods for distributed and cooperative systems. *Proc. 3rd Int. Coll. on Theoretical Aspects of Computing (1-day tutorial at ICTAC'06), Tunis, Tunisia*, volume 4281 de *Lecture Notes in Computer Science*. Springer, novembre 2006.

Communications dans des conférences internationales avec comité de lecture

LCR

[KK-LP-MR-FORTE-07] K. Klai, L. Petrucci et M. Reniers. An incremental and modular technique for checking LTL\X properties of Petri nets. *Proc. 27th International Conference on Formal Methods for Networked and Distributed Systems (FORTE'07), Tallinn, Estonia*, volume 4574 de *Lecture Notes in Computer Science*. Springer, juin 2007.

[ChoppyHatHeisAPSEC06] C. Choppy, D. Hatebur et M. Heisel. Component composition through architectural patterns for problem frames. Pankaj Jalote, éditeur, *Proc. of the Asia Pacific Software Engineering Conference (APSEC-2006)*, pages 27–34. IEEE, 2006.

[ChoppyReggioASWEC06] C. Choppy et G. Reggio. Requirements Capture and Specification for Enterprise Applications : a UML Based Attempt. Jun Han et Mark Staples, éditeurs, *Proc of the Australian Software Engineering Conference (ASWEC-2006)*, pages 19–28. IEEE, 2006.

[DD-MM-Calculemus2005] David Delahaye et Micaela Mayero. Quantifier Elimination over Algebraically Closed Fields in a Proof Assistant using a Computer Algebra System. 151(1) :57–73, 2006.

[LH-FK-LP-NT-FORTE-06] L. Hillah, F. Kordon, L. Petrucci et N. Trèves. PN standardisation : a survey. *Proc. 26th International Conference on Formal Methods for Networked and Distributed Systems (FORTE'06), Paris, France*, volume 4229 de *Lecture Notes in Computer Science*, pages 307–322. Springer, septembre 2006.

[ChoppyReggioFASE04] C. Choppy et G. Reggio. Improving Use Case Based Requirements Using Formally Grounded Specifications. *Fundamental Approaches to Software Engineering (FASE)*, volume 2984 de *Lecture Notes in Computer Science*, pages 244–260. Springer, 2004.

[ChoppyReggioIASSE04] C. Choppy et G. Reggio. Using UML for Problem Frame Oriented Software Development. Walter Dosch et Narayan Debnath, éditeurs, *Proc of the ISCA 13th Int. Conf. on Intelligent and Adaptative Systems and Software Engineering (IASSE-2004)*, pages 239–244. The International Society for Computers and Their Applications (ISCA), 2004.

[LK-LP-ICATPN04] L. Kristensen et L. Petrucci. An approach to distributed state space exploration for coloured Petri nets. *Proc. 25th Int. Conf. Application and Theory of Petri Nets (ICATPN'2004), Bologna, Italy, June 2004*, volume 3099 de *Lecture Notes in Computer Science*, pages 474–483. Springer, juin 2004.

[CL-LP-ACSD04] C. Lakos et L. Petrucci. Modular analysis of systems composed of semiautonomous subsystems. *Proc. 4th Int. Conf. on Application of Concurrency to System Design (ACSD'04), Hamilton, Canada, June 2004*, pages 185–194. IEEE Comp. Soc. Press, juin 2004.

[ChoppyHeiselWADT] C. Choppy, D. Hatebur et M. Heisel. Use of Patterns in Formal Development : Systematic Transition From Problems to Architectural Designs. *Recent Trends in Algebraic Development Techniques*, volume 2755 de *Lecture Notes in Computer Science*, pages 205–220. Springer, 2003.

Workshops et autres conférences avec comité de lecture

[BCC-DX07] O. Bertrand, P. Carle et C. Choppy. Chronicle modelling using automata and colored Petri nets. *The 18th International Workshop on Principles of Diagnosis (DX-07)*, pages 229–234, 2007.

[BCC-RJCIA07] O. Bertrand, P. Carle et C. Choppy. Vers une exploitation des simulations distribuées par les chroniques. *8e Rencontres nationales des Jeunes Chercheurs en Intelligence Artificielle (RJCIA 07)*, pages 15–30, 2007.

[SM-LP-MOCA-06] S. Mazouz et L. Petrucci. Modular discrete pseudo-state graphs for time Petri nets. *Workshop on Modelling of Objects, Components and Agents (MOCA'06, associated with PETRI NETS'06), Turku, Finland*, numéro 272 Universität Hamburg Bericht, pages 223–244, juin 2006.

LCR

[JB-GG-LP-MOMPES-05] J. Billington, G. E. Gallasch et L. Petrucci. Transforming couloured Petri nets to counter systems for parametric verification : A stop-and-wait protocol case study. *Proc. 2nd workshop on MOdel-based Methodologies for Pervasive and Embedded Software (MOMPES'05, satellite of ACSD'05), Rennes, France*, volume 39, pages 37–55. TUCS general publication, juin 2005. ISBN 952-12-1556-9.

[CL-LP-CPN-05] C. Lakos et L. Petrucci. Distributed and modular state space exploration for timed Petri nets. *Proc. Workshop on Practical Use of Coloured Petri Nets, Aarhus, Denmark*, pages 191–210, octobre 2005. Proceedings published as Report DAIMI-PB 576, Aarhus, DK.

[ChoppyHeisel04a] C. Choppy et M. Heisel. Une approche à base de "patrons" pour la spécification et le développement de systèmes d'information. *Proceedings Approches Formelles dans l'Assistance au Développement de Logiciels - AFADL'2004*, pages 61–76, 2004.

[CC-LP-CPN-04] C. Choppy et L. Petrucci. Towards a methodology for modelling with Petri nets. *Proc. Workshop on Practical Use of Coloured Petri Nets, Aarhus, Denmark*, pages 39–56, octobre 2004. Proceedings published as Report DAIMI-PB 570, Aarhus, DK.

[ChoppyReggioIWAAPF04] C. Choppy et G. Reggio. A UML-Based Method for the Commanded Behaviour Frame. K. Cox, J.G. Hall et L. Rapanotti, éditeurs, *Proc. of the 1st International Workshop on Advances and Applications of Problem Frames (IWAAPF 2004)*, pages 27–34. An ICSE 2004 workshop, IEEE, 2004.

Rapports

[CC-LP-GR07r] C. Choppy, L. Petrucci et G. Reggio. A method for modelling with coloured nets. Rapport Technique, Universite Paris 13, 2007.

[Darondeau-Petrucci] Philippe Darondeau et Laure Petrucci. Modular automata 2 distributed Petri nets 4 synthesis. Rapport Technique 6192, INRIA, mai 2007.

[CC-LP-GR06r] C. Choppy, L. Petrucci et G. Reggio. A method for modelling with place/transitions nets. Rapport Technique, Universite Paris 13, 2006.

[FK-LP-NWI-04] F. Kordon et L. Petrucci. Proposal for an addendum to ISO/IEC 15909-1, novembre 2004. NWI submitted by AFNOR to ISO/IEC JTC1-SC7-WG19, ref. MAL-012.

[Reggio :2003 :CLC] G. Reggio, E. Astesiano et C. Choppy. CASL-LTL : A CASL extension for dynamic reactive systems version 1.0 – summary. Rapport Technique DISI-TR-03-36, Univ. of Genova, 2003.

2.5 Composante Bases de Données / Entrepôts de Données

Participants : F. Boufares, K. Lellahi, S. Hamdoun, D. Berrabah (P5), M. Badri (P5)

Outre l'université de Paris 5, les recherches sont menées en collaboration forte avec l'université de Tunis, de Sfax, et d'autres membres du LIPN (H. Bennaceur, Y. Bennani, A. Osmani).

2.5.1 Thèmes de recherche

L'environnement informationnel actuel se caractérise par des données surabondantes (volumineuses) et éparpillées. Celles-ci sont également hétérogènes. En effet, avec l'apparition du web et le développement des différentes représentations et formats des documents, les données peuvent être de plusieurs catégories : structurées (relationnelles, objet), semi-structurées (HTML, XML, graphes) ou même non structurées (texte, images, son). Dans un tel contexte, le besoin d'intégration se fait de plus en plus sentir. Cependant, pour répondre à ce besoin, le développement des applications d'intégration (telles que pour un traitement élaboré de données, pour la construction des entrepôts de données ou des systèmes d'aide à la décision) doit composer avec la répartition des sources et l'hétérogénéité des données et de leurs structures. Le but de notre travail est donc l'intégration de données hétérogènes pour la construction, par exemple, de nouveaux ensembles de données (des entrepôts de données), de systèmes d'information intégrés.

LCR

- **Bases de Données** L'intégration de données hétérogènes nécessite l'analyse et la validation des schémas conceptuels de données (leurs structures) modélisés avec des formalismes différents tels que «EER, Extended Entity-Relationship» et/ou «UML, Unified Modeling Language». Nous faisons une étude détaillée de la cohérence globale de plusieurs catégories de contraintes définies sur les données et des conflits éventuels entre elles. En effet, un grand nombre de contraintes a été très largement étudié dans la littérature. Cependant, peu de travaux sont réalisés pour la détection et la localisation de ces conflits. Nous utilisons des techniques de programmation linéaire et de la logique des prédicats afin de détecter et localiser ces derniers. Par exemple, les contraintes de multiplicités sont modélisées sous forme de systèmes linéaires dont la résolution utilise les techniques de la programmation linéaire. Notre objectif est de développer un «Intégrateur-Compilateur de schémas de Base de Données» en utilisant plusieurs formalismes, d'envisager des passerelles entre ces derniers et d'enrichir les langages d'expression des contraintes.
- **Entrepôts de Données Hétérogènes** Notre but est la construction, la modélisation et la maintenance d'Entrepôts de Données Hétérogènes. Nous considérons ainsi des données sources totalement hétérogènes (des données structurées telles que le relationnel ou l'objet relationnel SQL2&3, des données semi-structurées telles que HTML, XML, et des données non structurées telles que les données du type multimédia MXD). Nous cherchons à proposer un modèle et un processus de construction d'entrepôt de données ayant des sources hétérogènes, et des algorithmes de maintenance de ce type d'entrepôts.

2.5.2 Perspectives de recherche

- **Bases de Données et Systèmes d'information**
  - Dans le cadre de l'intégration de systèmes d'information et d'applications complexes (dont les données sont modélisées et représentées avec plusieurs formalismes différents tels que EER, UML, XML,...), établir des règles de fusion (d'intégration) et vérifier la cohérence globale du nouveau système obtenu. Proposer un formalisme pour enrichir l'expression des contraintes (exprimées dans des environnements cibles tels OCL, Z, B ou SQL) afin de détecter et localiser des conflits éventuels.
  - Dans le cadre d'UML, proposer des règles de passage d'un diagramme à l'autre et en vérifier la cohérence globale.

- **Intégration de données et Entrepôts de Données**
  - Réaliser un outil de construction d'entrepôts de données dont les sources sont totalement hétérogènes structurés, semi-structurées et non structurées.
  - Evaluer la quantité de données non utilisées dans un entrepôt de données afin de réduire sa taille et d'améliorer les performances pour le maintenir.
  - Définir les procédures à suivre afin d'assurer la sécurité des données et la confidentialité de certains types d'informations extraites des dépôts de données. Définir un langage de contrôle de données dans les bases de données multidimensionnelles.

LCR

2.5.3 Séjours Scientifiques, Invitations, Tutoriels

- (F. Boufarès) Université de Tunis 2, séjours d'une semaine à Tunis en Décembre 2003 et Décembre 2004 sur les thèmes : Bases de Données Avancées ; Cohérence globale des contraintes et Exploitation Intelligente des Données et Entrepôts de Données Hétérogènes.
- (F. Boufarès) Tutorial MCSEAI'04, May 2004 : 8th Maghrebian Conference on Software Engineering and Artificial Intelligence (MCSEAI'04), 9th-12th May 2004, Sousse, Tunisia.

2.5.4 Autres

- (F. Boufarès) Lecteur scientifique de plusieurs articles pour le journal IJCS (International Journal of Computing and information Sciences, 2004), pour les conférences MCSEAI'2004, INFORSID'2006, SITIS'2006.
- (F. Boufarès) Co-encadrement de plusieurs thèses avec C. F. Ducateau de Paris 5, K. Lellahi de Paris 13 et S. BenAhmed de Tunis 2.

2.6 Séminaires d'équipe

L'équipe organise régulièrement un séminaire (au moins 15 fois par an) en invitant plus particulièrement de jeunes chercheurs dans cette discipline, mais aussi des chercheurs confirmés français ou étrangers présentant de nouvelles approches (P. Coppola, C. Laneve, B. Valiron, J.-Y. Girard, C. Faggian, K. Terui, H. Mairson, J. Billington ... pour n'en citer que quelques-uns). La liste complète est accessible sur le site web du laboratoire.

Chapitre 3
OCAD (Optimisation Combinatoire et Algorithmique Distribuée)

Sommaire

3.1 Membres de l'équipe 116
3.1.1 Évolution de l'équipe 118
3.2 Présentation scientifique générale 119
3.3 Optimisation Combinatoire 120
3.3.1 Thèmes de recherche 120
3.3.2 Résultats significatifs 121
3.3.3 Perspectives de recherche 123
3.3.4 Diffusion de la recherche 124
3.3.5 Formation à la recherche 127
3.3.6 Publications principales 130
3.4 Algorithmique Distribuée 132
3.4.1 Thèmes de recherche 132
3.4.2 Résultats significatifs 135
3.4.3 Perspectives de recherche 137
3.4.4 Diffusion de la recherche 141
3.4.5 Formation à la recherche 145

3.4.6 Publications principales 146
3.5 Séminaire 148

3.1 Membres de l'équipe

Membres au 01/07/2007 (les membres ayant quitté le LIPN figurent en italique) :

Permanents

	Nom	*Prénom*	*Situation*	*Institution*
	Alfandari	Laurent	MCF	ESSEC
	Banderier	Cyril	CR	CNRS
	Bennaceur	Hachemi	MCF	IUT Cergy
	*Bourgeois*	*Philippe*	MCF	IUTV
	Butelle	Franck	MCF	IUTV
	Cérin	Christophe	PU	IUTV
	Dubacq	Jean-Christophe	MCF	IUTV
	Finta	Lucian	MCF	IUTV
	Lavault	Christian	PU	IG
	Létocart	Lucas	MCF	IG
	*Nagih*	*Anass*	PU	Un. Metz
	Pekergin	Ferhan	MCF	IUTV
	Plateau	Gérard	PU	IG
	Ravelomanana	Vlady	MCF	IG
	Toulouse	Sophie	MCF	IG
	Valencia-Pabon	Mario	MCF	IG

OCAD

Doctorants

	Nom	*Prénom*	*Situation*	*Institution*
	Abbes	Heithem	B	Franco-Tunisienne
	Andriamampinina	Tsiriniaina	ATER	IG
	Fkaier	Hazem	B	Franco-Tunisienne
	Kenniche	Hichem	AM	IG
	Laval	Olivier	AM	IG
	Lemalade	Jean-Luc	B	Franco-Malgache
	Rasoanaivo	Andry	B	AUF
	Sadki-Fenzar	Jalila	AM	IG
	Touati	Nora	AM	IG

Impétrants

	Nom	*Prénom*	*Situation*	*Institution*
	*Castro de Andrade*	*Rafael*	ATER	IG
	Djerourou	Fayçal	ATER	IG
	Hakem	Mourad	ATER	IG
	Khemmoudj	Idir	B	EDF
	*Paillard*	*Gabriel*	B	Brésil
	*Marcos*	*Nicolas*	B	SNCF
	*Thiongane*	*Babacar*	ATER	IG

OCAD

Autre personnel temporaire du 01/09/2002 au 01/09/2007 :

Nom	*Prénom*	*Situation*	*Début*	*Fin*	*Financement*
Marcos	Nicolas	Chef projets recherche	02/2006		SNCF
Culus	Jean-François	post-doc	10/2006	09/2007	CNRS (STIC)
Quivet	Alexandre	Ingénieur de Recherche	01/2005	10/2005	ACI
Viola	Alfredo	poste rouge	07/2004	12/2004	CNRS (STIC)
Le Bars	Jean-Marie	délégation CNRS	01/2003	12/2004	CNRS
Maculan	Nelson	poste rouge	10/2002	09/2003	CNRS (STIC)

Personnels invités du 01/09/2002 au 01/09/2007 (au moins une semaine) :

N.b. : Les financements attribués à P13 désignent les professeurs invités pour 1 mois par l'université Paris 13.

Prénom	*Nom*	*Origine*	*Financement*
2007			
Jean	Cardinal	Un. libre de Bruxelles (ULB, Belgique)	P13
Tadeusz	Czachorski	Académie des Sciences de Pologne	CNRS & PAN
Krzysztof	Grochla	Académie des Sciences de Pologne	CNRS & PAN
M. Cristina	Pinotti	Un. degli Studi di Perugia (Italie)	CNRS & Un. Perugia
Arthur	Randrianarivony	Un. d'Antananarivo (Madagascar)	P13
2006			
Monique	Guignard	Wharton School, Un. of Pennsylvania (USA)	P13
Bernhard	Gittenberger	T.U., Wien (Autriche)	Égide-Amadeus
Hsien-Kuei	Hwang	Academia Sinica, Taipei (Taiwan)	P13
Mohamed	Jemni	ESSTT (Tunis)	MCF invité
Marcus	Kuba	T.U., Wien (Autriche)	Égide-Amadeus
Alois	Panholzer	T.U., Wien (Autriche)	Égide-Amadeus
Czachorski	Tadeusz	Académie des Sciences de Pologne	P13
2005			
Heithem	Abbes	ESSTT (Tunis)	inria-STIC
Hazem	Fkaier	ESSTT (Tunis)	inria-STIC
Bernhard	Gittenberger	T.U., Wien (Autriche)	Égide-Amadeus
Mohamed	Jemni	ESSTT (Tunis)	inria-STIC
Nelson	Maculan	Un. fédérale de Rio de Janeiro (Brésil)	PU (Docteur honoris causa)
Stephan	Olariu	Old Dominion Un. Norfolk (USA)	P13
Alois	Panholzer	T.U., Wien (Autriche)	Égide-Amadeus
2004			
Nelson	Maculan	Un. fédérale de Rio de Janeiro (Brésil)	P13
François	Soumis	École Polytechnique de Montréal (Canada)	P13 & ACI-Jeunes Chercheurs
Czachorski	Tadeusz	Académie des Sciences de Pologne	P13
Alfredo	Viola	Un. de la República (Uruguay)	Poste rouge CNRS
2003			
Nelson	Maculan	Un. fédérale de Rio de Janeiro (Brésil)	P13 & Poste rouge CNRS

3.1.1 Évolution de l'équipe

Permanents

Départs

2002 V. Gabrel, MCF (composante OC), mutation à Paris 9
2003 V. Zissimopoulos, PU (composante OC), mutation à l'Un. d'Athènes
2004 Ph. Bourgeois, MCF (composante OC), décès
2006 A. Nagih, MCF (composante OC), promotion PU à l'Un. de Metz

Arrivées

2002 L. Alfandari, PA ESSEC, composante OC (auparavant LAMSADE-Paris 9)
2003 L. Létocart, MCF IG, composante OC (thèse CEDRIC-CNAM)
2004 S. Toulouse, MCF IG, composante OC (thèse LAMSADE-Paris 9)
2005 C. Cérin, PU IUTV, composante AD (auparavant MCF Amiens, LARIA)
2006 M. Valencia-Pabon, MCF IG, composante AD (thèse LRI-Paris 11)

Depuis 2002, l'équipe OCAD a subi plusieurs modifications qui ont infléchi sa problématique d'ensemble en renouvelant les priorités de certain thèmes de recherche. Pour la composante *optimisation combinatoire*, l'arrivée en 2002 d'un professeur associé employé à l'ESSEC et le recrutement de deux maîtres de conférences en 2003 et 2004 n'ont pas contrebalancé les départs de trois maîtres de conférences et d'un professeur ces dernières années. En *algorithmique distribuée*, l'arrivée d'un chargé de recherche CNRS en 2002 a renforcé le thème combinatoire de la composante. Quant aux recrutements d'un professeur et d'un maître de conférences, ils ouvrent de nouvelles perspectives de recherche. Depuis 2005, se développe ainsi un nouveau thème de

travail autour de Christophe Cérin : *Grappes et grilles de calcul* (*Grid Computing*) — *i.e.* la conception d'outils pour les problèmes parallèles de grande taille sur des grappes et des grilles de calcul.

3.2 Présentation scientifique générale

L'informatique définit, étudie et manipule des structures finies ou discrètes à travers des algorithmes. L'objectif de l'équipe OCAD (*Optimisation Combinatoire* (OC) et *Algorithmique Distribuée* (AD)) est l'analyse et l'optimisation fines de systèmes complexes discrets, séquentiels ou parallèles, pouvant présenter une forte composante aléatoire. Un cadre formel, naturel et cohérent pour mener ce travail est la combinatoire. Ainsi la recherche conduite au sein de l'équipe OCAD mêle étroitement trois domaines complémentaires : la combinatoire, l'algorithmique et la complexité. Ce cadre intègre notamment l'optimisation de problèmes difficiles et de problèmes de grande taille par des algorithmes efficaces (approchés, voire exacts pour certaines instances) ou l'étude quantitative d'algorithmes probabilistes sur des structures ou des objets combinatoires.

OCAD

3.3 Composante Optimisation Combinatoire

Participants : L. Alfandari, H. Bennaceur, R. Castro de Andrade, F. Djerourou, M.I. Khemmoudj, O. Laval, J.-L. Lemalade, L. Létocart, N. Marcos, A. Nagih, F. Pekergin, G. Plateau, A. Rasoanaivo, J. Sadki-Fenzar, N. Touati, S. Toulouse

3.3.1 Thèmes de recherche

Initié en partie depuis septembre 2004 par l'ACI JC (cf. § 3.3.4), le projet de recherche principal de la composante OC est centré sur les méthodes de décomposition (schémas itératifs) pour la résolution exacte de problèmes difficiles d'optimisation combinatoire, méthodes issues de la programmation mathématique. Plus précisément, ce projet s'intéresse à l'accélération et au pilotage de tels schémas sous deux angles spécifiques : la *résolution approchée* d'une part (pour trouver « rapidement » de « bonnes » solutions) et les méthodes de *stabilisation* et de *réoptimisation* d'autre part (pour accélérer la convergence en fin de résolution).

Par ailleurs, la composante OC se consacre à la poursuite de travaux fondamentaux et appliqués en *optimisation en nombres entiers*. L'aspect fondamental concerne plus spécifiquement la modélisation, les relaxations, la complexité et l'approximation, les heuristiques, les métaheuristiques, et les méthodes de résolution exacte, tandis que les applications sont tournées vers le transport, le dimensionnement de réseaux et la gestion de production agricole.

Enfin, le groupe OC poursuit également des recherches dans le domaine de la *satisfaction et de l'optimisation des systèmes de contraintes*. En particulier, des méthodes combinant des techniques d'optimisation combinatoire et de satisfaction de contraintes ont été conçues, et appliquées notamment aux bases de données et à la production d'énergie nucléaire.

Optimisation en nombres entiers

OCAD

L'optimisation combinatoire a pour vocation de construire, pour des problèmes d'optimisation en variables mixtes ou en nombres entiers, des solutions exactes ou approchées. Les approches émergentes sont issues de la programmation mathématique, qui offre des formalismes unifiés pour une large gamme de problèmes et permet ainsi d'élaborer des méthodes de résolution génériques efficaces. La validation peut alors se faire à la fois sur un plan théorique — notamment par l'analyse du double point de vue de l'approximation et de la complexité — et sur un plan expérimental.

1 Résolution efficace de problèmes d'optimisation en nombres entiers de très grande taille

S'appuyant sur une forte complémentarité des domaines de recherche de ses différents membres, le projet de recherche de la composante OC s'inscrit plus précisément dans le cadre de la résolution globale et efficace de problèmes d'optimisation en nombres entiers de très grande taille, avec des volets théoriques, algorithmiques et applicatifs.

Les modèles étudiés sont pour la plupart issus des domaines de la planification des ressources dans les systèmes de production et de la gestion des flux dans les réseaux de transport. Dans ces domaines, la prise en compte récente des contextes d'incertitude d'une part, d'éléments qualitatifs d'autre part (notamment dans les réseaux), a mené à l'introduction de nouvelles familles de modèles, intégrant la formulation de contraintes particulières, ainsi qu'un nombre toujours accru de variables de décision. Les problèmes résultants sont dès lors très difficiles, de par leur taille (plusieurs milliers de variables et de contraintes), mais aussi de par leur structure combinatoire (présence de variables discrètes), et nécessitent, pour leur résolution effective, l'élaboration d'outils théoriques et algorithmiques spécifiques.

Par ailleurs, le contexte opérationnel est en demande d'un arbitrage constant entre la double exigence de qualité des solutions fournies et de temps de calcul raisonnables. Alors qu'en planification stratégique, l'intérêt de l'étude de différents scénarios prime sur les temps de calcul nécessaires pour les obtenir, en planification opérationnelle, le temps de réponse devient crucial : en cas de perturbation majeure, il faut reconstruire le plan de production en temps réel, tout en minimisant l'écart par rapport au plan prévisionnel afin de réduire l'impact sur la suite des

opérations. De nouveau, les outils actuels ne sont pas capables de répondre à ces deux exigences sur des problèmes rendus de plus en plus complexes.

Pour pallier les insuffisances des approches existantes prises séparément, on souhaite élaborer des schémas hybrides sachant tirer profit des avantages de chacune d'elles : la capacité de produire de bonnes solutions réalisables en temps de calcul acceptable pour les approches de type (méta)heuristique d'une part, la vision globale du problème propre aux approches basées sur la programmation mathématique, comme la génération de colonnes, d'autre part. Les schémas hybrides sont en effet souvent plus robustes et permettent de traiter une classe plus large d'instances. L'efficacité de la résolution globale peut aussi être améliorée par l'utilisation de techniques de réoptimisation exploitant l'historique des résolutions des instances successives d'un même problème. Une telle approche, combinant hybridation et réoptimisation, permet d'élaborer des méthodes de résolution qui, non seulement fournissent des solutions de très bonne qualité en temps de calcul raisonnable, mais aussi permettent de recalculer ces solutions en cas de faible perturbation des données.

Ces recherches sont élargies à des modèles et problèmes émergents, prenant en compte la qualité des solutions produites, l'introduction de l'incertitude, ou encore la notion d'aide à la décision (notion de robustesse, diversification des solutions, modélisation multicritère, etc.).

2 Aspects fondamentaux et appliqués

Le projet central du goupe OC est nourri des travaux fondamentaux et appliqués qui sont poursuivis par ailleurs. Sous l'angle de la programmation mathématique elle-même, des efforts particuliers sont produits sur l'étude de relaxations fortes pour les problèmes d'optimisation en variables 0-1, linéaires et non linéaires. Concernant la partie approximation polynomiale, elle se concentre plus particulièrement sur la classification des problèmes d'optimisation combinatoire, notamment sous l'angle encore peu développé de l'approximation différentielle. Par ailleurs, certaines familles de problèmes (notamment, recouvrement et partitionnement) font l'objet de recherches algorithmiques spécifiques pour leur résolution ou pour l'amélioration de leur degré d'approximation.

OCAD

Satisfaction et optimisation des systèmes de contraintes

Les activités de recherche de ce thème concernent le raisonnement à base de contraintes. Dans ce domaine, la connaissance est décrite sous forme de contraintes ou relations (symboliques ou numériques). L'objectif est d'étudier la cohérence (ou la satisfaction) et l'optimisation du système de contraintes associé à une base de connaissances. La conception de méthodes pour ces problèmes se base essentiellement sur les outils de deux domaines différents, mais très complémentaires, à savoir la Recherche Opérationnelle (RO) et l'Intelligence Artificielle (IA).

Le travail actuel concerne la combinaison des techniques d'optimisation combinatoire (*relaxation lagrangienne*, *branch and bound*, *méthodes de recherche locale*) et les techniques de satisfaction de contraintes pour résoudre les problèmes CSP et CSP valués.

Modélisation stochastique de systèmes discrets

La recherche de ce thème concerne la modélisation de systèmes à événements discrets par des processus stochastiques et au développement de méthodes de résolution. Un intérêt particulier est attribué à la modélisation et l'évaluation des performances des systèmes complexes, dont les modèles sont de très grande taille dans leur représentation classique. Pour contourner les problèmes dus aux espaces d'états de modèles comprenant des mécanismes de contrôle complexes, on peut recourir à une modélisation par des processus continus. Cette approche consiste, d'une part, à reconstituer l'information perdue du fait de la transformation discret-continu et, d'autre part, à développer les algorithmes numériques nécessaires à leur résolution.

3.3.2 Résultats significatifs

Optimisation en nombres entiers

1. Au sein d'un schéma itératif, une succession d'instances d'un problème dit auxiliaire doit être résolue. L'introduction des techniques de réoptimisation dans un tel schéma consiste

à accélérer la résolution d'une instance de la suite en tirant profit des résolutions des instances qui la précèdent.

L'élaboration de ces techniques n'a de sens que si les données des instances sont suffisamment voisines. Pour mieux appréhender cette question de la réoptimisation, il a été décidé de l'entamer dans le cadre de problèmes classiques d'optimisation. C'est pourquoi les travaux se sont d'abord orientés vers la résolution efficace du dual lagrangien du problème du sac à dos à deux dimensions (biknapsack) en variables 0-1. Réalisés dans le cadre de la thèse de Babacar Thiongane, ils ont permis des avancées significatives relativement à l'évolution des données de l'objectif et d'obtenir ainsi une large base de résulats théoriques, algorithmiques et expérimentaux :

- état de l'art sur la *postoptimisation* (analyse de sensibilité, paramétrisation et réoptimisation) en programmation en nombres entiers [RE-1],
- étude de la *complexité* du problème de la paramétrisation en programmation linéaire en 0-1 et élaboration d'un nouvel algorithme,
- *réoptimisation* dans le dual lagrangien du biknapsack en variables 0-1 : élaboration de deux approches (MARIE et ASIA) pour la résolution efficace du dual lagrangien par des méthodes de sous-gradients adaptées [RI-25, RI-46] ; combinaison de métaheuristiques et de techniques de réoptimisation pour la construction d'heuristiques performantes pour le problème de biknapsack en variables 0-1 [RI-46].

2. Des études ont été menées pour le dimensionnement de réseaux de télécommunications. Dans le cas général, la génération de colonnes adaptée à l'optimisation en nombres entiers a été appliquée avec succès en utilisant un *Branch-and-Bound* (méthode par séparation et évaluation) et des coupes géométriques pour éviter d'engendrer des colonnes inadéquates (thèse d'Irène Loiseau). Lorsque les demandes sont incertaines, des techniques issues de l'optimisation stochastique en nombres entiers ont été appliquées (*Branch-and-Cut* et méthodes « à la Benders ») [RI-23, RI-44, RI-62].

OCAD

3. Le statut, en termes de complexité (résolution exacte en temps polynomial) et d'approximabilité (ordre du degré d'approximation que l'on peut espérer obtenir en temps polynomial), de différents problèmes a été décidé : ces résultats concernent des problèmes de multiflots entiers maximaux et de multicoupes minimales [RI-33], notamment dans les graphes en anneaux [So-3], ainsi que certains problèmes de partionnement et de recouvrement d'un graphe par des chaînes de longueur bornée, pour lesquels de meilleures bornes d'approximation classique et différentielle ont également été obtenues [CL-32].
4. Des travaux ont été réalisés sur l'élaboration de modélisations sous forme de PLNE (programmes linéaires en nombres entiers) d'ordre polynomial pour la détermination de sous-graphes connexes de poids minimal d'un graphe valué (par exemple, chemin élémentaire, arbre et cycle) [RE-3].
5. Des schémas d'approximation utilisant des méthodes de programmation dynamique spécifiques ont été développés et analysés pour des sous-problèmes apparaissant dans des problèmes classiques de couverture de tâches en transport et de localisation d'entreprise [RI-83].

Satisfaction et optimisation des systèmes de contraintes

6. Une réécriture originale du problème SAT en un CSP binaire a été exploitée pour montrer qu'il existe de nombreux liens algorithmiques et des concepts équivalents utilisés pour ces deux problèmes [RI-20]. Différents types de méthodes améliorant la résolution du problème Max-CSP ont par ailleurs été développées :
   - méthodes de recherche locale intégrant les techniques de consistance locale (consistances d'arc) et de réduction du voisinage,
   - méthodes exactes hybrides combinant les techniques d'optimisation combinatoire et de satisfaction de contraintes [CI-72]. La plupart de ces méthodes ont été implémentées et testées sur des instances aléatoires, académiques et réelles [CI-103].

Modélisation stochastique de systèmes discrets

7. Par une modélisation qui s'appuie sur l'approximation de diffusion, le comportement transitoire des flux de classes différentes dans les réseaux de communications modélisés par des

réseaux de files d'attente a été étudié. L'approximation de diffusion est utilisée dans les modèles à classes multiples pour contourner le problème de l'influence de certains types de transfert de données sur le reste du trafic, tout en compensant la perte d'information due à la transformation des processus discrets en continus [RI-32]. Pour représenter les mécanismes de contrôle complexes des réseaux de communications, des modèles de diffusion à sauts multiples sont développés. La validité de la combinaison des sauts avec l'approximation de diffusion pour un modèle de routeur opto-électronique a été montrée et une méthode de résolution numérique pour ce type de modèle a été développée [CO-93].

3.3.3 Perspectives de recherche

Optimisation en nombres entiers

1. Concernant la mise en œuvre de schémas itératifs de type génération de colonnes, nos principaux travaux portent sur les deux points suivants :
   - *la réoptimisation*, méthode permettant d'améliorer la complexité des itérations dans la phase de convergence (une thèse en cours).
   - *l'approximation*, qui offre une justification théorique du comportement des heuristiques (une thèse en cours),

   (a) L'objectif actuel est la transposition de techniques de réoptimisation (initiées par les travaux sur les problèmes de sac-à-dos en variables 0-1) à un schéma de type génération de colonnes pour la résolution des problèmes de très grande taille, notamment dans le cadre de la gestion en temps réel des flux dans des réseaux de transport. Ces recherches se poursuivent en priorité vers la réoptimisation de problèmes en nombres entiers avec application au sac à dos quadratique [CO-165] et au plus court chemin avec contraintes de ressources [CO-150]. La problématique abordée est celle de la résolution globale efficace d'une suite d'instances du problème traité. Trois axes sont donc développés :
   - dans le cadre des dualités lagrangiennes et agrégées, la prise en compte de l'évolution des données du problème (méthodes de sous-gradient, génération de colonnes),
   - l'hybridation de métaheuristiques et d'heuristiques duales,
   - l'incrémentalité, c'est-à-dire la prise en compte de l'évolution du nombre de contraintes et/ou de variables.

   (b) En approximation, une perspective est d'étudier l'hybridation de méta-heuristiques et d'heuristiques à garantie de performance à la mise en œuvre de schémas itératifs. Plus précisément, il s'agirait de construire des populations de solutions diversifiées et de qualité contrôlée, de sorte à mieux piloter le schéma itératif de résolution.

   (c) Il est également nécessaire de compléter l'expertise de la composante dans le domaine des approches polyédrales, de sorte à entamer une voie de recherche complémentaire aux travaux actuels sur l'accélération de la convergence des schémas itératifs. En effet, de telles approches permettraient de diminuer le nombre d'itérations de ces schémas par la diminution du domaine de recherche à chaque itération.

2. Concernant l'approximation polynomiale au pire des cas, un axe privilégié est celui de la classification des problèmes d'optimisation discrète sous l'angle différentiel, au travers de deux défis essentiels : d'une part, décider du degré d'approximation différentielle de problèmes de type MaxSAT/MaxCCSP ; d'autre part, établir un cadre pertinent de réduction pour la préservation du rapport constant d'approximation différentiel, en menant une étude d'ordre structurel sur les instances des problèmes à lier par réduction. L'étude des problèmes de logique, par des méthodes à base de voisinages, a fait l'objet d'un post-doctorat CNRS en 2006-2007 qui a permis d'améliorer les bornes pour l'approximation différentielle des problèmes Max$k$SAT. Notons que ce projet est en lien étroit avec le groupe de travail *Approximation & Complexité* (cf. § 3.3.4).
3. La résolution exacte de problèmes non linéaires en nombres entiers, et plus particulièrement de problèmes à objectif quadratique, contraintes linéaires avec des variables bivalentes, est actuellement à l'étude par l'intermédiaire de méthodes lagrangiennes et de

OCAD

techniques de convexification [CO-163].

Satisfaction et optimisation des systèmes de contraintes

4. Les perspectives concernent essentiellement les applications des techniques CSP et de l'optimisation combinatoire aux problèmes réels industriels. Plus précisément, il s'agit d'approfondir l'étude d'approches combinant la programmation par contraintes et la recherche locale (prenant en compte un autre type de contraintes du problème de placement des arrêts des réacteurs nucléaires).

   À court terme, il est prévu d'étudier l'extension de certaines formes de consistance locale aux CSP pondérés (*WCSP*) et de développer d'autres formes de consistance locale basées sur la consistance d'arc.

   Un autre objectif est également de travailler sur le logiciel de référence de *WCSP Toolbar* pour y intégrer les résultats obtenus et évaluer les performances des méthodes élaborées.

Modélisation stochastique de systèmes discrets

5. Dans le cadre des travaux de modélisation des systèmes à événements discrets, le premier objectif est le développement de méthodes de résolution robustes et efficaces en temps d'exécution. Une étude de comparaison de coûts s'impose, afin de déterminer l'adéquation entre la complexité des modèles et l'efficacité des différentes méthodes. Ensuite, on peut envisager l'étude de plusieurs systèmes complexes par des modèles de diffusion à sauts multiples : serveurs web, centres d'appels, routeurs optique-électroniques, etc.

Applications

6. Bien que la recherche en OC se fonde sur le développement et l'étude théorique de méthodes génériques de résolution, les techniques développées s'appliquent à de nombreux domaines (transport, réseau, etc.). Il convient naturellement de continuer à les éprouver auprès des industriels, en consultation ou sur la base de projets.

OCAD

3.3.4 Diffusion de la recherche

- **Évaluation de la recherche**

Comités éditoriaux

- Anass Nagih : Bulletin de la Société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF) en 2003.
- Gérard Plateau : Journal of Heuristics depuis 1995
- Gérard Plateau : European Journal of Operational Research (2000-2006)
- Gérard Plateau : RAIRO - Operations Research depuis 1986

Comités de programmes

- Gérard Plateau : MIC'2003, 5th Metaheuristics International Conference, Kyoto, Japon, 25-28 août 03.
- Gérard Plateau : MIC'2005, 6th Metaheuristics International Conference, Vienne, Autriche, 22-26 août 05.
- Gérard Plateau : ALIO/EURO Vth Meeting, Paris, 26-28 octobre 05.
- Gérard Plateau : APMOD'06, 8th Applied mathematical Programming and MODelling meeting, Madrid, Espagne, 19-21 juin 06.
- Gérard Plateau : MIC'2007, 7th Metaheuristics International Conference, Montréal, Canada, 25-29 juin 07.
- Gérard Plateau : NCP'07, Non Convex Programming, INSA, Rouen, 17-21 décembre 07.

Comités de lecture ou d'évaluation

- Laurent Alfandari. Rapporteur d'une demi-douzaine d'articles pour des revues et conférences internationales (Comp. Opt. and Appli., Eur. J. of OR, RAIRO-OR, etc.).

- Lucas Létocart. Rapporteur de trois articles pour des revues internationales (Dis. Appl. Maths, RAIRO-OR, TOP).
- Anass Nagih. Rapporteur de six articles pour des revues et conférences internationales (Eur. J. of OR, RAIRO-OR, CP-AI-OR, ROADEF).
- Gérard Plateau. Rapporteur de vingt articles pour des revues et conférences internationales (Eur. J. of OR, OR, Workshop Math. Programming, MIC, ROADEF).

Autres responsabilités d'évaluation

- Gérard Plateau : FQRNT (Fonds québécois de la recherche sur la nature et les technologies) : évaluation d'une demande de projet en janvier 2004 (Québec)
- Gérard Plateau : CRSNG (Conseil de recherche en sciences naturelles et en génie du Canada) : évaluation d'une demande de projet en janvier 2004 (Ottawa)
- Gérard Plateau : CNRS : expertises de projets Agence Nationale de la Recherche 2005
- Gérard Plateau : Ministère : expertises de dossiers pour la DS1 puis de la DSPT9 - Master en 2004 et en 2007.
- Gérard Plateau : Membre du comité scientifique du GDR RO : expertise de projets 2005-2006.
- Gérard Plateau : Région Champagne-Ardenne : Expertise d'un dossier d'allocation post-doctorale en mars 2004
- Gérard Plateau : Membre des comités scientifiques des laboratoires d'informatique de Tours et d'Avignon en 2007.

Jurys hors LIPN

Jury de thèse : G. Plateau (6, rapporteur :5), A. Nagih (2, rapporteur :1), S. Toulouse (1, rapporteur :1).
Jury d'habilitation à diriger des recherches : G. Plateau (1, rapporteur :1).

OCAD

• Organisation de la recherche

Organisation de conférences et événements scientifiques

- Anass Nagih : Membre du comité d'organisation du Prix Robert Faure 2003 (prix récompensant une contribution originale dans le domaine de la Recherche Opérationnelle et l'Aide à la Décision).
- Laurent Alfandari et Lucas Létocart : Membres du comité d'organisation des Journées Franciliennes de Recherche Opérationnelle de 2003 à 2006 : organisation de 8 journées
- Anass Nagih : organisation de la participation du LIPN à la manifestation *Savante Banlieue*, Villetaneuse, 13-14 octobre 2005.

Gestion de groupes de travail Un nouveau groupe de travail, *Approximation & complexité*, a été créé en septembre 2006. Sa vocation première est la structuration de **NPO**, c'est-à-dire l'étude de la complexité d'approximation des problèmes difficiles d'optimisation combinatoire, en attaquant les problèmes sous un angle structurel plutôt qu'algorithmique. L'idée est alors de faire concourir à l'objectif de classification différentes approches et savoir-faire qui sont présents au laboratoire dans d'autres contextes de recherche : l'optimisation combinatoire (groupe OC), la combinatoire analytique (groupe AD), la combinatoire algébrique (LCR) et la théorie de la complexité (OCAD & LCR).

• Contrats

Type	Coordinateur	Resp. scient.	Nom	Financeur	Période	Durée (mois)	Budget global (keuros)	Part LIPN (keuros)
ACI	LIPN	A. Nagih	COGiTO	MENESR	09/2004-08/2007	36	66	66
CIFRE	SNCF	G. Plateau		SNCF	10/2002-09/2005	36	22,5	22,5
Contrat	EDF	H. Bennaceur		EDF	05/2003-04/2006	36	75	
Contrat	EDF	H. Bennaceur		EDF	08/2006-01/2008	18	83	10
Veille	SNCF	A. Nagih		SNCF	02/2005-11/2007	34	22,5	22,5

ACI COGiTO Le projet de recherche d'OC bénéficie (depuis septembre 2004) du soutien d'une ACI Jeunes Chercheuses et Jeunes Chercheurs, composée d'Anass Nagih, Laurent Alfandari, Lucas Létocart, Agnès Plateau (CNAM) et Sophie Toulouse, sur le thème *Planification et gestion optimisée des ressources en transport : réoptimisation et hybridation de la génération de colonnes et des métaheuristiques* (2004-2007). Ce projet, qui va de la modélisation à la réalisation informatique via la résolution, a notamment offert le cadre de collaborations avec des partenaires académiques (GERAD) et industriels (SNCF). Il a également permis de réaliser courant 2005 une plate-forme logicielle d'optimisation appelée COGiTO.

OCAD

CIFRE SNCF Dans le cadre de la thèse CIFRE SNCF de Nicolas Marcos, le travail a concerné la gestion de production et la planification en transport ferroviaire. Plus spécifiquement, un outil générique d'aide à la décision pour la gestion des locomotives fret à la SNCF a été élaboré grâce à de nouvelles modélisations du problème utilisant l'aspect multiflot de cette application [CO-73].

Contrats EDF L'étude du *problème de placement des arrêts des réacteurs nucléaires* a été développée dans le cadre de la thèse d'Idir Khemmoudj. Il s'agit d'un problème d'ordonnancement couplé à un problème de planification de la production (contrat EDF-LIPN). Dans un premier temps, la contribution de la programmation logique par contraintes à la résolution de ce problème a été évaluée. Un système combinant la programmation par contraintes et la recherche locale pour construire des plans d'arrêts avec un meilleur profit a ensuite été développé.

Veille SNCF Le contrat d'étude et de veille sur les approches de modélisation et techniques de résolution exploitables pour optimiser la planification de ressources de production mobiles entre la SNCF-DI&R et les membres de l'ACI Jeunes Chercheurs vise à recenser les formalisations existantes (et récentes) qui sont envisageables pour les problèmes de planification de ressources mobiles dans le domaine des transports, ainsi que la complexité de ces modélisations. Un panorama comparatif des approches de résolution existantes (et récentes) pour ces modèles (approches standards ou dédiées, optimales vs approchées) est dressé. Enfin, les performances numériques de certaines approches sont évaluées, à partir notamment de problèmes concrets fournis par la SNCF-DI&R.

• Dissémination

L'année 2005 a été l'occasion de rédiger trois chapitres sur les problèmes de sac-à-dos, de transport aérien et du voyageur de commerce, dans un ouvrage consacré à l'optimisation combinatoire [CL-26, CL-31, CL-29].

Invitations, tutoriels

- Anass Nagih : Conférencier-invité par la Direction de la Stratégie de la SNCF, la Direction des Opérations et la Direction de l'Innovation et de la Recherche dans le cadre du projet "Mieux concevoir pour mieux produire" ; titre "Planification des ressources humaines", siège de la SNCF Paris (17 juin 2003).
- Gérard Plateau : Tutoriel pour les Journées Franciliennes de Recherche Opérationnelle ; titre "Sur le sac à dos en variables 0-1", Paris (13 janvier 2006).

Actions de vulgarisation de la recherche

- Lucas Létocart : Présentation d'une démonstration "Courses et jeux de fourmis" à la manifestation *Savante Banlieue*, Villetaneuse, 13-14 octobre 2005.

- **Collaborations**

Collaborations universitaires

1. Les collaborations d'Anass Nagih avec François Soumis et Yassir Miladi (école Polytechnique de Montréal et GERAD) s'articulent principalement autour de :
   - la construction de rotations d'équipages en planification opérationnelle ;
   - les heuristiques lagrangiennes pour le problème de plus court chemin contraint ;
   - la réduction de l'espace des états dans un algorithme de type programmation dynamique.
2. Une collaboration entre les membres de l'ACI Jeunes Chercheurs et Jeunes Chercheuses et le GERAD a été mise en place sous la forme d'invitations réciproques d'enseignants-chercheurs.
3. Dans le cadre de la thèse de Jean-Luc Lemalade, financée par une bourse de coopération franco-malgache, en partenariat avec le CIRAD Montpellier et le FOFIFA Antananarivo, une étude porte sur l'optimisation de la gestion des ressources agricoles en milieu forestier tropical, dans le but d'éviter l'extension des surfaces à cultiver en gérant les parcelles en milieu de façon originale (cloisonnement, bandes anti-érosion, utilisation optimale des jachères et rotation des cultures) menant à un équilibre entre l'espace agricole et l'espace forestier [CO-114].
4. L'application suivante a été développée par Hachemi Bennaceur et Idir Khemmoudj : *Contraintes dans les bases de données* (en collaboration avec F. Boufares).
   L'étude de la validité d'un schéma relationnel « Entité-Association » d'une base de données repose sur la vérification de la cohérence du système de contraintes associé à ce schéma. Plus précisément, elle a porté sur différents types de contraintes (cardinalités, dépendance fonctionnelles, rôle, etc.) et sur les méthodes pour résoudre les systèmes de contraintes permettant de décider si une base de données est instanciable. Une extension de ces résultats à la modélisation UML a également été proposée [RI-18].
5. évaluation des performances dans le projet CNRS-PAN (Pol 16684) (coopération franco-polonaise). Projet de coopération bilatérale CNRS-PAN : *Analyse transitoire des réseaux à service intégrés par l'approximation de diffusion.* 2004-2005. Fehran Perkergin.
6. Projet de coopération bilatérale CNRS-PAN (Pol 16684) : *Processus de diffusion à sauts multiples en modélisation des systèmes de télécommunications.* 2006. Fehran Perkergin.

OCAD

3.3.5 Formation à la recherche

Thèses soutenues

1999-2003 Babacar Thiongane : *Ré-optimisation dans le dual lagrangien du biknapsak en variables 0-1*
Financement de la thèse : Allocation de Recherche du Ministère
Soutenue le 24 février 2003
Directeurs : Gérard Plateau et Anass Nagih
Situation actuelle du diplômé : Post-doc au Canada.

2001-2005 Irène Loiseau : *La génération de colonnes en nombres entiers*
Soutenue le 29 juin 2005
Directeurs : Gérard Plateau et Nelson Maculan
Situation actuelle du diplômé : Enseignant Chercheur à l'Université de Buenos Aires, Argentine.

2002-2006 Nicolas Marcos : *Modélisation et Optimisation de la Gestion du Matériel Roulant à la SNCF*
Soutenue le 23 novembre 2006
Directeurs : Gérard Plateau, Anass Nagih et Laurent Alfandari
Situation actuelle du diplômé : Chef de projets de recherche à la SNCF, Direction de l'Innovation et de la Recherche, Paris.

2002-2007 Mohand Ou Idir Khemmoudj : *Modélisation et résolution de systèmes de contraintes : Application au problème de placement des arrêts et de la production des réacteurs nucléaires d'EDF*
Soutenue le 04 avril 2007
Directeurs : Hachemi Bennaceur et Marc Porcheron
Situation actuelle du diplômé : Post-doc EDF-LIPN, Paris.

2002-2007 Fayçal Djerourou : *Contribution à la résolution complète et incomplète du problème de satisfaction de contraintes pondérées*
Soutenue le 15 juin 2007
Directeur : Hachemi Bennaceur
Situation actuelle du diplômé : Ingénieur en informatique, Paris.

HDR soutenues

2003 Alain Faye : *Contributions à la programmation quadratique en 0-1*
Soutenue le 17 décembre 2003
Coordinateur : Gérard Plateau
Situation actuelle du diplômé : Maître de Conférences HDR à l'IIE-CNAM.

2004 Anass Nagih : *Contributions à la résolution des problèmes en nombres entiers : réoptimisation, problèmes de grande taille et planification en transport*
Soutenue le 14 décembre 2004
Coordinateur : Gérard Plateau
Situation actuelle du diplômé : Professeur à l'Université de Metz.

2007 Frédéric Roupin : *Algorithmes Combinatoires et Relaxations par Programmation Linéaire et Semidéfinie. Application à la Résolution de Problèmes Quadratiques et d'Optimisation dans les Graphes.*
Soutenue le 24 novembre 2006
Coordinateur : Gérard Plateau
Situation actuelle du diplômé : Maître de Conférences HDR au CNAM.

Doctorants

nom	date	position actuelle pour thèses
O. Laval	10/2005-10/2008	Allocataire - moniteur
J.-L. Lemalade	10/2002-12/2006	bourse coopération franco-malgache
A. Rasoanaivo	10/2006-10/2009	bourse AUF
J. Sadki-Fenzar	10/2006-10/2009	Allocataire - monitrice
N. Touati	10/2005-10/2008	Allocataire - monitrice

OCAD

Participations à des écoles (printemps, été, hiver, automne)

- Olivier Laval : Séminaire de Mathématiques supérieures (SMS) : Optimisation combinatoire : Méthodes et applications, Département de mathématiques et de statistique (DMS) de l'Université de Montréal, 19-30 juin 2006.
- Nora Touati : Ecole printanière sur les tournées de véhicules, HEC Montréal, 10-12 mai 2006.
- Nora Touati : GERAD Summer School : Column generation, Ecole Polytechnique de Montréal, 15-19 mai 2006.

Cours de DEA / Master 2 recherche

- 2003 à 2004. (L. Létocart, A. Nagih) : *Méthodes Avancées en Optimisation Discrète*, DEA IA-OC, Paris 8 - Paris 13.
- 2004 à 2005. (L. Létocart, A. Nagih, S. Toulouse) : *Méthodes Avancées en Optimisation Discrète*, DEA IA-OC, Paris 8 - Paris 13.
- 2003 à 2005. (G. Plateau) : *Optimisation en nombres entiers*, DEA IA-OC, Paris 8 - Paris 13.
- 2005 à 2007. (L. Létocart, A. Nagih, S. Toulouse) : *Méthodes Avancées en Optimisation Combinatoire*, Master Mathématiques-Informatique 2^{e}année recherche, Paris 13.
- 2005 à 2007. (L. Alfandari, L. Létocart) : *Aide à la Décision*, Master Mathématiques-Informatique 2^{e}année recherche, Paris 13, et Master 2^{e}année Modélisation Informatique des Connaissances et du Raisonnement, Paris 13.
- 2005 à 2007. (G. Plateau) : *Optimisation combinatoire*, Master Mathématiques-Informatique 2^{e}année recherche, Paris 13. Option du Master Mathématiques-Informatique-Technologies 2^{e}année recherche (*Mathématiques et Applications au codage et à la cryptologie*), Paris 8.
- 2005 à 2007. (Sophie Toulouse) : *Optimisation Combinatoire*, DEA IGMA (*Informatique et Génie Mathématique*), Université d'Antananarivo (Madagascar).

Stages de DEA / Master 2 recherche

- 2003 (A. Nagih, G. Plateau, S. Lenestour (Air France)) : T. Soulimane (DEA IA-OC Paris 13, Air France) : *Application des métaheuristiques au problème de reconstruction des rotations d'équipages*
- 2003 (A. Nagih, G. Plateau, D. De Almeida (SNCF)) : S. Nicolas (DEA IA-OC Paris 13, SNCF) : *Optimisation de la gestion des lignes à voie unique*
- 2003 (A. Nagih, G. Plateau) : A. Birri (DEA IA-OC Paris 13, GDF) : *élaboration d'une méthode de dimensionnement des réseaux de transport de gaz*
- 2004 (A. Nagih, D. De Almeida (SNCF)) : A. Quivet (DEA IA-OC Paris 13, SNCF) : *Approches de résolution pour des modèles d'optimisation de la gestion des locomotives : étude et implantation*
- 2004 (A. Nagih, G. Plateau) : E. Dumont (DEA IA-OC Paris 13, Eurodécision) : *Optimisation de la diversité des cablages dans le domaine de l'automobile*
- 2005 (L. Létocart, A. Nagih) : N. Touati (DEA IA-OC Paris 13, LIPN) : *Ré-optimisation au sein d'un schéma de génération de colonnes*
- 2005 (L. Létocart, A. Nagih, G. Plateau) : R. Ouaras (DEA 103 Paris 9-Dauphine, LIPN) : *Ré-optimisation du sac à dos quadratique*
- 2005 (L. Létocart, A. Nagih) : G. Elain (DEA IA-OC Paris 13, Thalès) : *Modèles réalistes de placement d'acteurs virtuels*
- 2005 (A. Nagih, D. De Almeida (SNCF)) : J. Tea (DEA IA-OC Paris 13, SNCF) : *Métaheuristiques pour des modèles d'optimisation de la gestion des locomotives : étude et implantation*
- 2005 (L. Alfandari, A. Nagih, A. Plateau (CNAM)) : S. Djabouabdallah (DEA IA-OC Paris 13, LIPN) : *étude de méthodes de résolution approchée des problèmes de planification*
- 2006 (L. Létocart, A. Nagih, G. Plateau) : T. Radakovic (Master MMI2 Paris 13, LIPN) : *Ré-optimisation du sac à dos quadratique*
- 2006 (L. Alfandari, A. Nagih, A. Plateau (CNAM)) : J. Sadki-Fenzar (Master MMI2 Paris 13, LIPN) : *Méthodes de décomposition pour la résolution de problèmes de planification*
- 2006. (V. Ravelomanana, S. Toulouse) : H. Razafisambatra (DEA IGMA Tananarive) : *Approximation de problèmes définis sur des collections d'ensembles.*
- 2006. (V. Ravelomanana, S. Toulouse) : S. Djabali (P6-UPMC, LIPN) : *Approximation en moyenne.*
- 2007 (L. Létocart, M-C. Plateau, G. Plateau) : A. Benouali (Master MMI2 Paris 13, LIPN) : *Méthodes lagrangiennes pour la résolution du sac-à-dos quadratique avec contrainte de cardinalité*
- 2007 (L. Létocart, N. Touati) : Z. Beddar (Master MMI2 Paris 13, LIPN) : *Résolution de modèles d'optimisation de grande taille : méthodes de décomposition et algorithme de génération de colonnes*

3.3.6 Publications principales

Chapitres de livres

[ocad :MT07d] Jérôme Monnot et Sophie Toulouse. *Complexity and approximation results for bounded-size paths packing problems*, *Combinatorial Optimization and Theoretical Computer Science : Interfaces and Perspectives*, Chapitre 18. ISTE (V. Paschos ed.), 2007.

[ocad :MT07a] Jérôme Monnot et Sophie Toulouse. *Le voyageur de commerce et ses variations : un tour d'horizon de ses résolutions*, *Optimisation Combinatoire 5 : problèmes paradigmatiques et problématiques nouvelles*, Chapitre 2. Hermès (V. Paschos ed.), 2007.

[ocad :PN07] Gérard Plateau et Anass Nagih. *Le problème du sac à dos en variables 0-1*, *Optimisation combinatoire 4 : problèmes paradigmatiques*, Chapitre 6, pages 163–190. Hermès (V. Paschos ed.), Paris, 2007.

[ocad :AN06] Laurent Alfandari et Anass Nagih. *Optimisation de la construction de rotations déquipages en transport aérien*, *Optimisation combinatoire 3 : applications*, Chapitre 1. Hermès (V. Paschos ed.), Paris, 2006.

Articles dans des revues internationales avec comité de lecture sélectif

[ocad :A07] Laurent Alfandari. Improved approximation of the soft-capacitated facility location problem. *RAIRO Operations Research*, 41 :83–93, 2007.

[ocad :ALMP06] Rafael Castro De Andrade, Abdel Lisser, Nelson Maculan et Gérard Plateau. Enhancing a branch-and-bound algorithm for two-stages stochastic integer problems. *Management Science*, 52(9) :1450–1455, 2006.

[ocad :TNP06] Babacar Thiongane, Anass Nagih et Gérard Plateau. Lagrangean heuristics combined with reoptimization for the 0-1 biknapsack problem. *Discrete Applied Mathematics*, 154 :2200–2211, 2006.

OCAD

[ocad :ALMP05] Rafael Castro De Andrade, Abdel Lisser, Nelson Maculan et Gérard Plateau. B & b frameworks for the capacity expansion of high speed telecommunication network under uncertainty. *Annals of Operations Research*, 140(1) :49–65, 2005.

[ocad :CLR05] M-C. Costa, L. Létocart et F. Roupin. Multicut and integral multiflow : a survey. *European Journal of Operational Research*, 162(1) :55–69, 2005.

[ocad :CP05b] Tadeusz Czachorski et Ferhan Pekergin. A novel diffusion process with jumps to study an electronic-optical edge router. *Archives of Theoretical and App. Informatics*, 17(2) :76–88, 2005.

[ocad :TNP05] Babacar Thiongane, Anass Nagih et Gérard Plateau. Adapted step size in a 0-1 biknapsack lagrangean dual solving algorithm. *Annals of Operations Research*, 139 :353–373, 2005.

[ocad :ALMP04] Rafael Castro De Andrade, Abdel Lisser, Nelson Maculan et Gérard Plateau. Telecommunication network capacity design for uncertain demand. *Computational Optimization and Applications*, 29(2) :127–146, 2004.

[ocad :B04] Hachemi Bennaceur. A comparison between CSP and SAT techniques. *Constraints : an International Journal*, 9(2) :123–138, 2004.

[ocad :BB04a] F. Boufarès et H. Bennaceur. Consistency problems in er-schemas for database systems. *Information Sciences Journal*, 163(4) :263–274, 2004.

Articles dans des revues nationales ou internationales

[ocad :MPL03] Nelson Maculan, Gérard Plateau et Abdel Lisser. Integer linear models with a polynomial number of variables and constraints for some classical combinatorial problems. *Pesquisa Operacional*, 23(1) :161–168, 2003.

[ocad :TNP03a] Babacar Thiongane, Anass Nagih et Gérard Plateau. Analyse de sensibilité pour les problèmes linéaires en variables 0-1. *RAIRO-Operations Research*, 37(4) :291–309, 2003.

Communications dans des conférences internationales avec comité de lecture

[ocad :KB06b] M. I. Khemmoudj et H. Bennaceur. Clique inference process for solving Max-CSP. *12th International Conference on Principles and Practice of Constraint Programming (Nantes, 2006)*, pages 246–250. Springer- Berlin Heidelberg, 2006.

[ocad :KBN05] M. I. Khemmoudj, H. Bennaceur et Anass Nagih. Combining arc-consistency and dual lagrangean relaxation for filtering CSPs. *2th International Conference on Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems (Praha, 2005)*, pages 258–272. Springer - Berlin Heidelberg, 2005.

Workshops et autres conférences avec comité de lecture

[ocad :LNP07] Lucas Létocart, Anass Nagih et Gérard Plateau. Re-optimization in the resolution of the 0-1 quadratic knapsack problem. *EURO XXII*, Prague, République tchèque, 2007.

[ocad :LPP07] Lucas Létocart, Marie-Christine Plateau et Gérard Plateau. New bounds for the 0-1 exact k-item quadratic knapsack problem. *ECCO XX*, Limassol, Chypre, 2007.

[ocad :TLN07] Nora Touati, Lucas Létocart et Anass Nagih. Diversification des solutions et réoptimisation pour la résolution de problèmes de plus court chemin avec contraintes de ressources dans un schéma de génération de colonnes. *ROADEF 2007*, pages 419–420, Grenoble, France, 2007.

[ocad :LANP06] Jean Luc Lemalade, Laurent Alfandari, Anass Nagih et Gérard Plateau. Un modèle de planification des rotations culturales dans un contexte de lutte contre la déforestation à madagascar. *7e congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision*, Lille, 2006.

[ocad :CPS05a] Tadeusz Czachorski, Ferhan Pekergin et Bartlomiej Szady. Forming optical packets, a diffusion approximation and simulation model. *Proceedings of EuroNGI (European New Generation Internet) 2nd Workshop IA.8.1*, Portugal Aveiro, 2005.

[ocad :MAGQNAP05] Nicolas Marcos, David De Almeida, Daniel Gauyacq, Alexandre Quivet, Anass Nagih, Laurent Alfandari et Gérard Plateau. Un outil générique d'aide à la décision pour la gestion des locomotives fret à la SNCF. *6e congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision*, pages 295–312, Tours, 2005.

OCAD

Travaux en cours de soumission

[ocad :BCLR06] C. Bentz, M-C. Costa, L. Létocart et F. Roupin. Multicut and integral multiflow in rings. RR 1050, Centre d'Etude et de Recherche en Informatique du CNAM, Conservatoire National des Arts et Métiers, 2006. Soumis à EJOR.

3.4 Composante Algorithmique Distribuée

Participants : T. Andriamampinina, C. Banderier, F. Butelle, C. Cérin, J-C. Dubacq, L. Finta, M. Hakem, H. Kenniche, C. Lavault, G. Paillard, A. Rasoanaivo, V. Ravelomanana, S-M. Sedjelmaci, M.Valencia-Pabon

3.4.1 Thèmes de recherche

L'algorithmique distribuée existe au sein de l'équipe OCAD depuis presque quinze ans. Quant aux méthodes d'étude, à la conception et à l'analyse des algorithmes distribués, elles ont bien évolué durant cette période. La mathématisation du domaine s'est manifestée tout particulièrement par l'utilisation d'outils et de méthodes d'analyse sans cesse plus raffinés, entre autres dans les mesures de complexité en temps et en messages.

La recherche et l'analyse d'algorithmes, c'est-à-dire de procédures effectives permettant de résoudre des problèmes ou bien de calculer la valeur d'une fonction en un nombre fini d'étapes, est particulièrement active autour des thèmes suivants :

- les problèmes de communication dans les réseaux d'interconnexion ;
- la modélisation et l'étude de problèmes issus des systèmes distribués ;
- l'analyse en moyenne de classes générales d'algorithmes et de protocoles distribués ;
- l'étude des algorithmes probabilistes,
- l'étude de propriétés fondamentales de familles d'objets et structures combinatoires : arbres, graphes, hypergraphes, permutations, modèles d'urnes, etc.

Les outils utilisés sont généralement issus des mathématiques classiques (analyse complexe, équation différentielles, asymptotique, probabilités) et les *séries génératrices* en constituent la « brique de base » par excellence. À partir des séries génératrices, de nombreuses techniques analytiques peuvent être mises en œuvre, allant de l'approximation asymptotique (réelle et complexe) aux méthodes stochastiques. Les résultats fournissent alors des informations très précises sur le comportement qualitatif ou quantitatif des systèmes étudiés (cf. Ph. Flajolet, D.E. Knuth, R. Motwani, etc.). Ce progrès vers des techniques rodées en mathématiques discrètes — la *combinatoire analytique*, la théorie de graphes et les *probabilités* — s'est trouvé accéléré ces dernières années par l'arrivée dans l'équipe de jeunes chercheurs spécialistes du domaine (Vlady Ravelomanana, Cyril Banderier, Mario Valencia-Pabon).

OCAD

• Conception et analyse d'algorithmes distribués

Les systèmes distribués se caractérisent par un processus de circulation et d'échange de l'information entre entités autonomes communicantes concourant à la réalisation d'un objectif commun. Le développement d'outils et de méthodes algorithmiques appropriées (combinatoire analytique, méthodes stochastiques, asymptotique complexe, etc.) permet d'analyser plus finement de nombreux algorithmes distribués, en particulier en moyenne. En retour, ces méthodes permettent de concevoir de nouvelles classes d'algorithmes distribués, dont la complexité en communication et/ou en temps est notablement améliorée. Les classes de réseaux (graphes) considérés possèdent des structures (topologies) très diverses et peuvent jouir de caractéristiques variées : réseaux synchrones ou asynchrones, anonymes ou avec identités, fautifs ou sûrs, fixes ou mobiles, information structurelle, etc.

Outre la recherche et la spécification (plus ou moins fouillées) de l'algorithme lui-même, les résultats portent prioritairement sur l'analyse de la complexité (au pire et en moyenne) de ces algorithmes (probabilistes ou non). Plus rarement, cette analyse de complexité est poussée plus loin : calcul de tous les moments, preuve d'existence et mise en évidence d'une distribution limite, etc. (cf. § 3.4.3).

Classes générales d'algorithmes distribués

De vastes *classes d'algorithmes distribués* sont conçues et analysées dans cet axe de recherche. Ces algorithmes (déterministes ou probabilistes) s'appliquent à divers problèmes cruciaux du distribué, dans un très grand nombre de topologies de réseaux et sous diverses contraintes matérielles et temporelles. Franck Butelle et Christian Lavault étudient le problème de l'*élection anonyme* dans les réseaux *quelconques* (plus particulièrement les anneaux) et celui de la recherche

d'un *arbre optimal avec contraintes* (de diamètre ou de degré minimaux) dans un graphe *quelconque*. Un grand nombre de *protocoles d'ordonnancement dans des réseaux quelconques* (sous diverses contraintes matérielles ou temporelles) sont l'objet des travaux de Franck Butelle, Mourad Hakem et Lucian Finta. Enfin, de nouvelles méthodes pour construire des classes d'algorithmes distribués ont été utilisées par Christian Lavault et Gabriel Paillard : *inversions de liens d'un réseau acyclique orienté*, *inversions d'arcs d'un (multi)graphe quelconque*, etc. (cf. § 3.4.3).

Dans un univers réparti, l'ordonnancement est un problème de contrôle qui peut être pris en charge par la coopération des processus communicants situés sur chaque site actif. L'impossibilité d'avoir un état global cohérent d'un système totalement réparti fait que la plupart des algorithmes proposés ne permettent pas d'avoir une solution optimale. Seules des heuristiques peuvent être utilisées, comme celle proposée par Franck Butelle et Mourad Hakem (cf. 3.4.2). Dans la littérature courante, les hypothèses formulées pour la conception d'algorithmes d'ordonnancement distribué temps réel s'avèrent trop simplificatrices. En général, ils ne traitent en effet que de tâches indépendantes et ils supposent que les coûts de migration de celles-ci ne sont pas trop onéreux. De fait, le problème relevant de la prise en compte des contraintes de précédence réparties demeure non résolu. C'est seulement grâce à des hypothèses réalistes qu'il est possible de concevoir un algorithme d'ordonnancement distribué temps réel (cf. 3.4.2).

Protocoles de communication dans les réseaux sans fil

L'algorithmique des réseaux sans fils se situe à la frontière de l'algorithmique distribuée et de la théorie des graphes. L'environnement (structures, caractéristiques) y est supposé *quelconque* (anonyme, synchrone ou asynchrone, tolérant aux fautes, etc.). C'est un champ d'investigation très fructueux, commun aux deux premiers axes de recherche du thème : la conception et l'analyse d'algorithmes distribués et l'étude de structures ou objets combinatoires à dominante aléatoire.

L'intérêt porté aux réseaux sans fils (réseaux mobiles ad hoc, capteurs, etc.) s'est considérablement accru ces dernier temps. Ceux-ci sont en effet utilisés en cas de sinistre (tremblement de terre, incendie, sauvetage, etc.) ou dans le cadre d'applications spécifiques diverses, militaires en particulier (« concret » et lucratif !). Contrairement aux réseaux mobiles cellulaires, qui nécessitent une infrastructure rigide, les *réseaux mobiles ad hoc* (ou MANET) et les *capteurs* doivent pouvoir s'utiliser très souplement et s'auto-organiser très rapidement pour fournir des services locaux spécifiques, souvent en urgence. Les capteurs sont ainsi des entités identiques minuscules à énergie (durée de vie des batteries) très limitée qui sont déployées massivement sur une région bien déterminée (du plan ou de l'espace) afin d'y mesurer des données locales (température, pression, hygrométrie, hydrométrie, mouvements divers, etc.) et les communiquer.

Un MANET est un *système distribué non centralisé* à *topologie variable*. Aucune unité privilégiée n'est en mesure de communiquer avec toutes les stations (anonymes). La situation et l'évolution globale du réseau leur reste donc inconnue. Le modèle de réseau considéré est « à sauts multiples » et un MANET se modélise par un graphe G non orienté, où une arête entre deux sommets (stations) indique que ceux-ci peuvent communiquer directement. Tous les problèmes suscités par un MANET G (structure, communication, mobilité, sécurité, etc.) supposent que certaines propriétés structurelles de G soient *a priori* connues de chaque station. La plupart des protocoles de contrôle fondamentaux, comme la diffusion ou l'élection, utilisent ce type d'hypothèse initiale. C'est ainsi qu'une diffusion dans G suppose, en général, la connaissance par les stations de certaines de ses caractéristiques exactes ou partielles (majorées) : son nombre de nœuds n, son degré maximum Δ, son diamètre D, etc.

L'objectif de cet axe de recherche est d'étudier et d'analyser de manière assez systématique les principaux *protocoles de contrôle* pour la communication dans les MANETS. Christian Lavault et Vlady Ravelomanana étudient les algorithmes probabilistes de *diffusions*, d'*initialisation*, d'*élection*, de *routages*, de *k-clustering*, d'*exclusion mutuelle*, de *tri*, etc. L'« initialisation » consiste à attribuer aléatoirement des identités distinctes à chaque entité (radio, capteur, etc.) du réseau ; elle s'apparente à une attribution de « codes » uniques (largeurs de bandes passantes, coloriage, (ré)étiquetage, etc.).

Par ailleurs, le modèle des *graphes aléatoires géométriques* (cf. 3.4.1) permet à Vlady Ravelomanana d'obtenir des mesures précises sur les propriétés générales et extrêmales de ces réseaux : taille, composantes, transitions de phase, régimes, etc.

• Systèmes distribués et algorithmique parallèle

Grid Computing

L'arrivée de Christophe Cérin en 2005, permet à la conception d'algorithmes parallèles ou distribués de bénéficier d'une nouvelle impulsion, dans un cadre moins théorique et mieux adapté aux besoins en développement : le *Grid Computing*.

L'évolution du calcul parallèle s'est progressivement orienté ces dernières années vers l'utilisation de composants standards. Ainsi est apparu le calcul sur *réseaux de stations de travail* (ou *Cluster Computing*). Ces nouvelles architectures parallèles atteignent des puissances de calcul de l'ordre du Téraflop. On est capable d'interconnecter plus de mille PC entre eux et de les faire communiquer par des interfaces répandues (de type Gigabit ETHERNET) ou spécialisées (de type MYRINET).

Depuis cinq ans est apparue la notion de *grille*, qui vise à mutualiser les ressources, par exemple celles des PC des grands centres de recherche ou des internautes, connectés à Internet. Les grilles sont donc des environnements logiciels qui autorisent les applications à intégrer, à afficher, à instrumentaliser des ressources qui sont détenues par des organismes ou des individus situés dans différentes zones géographiques. Les paradigmes *grappe* et *grille* ont en commun le souci de partager des ressources de calcul, de stockage dans un but d'économie des moyens mis en œuvre.

Le projet d'ANR *SafeScale* (*Security And Fault-tolerance to Exploit Safety ambient Computing in lArge scaLe Environments* – `https://www-lipn.univ-paris13.fr/safescale`) a été retenu fin 2005. Il est piloté par le LIPN avec pour partenaires l'ENSTB, l'Irisa, le LIG, et l'université Joseph Fourier.

Le LIPN s'est positionné sur des aspects applicatifs. Il propose d'étudier deux applications : le partitionnement de données et l'évaluation (détermination) de polynômes. Pour ces deux applications, il s'agit de distribuer des calculs et des données sur Internet et de vérifier le calcul produit en le comparant avec des résultats produits sur des sites de confiance.

OCAD

Pour le partitionnement de données, il s'agit de le faire par des techniques de vol de cycles plutôt que par des techniques statiques qui prédéterminent ce qu'il faut envoyer sur chacun des sites. Pour l'évaluation de polynômes, il s'agit d'implémenter dans un premier temps la méthode actuelle qui se prête bien à une distribution à large échelle et qui a été développée en particulier par Frédéric Toumazet de l'équipe LCR et membre de *SafeScale*. Dans un deuxième temps on greffera une certification des résultats mais il y a déjà un travail important à fournir pour construire l'architecture matérielle de calcul. Elle se fera très certainement à base d'EXTREMWEB (`http://www.xtremweb.net`).

PGCD de deux entiers

Les travaux de recherche de Sidi Mohamed Sedjelmaci ont permis, quant à eux, de maintenir un très bon niveau de publications internationales en arithmétique parallèle (étude et analyse d'*algorithmes parallèles du PGCD de deux entiers*).

Le PGCD est une opération arithmétique fondamentale dont les applications sont multiples. La plupart des logiciels scientifiques utilisent l'algorithme séquentiel polynomial de Schönhage, de complexité en temps $O\left(n(\log n)^2 \log\log n\right)$ pour des entiers de n bits. En calcul parallèle, on ignore si ce problème est dans la classe NC de P ou bien P-complet (la classe la plus « difficile » de P). Deux directions de recherche s'imposent donc : tenter une démonstration directe de l'appartenance du PGCD à la classe P-complet ou bien exhiber un algorithme parallèle de PGCD appartenant à NC.

• Algorithmique, combinatoire et complexité

Parmi les objets ou les structures les plus utilisés, on peut citer les mots, les arbres, les graphes et hypergraphes, les cartes, les chemins, les permutations, les partitions, etc. Parmi les exemples classiques d'opérations ou de manipulations, citons les bijections, les codages, l'énumération, la génération aléatoire, la visualisation, le coloriage et le (ré)étiquetage.

Les problèmes combinatoires suscités par l'algorithmique peuvent être de nature diverse :

- l'énumération et comptage d'objets permettant, en particulier, d'effectuer des analyses exactes et en moyenne d'algorithmes,

- la représentation et la manipulation efficaces d'objets mathématiques variés (graphes, nœuds, groupes, séries formelles, etc.),
- les calculs explicites sur ces objets,
- la mise à jour de correspondances entre ces objets à travers des bijections,
- la génération aléatoire, qui a pour objectif d'engendrer des objets combinatoires, non pas de façon exhaustive mais selon une loi de probabilité donnée ;
- les études purement combinatoires des propriétés de certains objets ou familles d'objets comme les mots, les arbres, les chemins ou bien encore les (hyper)graphes.

Propriétés de structures et de familles d'objets combinatoires

Les méthodes de la combinatoire analytique permettent de résoudre quantité de problèmes soulevés en algorithmique. Trois directions de recherche prédominent au sein de la composante AD.

- D'une part, l'étude par Vlady Ravelomanana et Christian lavault des propriétés générales et extrêmales des graphes et hypergraphes aléatoires $\mathbb{G}(n,p)$, $\mathbb{G}(n,r)$ ou $\mathbb{H}(n,p)$ et de leurs algorithmes (processus d'apparition de composantes, taille de composantes, transitions de phase, régimes, couplages, etc.).
- D'autre part, l'analyse par Cyril Banderier de la combinatoire des marches aléatoires sur $\mathbb{Z}$ ou $\mathbb{N}$, de chemins dirigés dans le plan (énumération en moyenne, asymptotique, etc.), en lien avec l'évolution d'une pile ou d'une file d'attente.
- Enfin, la théorie de la coloration des graphes. Le partitionnement d'un ensemble d'objets en classes vérifiant certaines règles est un processus fondamental pour beaucoup de problèmes combinatoires. Un ensemble de règles simples peut nous permettre de savoir pour chaque couple d'objets si ils peuvent ou non appartenir à la même classe : cette problématique concerne en particulier la théorie de la coloration des graphes. Une des directions de recherche de Mario Valencia-Pabon, Vlady Ravelomanana et Christian lavault consiste à étudier certains généralisations de problèmes de colorations des sommets ou des arêtes d'un graphe que l'on trouve souvent dans des problèmes d'ordonnancement de tâches.

Mesure de complexité : l'analyse perturbative (*Smoothed Analysis*)

Il s'agit d'une mesure de complexité intermédiaire entre le pire des cas et la complexité en moyenne qui semble adéquate pour quantifier la « difficulté » d'un problème. Il paraît ainsi naturel de s'intéresser aux nombreux problèmes classés NP ou assimilés (d'optimisation et de combinatoire).

En effet, être « worst-case-NP » est fort différent d'être « average-case-NP », lui-même fort différent d'être « average-case-NP sur une région suffisamment grande ». C'est de ce dernier point dont on a besoin, par exemple, pour avoir un cryptosystème fiable : le pire des cas dans le *paysage de complexité* de l'algorithme, ne doit pas être un pic isolé mais un plateau. Comment quantifier et étudier cette notion de « problème difficile sur une grande région » ? C'est ici qu'intervient le concept de *smoothed analysis* (traduit par *analyse perturbative*) : où sont les instances difficiles ? Plus précisément, la complexité perturbée est ainsi définie comme le maximum sur les entrées du temps moyen mis par l'algorithme à s'exécuter sur les perturbations de l'entrée. *L'analyse perturbative* étudie donc la complexité d'un algorithme (ou d'une valeur particulière) autour d'une instance perturbée. Comment se modifie le paysage de complexité d'un algorithme lorsque l'on perturbe les instances d'entrée autour d'éventuelles régions « difficiles » ? L'analyse perturbative mesure la performance moyenne d'un algorithme autour d'une instance, permettant ainsi de déceler ces régions difficiles (cf. § 3.4.2 – résultats de Cyril Banderier).

OCAD

3.4.2 Résultats significatifs

• Conception et analyse d'algorithmes distribués

Classes générales d'algorithmes distribués

Dans ce domaine particulièrement vaste, de nombreux résultats significatifs ont été obtenus par Franck Butelle, Lucian Finta, Mourad Hakem (ordonnancement) et par Franck Butelle, Christian Lavault, Mario Valencia-Pabon (arbres optimaux contraints). On peut ainsi citer un

grand nombre de *protocoles d'ordonnancement dans les réseaux quelconques* (parallèles ou distribués et sous diverses contraintes matérielles ou temporelles) [RI-45, CI-161, CI-45], plusieurs bons algorithmes de recherche d'un *arbre optimal avec contraintes* (de diamètre ou de degré minimaux) dans un graphe *quelconque* [RI-16, RI-84], l'analyse poussée d'une *élection anonyme* dans les anneaux [RI-51], et de nouvelles méthodes de conception de classes d'algorithmes distribués : *inversions de liens d'un réseau acyclique orienté*, *inversions d'arcs d'un (multi)graphe quelconque* [CI-70], etc.

Grâce à des hypothèses plus réalistes (réseau arbitrairement grand et totalement distribué, groupes de tâches sporadiques dépendantes caractérisées par leurs échéances temporelles, pouvant apparaître sur n'importe quel site et à n'importe quel moment, etc.), Franck Butelle, Lucian Finta et Mourad Hakem ont récemment proposé un algorithme original d'ordonnancement distribué temps réel [CI-161].

Protocoles de communication dans les réseaux sans fil

Dans le domaine des *protocoles de contrôle* pour la communication dans les réseaux sans fil, les résultats les plus significatifs se centrent autour d'études et d'analyses assez systématique des algorithmes probabilistes classiques du domaine (Christian Lavault et Vlady Ravelomanana). Plus précisément, la *diffusion* [RI-71, IN-7, RI-64], l'*initialisation* [CI-49, IN-7, RI-64], l'*élection* [RI-71], le *k-clustering* [CI-63], l'*exclusion mutuelle*, le *routage* [**?**, RI-64], etc. Il faut noter de plus que la plupart des protocoles que nous avons proposés sont « efficaces en énergie », i.e. ils consomment très peu d'énergie. Dans [RI-71], par exemple, les stations ne restent éveillées qu'au plus $O(\log\log(n))$ unités de temps pour une élection de complexité en temps $O(\log(n))$.

Sur les graphes géométriques aléatoires pouvant servir de support de modélisation aux réseaux sans fil, Vlady Ravelomanana [RI-9] a quantifié les degrés des sommets ainsi que les taux de couverture durant le régime supercritique. Ces résultats ont permis de concevoir des protocoles d'initialisation [RI-64] lors des phases de déploiements aléatoires des stations de ces réseaux.

OCAD

• Systèmes distribués et algorithmique parallèle

Grid Computing

Christophe Cérin et Jean-Christophe Dubacq ont obtenu récemment un résultat significatif d'une partie des objectifs de leurs recherches dans l'ANR *SafeScale* [CI-123], dont le LIPN est porteur. Ce travail propose des méthodes exactes et approchées pour déterminer les quantités de données à envoyer à chaque processeur d'un grappe de processeurs à partir d'une configuration donnée. Le schéma de partitionnement, réalisé dans un contexte où les processeurs sont hétérogènes, tient compte de la complexité (connue ou estimée) du traitement qui le suit.

PGCD de deux entiers

Le meilleur algorithme parallèle proposé à ce jour demeure celui de Chor et Goldreich (1990) de complexité $O(n/\log n)$ sur un modèle de PRAM *CRCW* avec $n^{1+\epsilon}$ processeurs ($\epsilon > 0$ constante quelconque). Les résultats les plus récents dans ce domaine utilisent des techniques de « réduction$k$-aire » (algorithmes « accélérés » de Jebelean et Weber). En particulier, l'amélioration de l'algorithme de Weber par Mohamed Sedjelmaci permet de proposer plusieurs algorithmes parallèles plus performants [CI-31].

• Algorithmique, combinatoire et complexité

Propriétés de structures et de familles d'objets combinatoires

Graphes et hypergraphes aléatoires

Vlady Ravelomanana travaille sur les *tailles des composantes géantes dans les graphes aléatoires* $\mathbb{G}(n,p)$ [RI-47, CI-63, RI-64]. Ses résultats montrent en particulier que ces composantes possèdent $O(\ell^{1/3}\, n^{2/3})$ sommets lorsque le nombre d'arêtes est exactement $n+\ell$, avec $\ell = o(n)$. Ces analyses fines montrent donc que durant la phase critique de transition la taille des plus grandes composantes d'un tel objet varie violemment de $O(\log n)$ à $O(n^{2/3})$ puis de manière plus douce de $O(n^{2/3})$ à $O(n)$ (en passant par toutes les phases intermédiaires). Ces résultats se généralisent aux hypergraphes aléatoires $\mathbb{H}(n,p)$ [RI-47, CI-63, RI-64]. Ce sont là les premiers

résultats concernant les tailles des composantes géantes durant ce phénomène, connu sous le nom de *Double-Jump* d'Erdős-Rényi.

Énumération de marches aléatoires sur $\mathbb{N}$

Ce sujet qui est l'objet d'études combinatoires depuis plus d'un siècle [RI-43] reste cependant au cœur de nombreuses disciplines [LI-3]. Dans le cas d'un jeu de sauts autre que $+1$ et -1, l'étude de ces marches sur $\mathbb{N}$ devient fort délicate, néanmoins, une approche initiée par C. Banderier et Philippe Flajolet montre que de nombreux paramètres se prêtent à une étude énumerative et asymptotique exacte. Plus récemment, C. Banderier et B. Gittenberger [CI-129] ont ainsi montré que l'aire sous une marche aléatoire a des moments algébriques, ouvrant la porte à une analyse fine de ce paramètre délicat relié à la fonction d'Airy.

Coloration d'arêtes d'un multigraphe

Le résultat obtenu par Jean Cardinal, V. Ravelomanana et M. Valencia-Pabon [CI-159] sur un type de coloration d'arêtes d'un multigraphe connu comme *chromatic edge strength* est intéressant dans le contexte des problèmes d'ordonnancement des tâches cycliques et parallèles. Il montre qu'il est toujours possible de trouver un ordonnancement pour ce type de tâches qui permet de minimiser à la fois le temps total d'exécution et le temps moyen avant qu'une tâche se termine. Ce problème est connu comme étant NP-difficile même si le (multi)graphe associé des tâches est bipartite, planaire et de degré maximum 3.

Noyaux de graphes dirigés

C. Banderier, J.-M. Le Bars et V. Ravelomanana [CI-57] ont étudié une classe de graphes orientés (digraphes) avec une contrainte de coloriage due à Von Neumann et issue de la théorie des jeux (de type jeu de Nim). Cette contrainte est équivalente à la notion de *noyaux d'un digraphe*, qui apparaît dans de nombreux domaines, non seulement en théorie des jeux mais aussi en théorie de la complexité, en intelligence artificielle (logique des défauts, argumentation en système multi-agents), loi 0-1 en logique monadique du second ordre, et en combinatoire (graphes parfaits), etc. La notion de noyau d'un digraphe débouche sur des questions difficiles (au sens de la NP-complétude, de la #P-complétude, ou de l'approximation). Cependant, les auteurs ont montré qu'il est possible d'utiliser les séries génératrices et leurs singularités afin d'obtenir, par exemple, des résultats sur la taille du noyau dans un circuit ou dans un digraphe unicircuit, et la probabilité de gagner à certains jeux de Nim. C'est un premier pas vers un traitement des noyaux dans des graphes plus complexes, avec, par exemple, une approche à la Wright. Cette méthode peut être appliquée plus généralement à n'importe quelle contrainte « locale» de coloriage dans des structures combinatoires décomposables.

Mesure de complexité : l'analyse perturbative

Les résultats de C. Banderier, R. Beier et K. Mehlhorn [CI-23] ont déjà exhibé des paysages de complexité assez variés ou surprenants; notamment certains dont le pire des cas classique n'est paradoxalement pas la région la plus difficile, etc. (Un succès de cette approche est, pour l'instant, une explication de la complexité polynomiale en pratique de la méthode du simplexe.)

Notons que cette étude fait intervenir de l'asymptotique de coefficients de fonctions de variables complexes (vérifiant des équations différentielles d'un certain type) qui apparaissent de façon ubiquitaire en analyse d'algorithmes; divers problèmes qui y sont liés demeurent toujours ouverts (cf. § 3.4.3).

OCAD

3.4.3 Perspectives de recherche

Malgré la diversité des thèmes de recherche de la composante AD et de leurs perspectives, un souci commun est de progresser toujours plus dans la compréhension profonde des modèles, de la structure et des propriétés essentielles des objets analysés. Pour cela, des progrès notables doivent être réalisés dans la recherche et la compréhension d'outils adéquats plus sophistiqués, donc de plus en plus mathématisés.

• Conception et analyse d'algorithmes distribués

Nous trouvons un exemple presque paradigmatique de tels besoins dans les analyses de complexité d'algorithmes distribués. Après s'être intéressés à la complexité dans le pire des cas, les

chercheurs ont borné la complexité *en moyenne* de ces algorithmes. Nous nous proposons de raffiner ces analyses en moyenne (asymptotique exacte de l'espérance, variance, etc.), voire d'inclure des preuves d'existence et de mise en évidence d'une *distribution limite* (lorsqu'elle existe). Seules des techniques de bon niveau mathématique peuvent autoriser ce type d'analyses : séries génératrices, approximations asymptotiques (réelles ou complexes), méthodes probabilistes, etc.

Classes générales d'algorithmes distribués

Le recrutement récent de Mario Valencia-Pabon procure déjà un élan et des compétences accrues (théorie des graphes, coloration, etc.)aux recherches en algorithmique distribuée. Un premier résultat, soumis en commun avec Christian Lavault, est la conception et l'analyse d'un algorithme performant de construction d'un *arbre de recouvrement minimum à degré mimimum* d'un graphe quelconque [RI-84]. Sur cette base, nous avons aussi commencé l'analyse en moyenne d'algorithmes de *routage par retournement d'arcs* (total ou partiel) dans les graphes sans cycle orientés (DAG) quelconques. (Des résultats sont déjà obtenus pour la ligne et les forêts arborescentes.)

Dans le vaste domaine de l'ordonnancement, Franck Butelle, Lucian Finta et Mourad Hakem prolongent leur étude de grandes classes de protocoles de contrôle distribués. Il s'agit d'abord d'algorithmes distribués d'ordonnancement sous *divers types de contraintes* (matérielles et temporelles), puis ensuite, petit à petit, de leur *relaxation partielle*. L'introduction d'algorithmes *tolérants aux fautes* (pannes franches et transitoires, fautes quelconque, fautes byzantines, etc.) d'une part et, d'autre part, de la prise en compte de la *finesse dans les délais de communication* dans les réseaux constituent autant de perspectives nouvelles vers la construction d'une palette très générale de ces classes de protocoles. Par ailleurs, ces travaux devraient s'intégrer en tant qu'outils pour le *Grid Computing* (GRID5000).

Protocoles de communication dans les réseaux sans fil

OCAD

Les problèmes en chantier dans cet axe de recherche sont fondamentaux pour la communication dans un MANET. Certains résultats pourraient être utilisés comme « briques de base » pour des protocoles de contrôle de haut niveau déjà cités. L'objectif principal de Christian Lavault, Vlady Ravelomanana et Mario Valencia-Pabon est de réussir à émuler un réseau MANET à simple saut à l'aide d'un réseau MANET à sauts multiples, sans aucune information structurelle *à priori* sur le graphe G engendré. Les hypothèses (très générales) sont donc les suivantes :

1. initialement, les stations sont *indiscernables* (symétriques et anonymes) ;
2. les stations ne possèdent *a priori* aucune connaissance structurelle (exacte ou partielle) de G, ni son nombre de noeuds n, ni son degré maximum Δ, ni son diamètre D (ou des majorants).

Il s'agit donc de surmonter l'absence d'information en évaluant l'ensemble des caractéristiques de G et de construire pas à pas des algorithmes de diffusion probabilistes efficaces afin d'approcher Δ, D et enfin n avec une erreur $\epsilon > 0$ faible.

Les méthodes stochastiques et l'asymptotique (réelle et complexe) appliquées aux séries génératrices multivariées permettent d'obtenir des résultats suffisamment précis pour espérer résoudre ces problèmes de manière satisfaisante.

• Systèmes distribués et algorithmique parallèle

Grid Computing

Deux ouvertures internes à Paris 13 sont amenées à se développer.

– La première consiste à partir de la souche du logiciel de calcul sur grille de PC XTREMWEB et de l'adapter dans un contexte bio-médical, une première collaboration entre le LIPN et l'UFR SMBH est mise en place en 2006 via un stage de deuxième année d'élève ingénieur de Paris 13.

Du coté SMBH, le Dr. Jean-Marie Rocchisani (MCU-PH Médecine Nucléaire, Paris 13) s'est intéressé depuis plusieurs années aux grilles médicales, notamment HEALTHGRID et plus particulièrement à leurs applications pour la plate-forme de simulation Monte Carlo GATE (`http://opengatecollaboration.healthgrid.org`). GATE est un outil performant de simulation des systèmes d'imagerie utilisant des rayonnements particulaire ou photonique, existants ou

en cours de développement. A. Martineau et J.-M. Rocchisani ont déjà validé la modélisation avec GATE d'une « gamma-caméra » dédiée à l'imagerie du petit animal et des petits organes.

Les objectifs sont donc de paralléliser pour la grille l'application GATE pour la plate-forme XtremWeb (besoin du SMBH) et d'offrir également un niveau de certification des images produites par le système de calcul XtremWeb (besoin du LIPN). Pour l'instant, dans le cadre de *SafeScale*, il n'y a que de la certification symbolique qui est envisagée c'est à dire que la certification consiste soit à évaluer un polynôme soit à vérifier que le résultat est trié puis de garder ou rejeter le calcul. Dans le cas de GATE, la certification pourra utiliser plusieurs critères quantitatifs d'évaluation de la qualité des résultats des simulations de deux systèmes d'imagerie de référence (un « Tomographe par Émission de Positons » et une « gamma-caméra »).

– La seconde est relative à la rénovation de la grappe commune LAGA-LIPN. Plusieurs financements (de l'ANR *SafeScale*, CNRS, interne à Paris 13) ont été injectés pour acheter des nœuds, du matériel d'interconnexion, de la climatisation. Cette infrastructure doit permettre aux deux laboratoires d'avoir une plate-forme de base afin de mettre au point des tests dimensionnants par exemple. Nous avons retenu une organisation matérielle à la Grid5000 c'est-à-dire hautement reconfigurable, seule garantie à notre avis pour que la plate-forme puisse être acceptée par le plus grand nombre de chercheurs des deux laboratoires. Les années qui arrivent vont permettre au groupe des grilles de commencer le travail de prototypage en interne avant déploiement dans Grid5000.

Au niveau national, concernant le projet *SafeScale*, le LIPN s'est en particulier positionné sur le plan des applications. Dans les 18 derniers mois du projet, nous avons en chantier :

1. Le résultat obtenu dans [CI-123] préfigure une partie du programme de recherche de Christophe Cérin et Jean-Christophe Dubacq dans l'ANR *SafeScale*. Une collaboration étroite avec l'Imag (LIG) autour des travaux de Daouda Traoré et de Serge Guelton pour les implémentations liées à l'intergiciel étudié Kaapi (*Kernel for Adaptative, Asynchronous Parallel and Interactive programming*). Nous travaillons sur le partitionnement de données et l'auto-adaptativité — comment faire recouvrir le temps de partitionnement avec la diffusion des données dans un contexte parallèle adaptatif (aux congestions réseaux par exemple).
   Nous avons également identifié le problème de la génération des éléments fréquents comme application pouvant bénéficier d'une technique adaptative de gestion des processus. L'idée est de commencer dynamiquement la génération des groupes de $k+1$ éléments fréquents avant d'avoir terminé la génération complète des groupes de k éléments (méthode de l'algorithme classique). L'intergiciel Kaapi doit nous permettre de le réaliser de manière souple.
2. En interne, Frédéric Toumazet doit proposer un code stable de son application de vérification de propriétés de polynômes avant la fin 2007 afin que nous puissions l'intégrer avec un intergiciel de type XtremWeb.

Au niveau international, deux co-tutelles de thèse viennent d'être mise en place avec l'École Supérieure des Sciences et Technologies de Tunis (ESSTT). Le travail est orienté respectivement vers l'algorithmique parallèle de gestion des communications entre grappes de grappes et l'ordonnancement par vol de cycle d'une part et la gestion de données dans des systèmes pair-à-pair de publication d'annonces.

L'opération de diffusion pour des grappes de grappes est étudiée dans le cadre d'heuristiques à partir d'arbres binomiaux convenablement définis. Par ailleurs nos travaux actuels et futurs se démarquent par le fait que nous visons à réaliser la diffusion lorsque les nœuds possèdent plusieurs cartes réseaux.

Un premier travail expérimental [Ra-31] autour de plusieurs systèmes pairs-à-pairs nous a permis de montrer en 2007 les limites de ces systèmes lorsqu'on les déploie sur un nombre relativement important de noeuds (300 processeurs). Nous envisageons plus particulièrement un travail sur le stockage des informations de gestion ou des données applicatives de programmes tournant sur des plate-formes de type grille de PC. En effet, les plate-formes actuelles sont fortement centralisées dans le sens ou un coordinateur central dispose de toutes les connaissances pour ordonnancer les travaux et, de plus n'utilisent pas les techniques pairs-à-pairs pour stocker les données générées en cours d'exécution. Ces deux aspects limitent les capacités de passage à l'échelle des intergiciels de grille de PC.

PGCD de deux entiers

Mohamed Sedjelmaci approfondit l'étude des algorithmes parallèles de calcul du PGCD en proposant deux variantes de l'algorithme d'Euclide, l'une *binaire*, l'autre sous forme *accélérée*.

• Algorithmique, combinatoire et complexité

Graphes et hypergraphes aléatoires

– Les *graphes aléatoires géométriques* sont minutieusement étudiés par Penrose. Il subsiste cependant plusieurs problèmes ouverts, dont le *rayon critique* des disques à partir duquel un graphe possède (asymptotiquement presque sûrement) un cycle hamiltonien. Vlady Ravelomanana a étudié jusqu'à présent les graphes aléatoires géométriques *homogènes*, i.e. dont tous les points possèdent le même rayon. Dans un futur proche, nous nous proposons d'étudier ces structures lorsque les centres des disques, dispersés aléatoirement, disposent d'un rayon selon une certaine loi de probabilité. Ces problèmes sont intimement liés à ceux soulevés par la communication dans les réseaux sans fils (capteurs en particulier). En effet, toute station très utilisée voit sa batterie s'affaiblir rapidement et, par conséquent, son rayon de transmission et de réception diminuer en proportion.

– Dans le domaine des (hyper)graphes aléatoires ($\mathbb{G}(n,p)$ ou $\mathbb{H}(n,p)$), nous nous intéressons aux problèmes liés aux transitions de phases. Typiquement, en augmentant la probabilité d'existence p d'une arête, l'objet passe d'un état « gazeux » à un état « liquide », puis d'un état « liquide » à un état « solide ». Nous avons pour objectif de caractériser de tels phénomènes dans un cadre général et unificateur, grâce aux techniques offertes par la combinatoire analytique et les méthodes probabilistes.

– Il s'agit aussi d'étendre des résultats sur les graphes [RI-47, **?**] à certaines variétés d'hypergraphes et ce, de la manière la plus générale possible. Notre second but est d'utiliser ces résultats afin de quantifier les transitions de phases dans les instances aléatoires des problèmes de satisfiabilité tels que 2-XORSAT, $(2+k)$-SAT ou SAT, qui sont des problèmes centraux en théorie de la complexité. Ensuite, nous nous proposons d'exploiter les résultats et les méthodes obtenus pour concevoir et analyser des algorithmes classiques d'optimisation combinatoire (couplages, ensembles indépendants, cliques, etc.). Ces travaux pourront enfin servir de briques de base pour les études d'approximation en moyenne d'algorithmes (cf. paragraphe *Approximation en moyenne et analyse perturbative* (Aperox) ci-dessous).

OCAD

Combinatoire des mots circulaires et des permutations

Ce projet vise à ébaucher un début d'*unification* de la recherche en combinatoire (analytique et algébrique) au LIPN ; ces premiers pas s'effectuent par une coopération (déjà en œuvre) entre les combinatoriciens de l'équipe OCAD et ceux du thème *Combinatoire* de l'équipe LCR. Le rapport de Master (2006) et le sujet de thèse de Hichem Kenniche (co-encadré par Gérard H. E. Duchamp et Christian Lavault) s'inscrivent dans cet objectif. En effet, l'étude et l'analyse d'algorithmes et de structures combinatoires dans des cycles de permutations complètent avec grand profit leur combinatoire algébrique.

La statistique des objets combinatoires relatifs aux permutations « classiques » (pics, creux, doubles montées-descentes) s'obtient aisément à partir de leurs séries génératrices. Il existe cependant peu de travaux généralisant ces résultats aux permutations dites « en anneau » (ou *unicycliques*), i.e. dans une même classe d'équivalence de cycle définie aux permutations circulaires près. Il en est de même pour l'application à l'analyse en moyenne d'algorithmes distribués d'élection avec identités sur un anneau (en cycle) qui procèdent par élimination des pics (creux) de ces identités. La série génératrice associée et l'asymptotique des coûts permettent de calculer les premiers moments de la complexité en temps et en messages, puis de prouver l'existence d'une distribution limite (si elle existe), etc.

Le but est, en effet, d'associer les approches analytique et algébrique et d'étendre leurs résultats de manière à lever les verrous scientifiques dus aux différentes représentations des objets combinatoires « circulaires » en général : sur un alphabet Σ, propriétés des colliers construits sur Σ^*, évaluation de la longueur moyenne des occurrences d'un symbole de Σ dans un collier, etc.

Approximation en moyenne et analyse perturbative (Aperox)

Le projet APEROX repose sur l'emploi d'outils mathématiques pour étudier la « qualité » d'algorithmes approchés pour des problèmes « difficiles » d'optimisation discrète (NP-h). APEROX associe ainsi deux domaines de recherche qui constituent, depuis plus de dix ans, deux axes forts d'OCAD : la mesure de complexité d'un algorithme et l'évaluation des solutions retournées par un algorithme, pour un problème donné ; il exploite des compétences et des complémentarités de l'équipe en analyse en moyenne d'algorithmes (Cyril Banderier et Vlady Ravelomanana) et en approximation polynomiale (Sophie Toulouse). Précisément, le projet se propose d'initier deux approches nouvelles :

- *l'approximation en moyenne*, qui mesure le degré d'approximation d'un problème sur des instances aléatoires ;
- *l'analyse perturbative*, déjà détaillée au § 3.4.1.

Concernant l'approximation en moyenne, les motivations sont les suivantes : outre le fait de disposer d'un indicateur complémentaire d'évaluation des algorithmes, l'analyse en moyenne du rapport d'approximation permet surtout de fournir des informations sur les problèmes qui sont non- ou mal approximables (tels, notamment, les problèmes MaxSAT, MaxCCSP, MaxIS, etc.) ; plus spécifiquement, de même que s'affranchir de l'exactitude pour la résolution des problèmes NP-difficiles (NP-h) permet de structurer NP-h en fonction du degré d'approximation des problèmes (classes (D)PTAS, (D)APX, (D)PolyAPX, etc.), s'affranchir du pire cas pour les problèmes zéro-approximables en différentiel (0-DAPX) devrait offrir une certaine structuration des problèmes 0-DAPX (ou plus largement, des problèmes non approximables à rapport constant) en fonction de leur degré d'approximation en moyenne (approximable à rapport moyen constant, polynomial, etc.).

Toujours dans le sens d'une évaluation plus fine des algorithmes, on pourrait envisager de faire l'analyse perturbative du rapport d'approximation ; si une telle démarche offrirait en effet un indicateur encore plus significatif de la qualité d'un algorithme, le défi semble cependant plus qu'ambitieux. Aussi, à défaut d'analyser le comportement du rapport lui-même, il serait intéressant d'étudier le comportement de la valeur d'une solution optimale, d'une pire solution, d'une solution approchée définie structurellement (ex. : un optimum local). Par ailleurs, faire l'analyse en moyenne d'une solution retournée au voisinage (perturbé) d'une pire instance (instance réalisant le pire cas d'un algorithme donné) donnerait plus de sens à l'analyse pire cas de l'algorithme considéré, en proposant une réponse à la question : le pire cas est-il significatif ?

OCAD

3.4.4 Diffusion de la recherche

• Évaluation de la recherche

Comités éditoriaux

- C. Cérin Éditeur en chef d'un numéro spécial de *Parallel Computing* en 2006-2007.
- C. Lavault Membre du comité de rédaction de *Studia Informatica Universalis* depuis 2001.
- C. Lavault Membre du comité scientifique de *Parallélisme, réseaux et répartition* depuis 1995.

Comités de programmes

- Christophe Cérin, Jean-Christophe Dubacq et Lucian Finta (2007). Membre du comité de programme de la Conférence *Grid and Pervasive Computing* (GPC-2007), France, 2-4 mai 2007.
- Jean-Christophe Dubacq (2007). Membre du comité de programme du numéro spécial de la revue *Parallel Computing* (PARCO), Elsevier.
- Christian Lavault (2007). *Co-chair* de la *3rd International Conference on Mobile ad-hoc and Sensor Networks* (MSN-2007), Beijing (Chine), 12-14 déc. 2007.
- Christian Lavault (2007). Membre du comité de programme de la Conférence *The 2007 International Symposium on Parallel and Distributed Processing with Applications* (ISPA-2007), Canada, 21-24 août 2007.
- Mario Valencia-Pabon (2007). Membre du comité de programme de la Conférence *Latin-American Algorithms, Graphs and Optimization Symposium* (LAGOS-2007), Chili, 25-29 nov. 2007.

- Christophe Cérin (2007). Program Chair pour un numéro spécial de la revue *Parallel Computing* (PARCO).
- Christophe Cérin (2006). Program Co-Chair de *Grid and Pervasive Computing* (GPC'2006). Membre des comités de programme de SBAC, AINA, GPC et de RenPar 2006. Finance Chair de HPDC'2006.
- Christian Lavault (2003, 2004, 2005). Membre des comités de programme d'OPODIS (*International Conference on Principles of Distributed Systems*) en 2003, 2004 et 2005.
- Cyril Banderier (2003). Responsable du comité de programme de la conférence *Marches aléatoires discrètes : théorie et applications en combinatoire, biologie algorithmique, informatique théorique, probabilités, physique statistique* (DRW-2003), IHP, 2003.

Comités de lecture ou d'évaluation

- Cyril Banderier. Rapporteur d'une vingtaine d'articles pour des revues et conférences internationales (Algorithmica, Disc. Math., JCT-a, JTNB, RSA, TCS, EJC, SLC, DRW, FPSAC, Gascom, Mathinfo, MFCS, IEEE, Proba. Stat., etc.).
- Franck Butelle. Rapporteur d'une demi-douzaine d'articles pour des revues et conférences internationales (Ann. of Telecom, Int. J. on Fondations of Comp. Science, IPDPS, etc.).
- Christophe Cérin. Rapporteur d'une demi-douzaine d'articles pour des revues et conférences internationales (IEEE Trans. on Comp., AINA, IPDPS, GPCSBAC, etc.). Finance Chair de HPDC'2006.
- Christian Lavault. Rapporteur d'une vingtaine d'articles pour des revues et conférences internationales. (Acta Inf., Dist. Comput., Inf. & Comput., IPL, Par. & Dist. Comput., Par. Proc. Letters, Studia Inf. Univ. , TCS, DISC, ICALP, IEEE Int. Conf. on CS and Appli., ICDCS, MFCS, OPODIS, SIROCCO, IEEE TPDS, etc.).
- Vlady Ravelomanana. Rapporteur d'une quinzaine d'articles pour des revues et conférences internationales. (Disc. Maths and TCS, IPL, J. of Algo., J. de Th. des Nombres de Bordeaux, Par. Proc. Letters, RSA, TCS, IEEE TPDS, IEEE/ACM Trans. on Net., DRW, FPSCAC, HICSS, IEEE Int. Conf. on CS and Appli., etc.).
- Mario Valencia-Pabon (2006-2007). Rapporteur d'une douzaine d'articles pour des revues et conférences internationales (App. Disc. Maths., Disc. Maths., J. of Comput. and App. Maths, STACS, LAGOS, etc.).

OCAD

Autres responsabilités d'évaluation

- Christophe Cérin (2003-2006). Membre du jury du prix de thèse de Spécif.
- Christian Lavault (2007). Rapport d'évaluation d'un projet du COFECUB (*COmité Français d'Évaluation de la Coopération Universitaire et scientifique avec le Brésil*) : *Modèles hybrides et adaptatifs pour la construction des systèmes répartis adaptés au facteur d'échelle (scalable).*
- Christian Lavault (2006). Rapport d'évaluation pour la nomination de la professeure P. Flocchini comme *Full Professor* de l'Université d'Ottawa (CRU – Ontario, Canada).
- Christian Lavault (2005). Rapport d'évaluation pour nomination de la professeure P. Flocchini à une Chaire de recherche de l'Université d'Ottawa (CRU – Ontario, Canada).

Jurys hors LIPN

Jurys de thèse : Cyril Banderier (rapporteur : 1). Christophe Cérin 3 (rapporteur : 2). Christian Lavault 16 (rapporteur : 7).

Jurys d'habilitation à diriger des recherches : Christian Lavault (2).

• Organisation de la recherche

Responsabilité de groupes de travail du CNRS

- Christian Lavault. Responsable du groupe *Combinatoire analytique* de MATHSTIC *Alea* dans le GDR 673 du CNRS *Informatique Mathématique* (IM).
  Responsable de l'équipe *Combinatoire et analyse d'algorithmes distribués* (CADAD) du groupe de travail *ALEA* dans le projet CNRS *MathSTIC* de décembre 2001 à janvier 2006 (laboratoires partenaires : LaRIA, université d'Amiens, et LIPN).

– Christophe Cérin. Responsable depuis 2007 du groupe *Sécurité, Sureté* dans le GDR ASR (Architecture, système, Réseaux).

Organisation de conférences et événements scientifiques

– Cyril Banderier (2005-2006) : coorganisation des Congrès *MATh.en.JEANS*.
– Christophe Cérin, J.-C. Dubacq et L. Finta (2007) : organisateurs locaux de la conférence *Grid and Pervasive Computing* (GPC-2007), France, 2-4 mai 2007.
– Christophe Cérin (2006) : organisateur en 2006 d'un numéro spécial de la revue *Parallel Computing* sur les *Large Scale Grids*.
– Christophe Cérin (2005) : co-organisateur en 2005 d'un workshop en marge de la conférence EUROPAR (Lisbonne) sur les système de Grilles (workshop sur invitation).
– Christian Lavault, V. Ravelomanana (2005) : membres du comité d'organisation (avec J.-M. Le Bars) des *Journées ALEA* 2005 (CIRM, SMF-CNRS, Luminy).
– Cyril Banderier, Christian Lavault, Vlady Ravelomanana (2003) : comité d'organisation de la conférence *Marches aléatoires discrètes : théorie et applications en combinatoire, biologie algorithmique, informatique théorique, probabilités, physique statistique* (DRW-2003), IHP, 2003.

Contrats

Type	Coordinateur	Resp. scient.	Nom	Financeur	Période	Durée (mois)	Budget global (keuros)	Part LIPN (keuros)
PAI	•	C. Banderier	Amadeus	Égide	01/2005-12/2006	24	18	9
ANR	•	C. Cérin	*SafeScale*	CNRS	01/2006-12/2008	36	337	70
ACI		C. Cérin	*Grid Explorer*	CNRS	06/2003-07/2006	36	2000	
AUF	•	C. Lavault	*Coop. AUF*	AUF	07/2006-12/2008	24	22	
ECOS	•	C. Lavault	*Action ECOS*	ECOS-Sud	01/2008-01/2011	36	63	

ANR *SafeScale* *Autres partenaires : universités de Grenoble, Rennes et Brest*

Le projet d'ANR *SafeScale – Security And Fault-tolerance to Exploit Safety ambient Computing in lArge scaLe Environments* est piloté par le LIPN et regroupe des laboratoires à Grenoble, Rennes et Brest. Quatre chercheurs du LIPN y sont représentés : Christophe Cérin, Jean-Christophe Dubacq, Lucian Finta et Frédéric Toumazet. Ce projet a été labélisé fin 2005.

Par ailleurs et en liaison avec les grilles, le LIPN s'est également positionné dans l'initiative fr-GRID qui vise à installer un portail d'accès à la grille opérationnelle. Ce projet est porté par Ph. d'Anfray (Renater) et C. Cérin. L'embryon de portail est sur `http://www.fr-grid.org`. Parmi les objectifs affichés de cette initiative figure la formation (une semaine de formation sera organisée début 2007).

ACI *Grid Explorer*, INRIA-STIC *Autres partenaires : Inria, universités de Grenoble, Tunis et Versailles*

Christophe Cérin participe à l'ACI *Grid Explorer* et à un projet Inria-STIC-université Tunisienne en cours de montage pour un renouvellement. Il sera porté par le LIPN avec comme partenaires Grenoble, Versailles, Inria Futur, ESSTT et la Faculté des Sciences de Tunis.

Coopération AUF *Autres partenaires : universités d'Antananarivo et de Bruxelles (ULB)*

Ce projet de coopération de l'AUF de deux années (2006-2008) regroupe trois universités francophones, l'*Université d'Antananarivo* (Tananarive, Madagascar), l'*Université Libre de Bruxelles* (ULB, Bruxelles, Belgique) et l'université Paris 13. Ses objectifs visent à la formation et

OCAD

au renforcement d'un troisième cycle en informatique théorique et mathématiques discrètes à l'université d'Antananarivo (formation doctorale IGMA, *Informatique et Génie Mathématique*). Il s'agit en particulier de cours dans le cadre du DEA IGMA, d'encadrements de doctorants et de stagiaires en Master recherche, ainsi que d'échanges interuniversitaires de chercheurs.
Co-responsables : C. Lavault (Paris 13) et A. Randrianarivony (Un. d'Antananarivo). Responsables locaux : J. Cardinal (ULB), C. Lavault et V. Ravelomanana (Paris 13), A. Randrianarivony (Un. d'Antananarivo).

Action ECOS-SECyT *Autres partenaires : universités de Buenos Aires, Nationale de la Plata (Argentine), de Chile, Columbia University, de Magdeburg et de Rio de Janeiro*

Le projet franco-argentin ECOS-SECyT *Coloration des graphes : généralisations et applications aux jeux combinatoires* a pour partenaires institutionnels ECOS-Sud (Ministère des Affaires Étrangères et Ministère de l'Éducation Nationale) et SECyT (Secretaría de Ciencia, Tecnología e Innovación Productiva). Cette coopération de recherche universitaire implique deux laboratoires, le LIPN (Un. Paris 13, France) et le Departamento de Computación (Facultad de Ciencias Exactas y Naturales, Un. de Buenos Aires, Argentine).
Cinq chercheurs du LIPN y participent : Christian Lavault (PU), Mario Valencia-Pabon (MCF), Sophie Toulouse (MCF), Vlady Ravelomanana (MCF) et Hichem Kenniche (doctorant).
Co-responsables : Christian Lavault (LIPN, Un. Paris 13) et Pablo Jacovkis (PU, Departamento de Computación, Un. de Buenos Aires).

Par ailleurs, Mario Valencia-Pabon participe à l'action de collaboration internationale *Théorie des graphes et optimisation combinatoire* pilotée par l'Université de Buenos Aires (Argentine) dans le cadre de l'ANPCyT. Les six partenaires de ce projet de recherche scientifique et technologique d'une durée de 38 mois sont l'Un. de Buenos Aires (Argentine), Columbia University (USA), l'Un. de Magdeburg (Allemagne), l'Un. de Paris 13 (France), l'Un. Nationale de la Plata (Argentine), l'Un. Fédérale de Rio de Janeiro (Brésil) et l'Un. de Chile (Chili). (Budget global : 225 000 euros.)

OCAD

• Dissémination

Actions de vulgarisation de la recherche

- Cyril Banderier (2003 à 2006) :
  Congrés *MATh.en.JEANS* (Paris, 2004-2006).
  Journées *Savante Banlieue* (Villetaneuse, 2003-2006) : stand et/ou miniconférence.
  Rencontres CNRS Jeunes *Sciences et Citoyens* (Futuroscope de Poitiers, 2005-2006).
  Intervention en collège à Epinay sur Seine (2004-2006).
  Rencontres CNRS Jeunes à Drancy en 2006.
  Salon 2006 de la culture et des jeux mathématiques (Paris, 2005-2006) :
  animation du stand CNRS/*MATh.en.JEANS*.
  Projet *Animath* sur Pi, journée au Palais de la découverte (juin 2006).
  Fête de la science, Math pour tous/Cité des Sciences (Paris, 13-15 octobre 2006).
- Christophe Cérin, Jean-Christophe Dubacq (2006) : Coordination de l'intervention du LIPN à *Savante banlieue*.

Invitations, tutoriels

- Cyril Banderier (2003) : Journées du LIPN, IG (26-27 mai 2003).
- Cyril Banderier (2003) : Rencontres mathématiques de Rouen « in honor of Claude Dellacherie», Rouen (18-20 juin 2003).
- Cyril Banderier (2004) : Séminaire d'algorithmique du GREYC, Caen (26 Octobre 2004)
- Cyril Banderier (2005) : *Joint Seminar*, TU Wien/U Wien. (15 novembre 2005)
- Cyril Banderier (2006) : Séminaire de théorie des nombres et de combinatoire, Lyon (3 octobre 2006)
- Christophe Cérin (2005-2006) : invitation annuelle à Tunis pour les *Journées du parallélisme* et dans le cadre d'un projet Inria STIC avec les universités tunisiennes.
- Christian Lavault (2006) : *Advanced Research Workshop on Information Security in Wireless Networks*, NATO-ARW, septembre 2006, Suceava (Roumanie). Tutoriel invité : *Link*

Reversal Routing Algorithms in MANETs.
- Vlady Ravelomanana (2006) : *Advanced Research Workshop on Information Security in Wireless Networks*, NATO-ARW, septembre 2006, Suceava (Roumanie). Tutoriel invité : *Random Geometric Graphs and Wireless Networks.*
- Christian Lavault et Vlady Ravelomanana (2005) : *Journées ALEA* 2005 (CIRM, SMF-CNRS, Luminy).

• Collaborations universitaires et invitations

- 1999-2004. Christian Lavault. Dans le cadre d'une coopération ERASMUS-SOCRATES , échanges avec l'*université Polytehnica* de Bucarest (Roumanie).
  Responsables : V. Cristea et A. Florea.)
- 2005- . Christophe Cérin. Projet Inria STIC-université Tunisienne qui vise la formation doctorale en Tunisie et s'appuie sur un programme de recherche entre Grenoble, Versailles, Villetaneuse et la faculté des sciences et l'ESSTT de Tunis.
- 2006-2008. Cyril Banderier, Christian Lavault, Vlady Ravelomanana, Sophie Toulouse. Coopération AUF entre trois universités francophones, l'*université d'Antananarivo* (Tananarive, Madagascar), l'*université Libre de Bruxelles* (ULB, Bruxelles, Belgique) et l'université Paris 13. Le projet consiste à développer la formation et le renforcement d'un troisième cycle en informatique théorique et mathématiques discrètes à l'université d'Antananarivo (formation doctorale IGMA, *Informatique et Génie Mathématique*).
  Responsables : C. Lavault (Paris 13) et A. Randrianarivony (Un. d'Antananarivo). Responsables locaux : J. Cardinal (ULB), C. Lavault et V. Ravelomanana (Paris 13), A. Randrianarivony (Un. d'Antananarivo).

3.4.5 Formation à la recherche

Thèses soutenues

2002-2006. Mourad Hakem : *Ordonnancements distribués*
Financement de thèse :
Soutenue le 16 décembre 2006
Directeur : Christian Lavault. Encadrant : Franck Butelle
Situation actuelle du diplômé : postdoctorant.

2001-2005. Gabriel Paillard : *Quelques contributions à l'algorithmes distribués*
Financement de thèse : allocation de recherche brésilienne
Soutenue le 28 juin 2005
Directeur : Christian Lavault
Situation actuelle du diplômé : assistant en Informatique à l'Universitad Federal du Ceara (Brésil) depuis septembre 2006.

Doctorants

nom	date	position actuelle pour thèses
T. Andriamampinina	10/03-12/07	AM IG
E. Levy	10/03-12/07	Assistant ULB (Belgique)
A. Rasoanaivo	10/06-12/09	bourse AUF
H. Kenniche	10/06-12/09	bourse IG

OCAD

Cours de DEA / Master 2 recherche

- 2003 à 2004. (F. Butelle) : *algorithmique du texte*, DEA ISCCA, Paris 8-EPHE.
- 2003 à 2004. (C. Lavault) : *Introduction à l'analyse des algorithmes*, DEA ISCCA, Paris 8-EPHE. *Introduction à l'analyse des algorithmes distribués*, DESS IED (enseignement à distance), Paris 8.
- 2003 à 2005. (C. Lavault) : *Algorithmique distribuée*, DEA IA-OC, IG.
- 2005 à 2007. (C. Banderier, C. Lavault, V. Ravelomanana) : *Combinatoire analytique*, 2^{e}année de Master Recherche Mathématiques-Informatique, IG. (Commun avec le Master Mathématiques-Informatique-Technologies 2^{e}année recherche (*Mathématiques et Applications au codage et à la cryptologie*), Paris 8.)

– 2006-2008 (C. Banderier, C. Lavault, V. Ravelomanana) : *Algorithmique et Combinatoire analytique*, DEA IGMA (*Informatique et Génie Mathématique*), Université d'Antananarivo (Madagascar).

Stages de DEA / Master 2 recherche

– 2003. (C. Lavault) : E. Levy (DEA IA-OC Paris 13, MPRI2) : *Analyse et conception d'un algorithme distribué de cycle hamiltonien pour les graphes aléatoires de type* $\mathbb{G}(n,p)$.
– 2005. (C. Lavault, V. Ravelomanana) : Thanh-Binh Doan (DEA IA-OC Paris 13) : *Algorithmique des réseaux radio.*
– 2006. (G. Duchamp, C. Lavault) : H. Kenniche (Master MMI2 Paris 13, LIPN) : *Combinatoire des mots circulaires et permutations.*
– 2006. (V. Ravelomanana, S. Toulouse) : H. Razafisambatra (DEA IGMA Tananarive) : *Approximation de problèmes définis sur des collections d'ensembles.*
– 2006. (V. Ravelomanana, S. Toulouse) : S. Djabali (P6-UPMC, LIPN) : *Approximation en moyenne.*

3.4.6 Publications principales

Livres

[ocad :BK03] Cyril Banderier et Christian Krattenthaler. *Discrete Random Walks : Theory and applications in combinatorics, computational biology, computer science, probability theory, and statistical physics.* Discrete Mathematics & Theoretical Computer Science, 2003.

Articles dans des revues internationales avec comité de lecture sélectif

OCAD

[ocad :LaVa08] Christian Lavault et Mario Valencia-Pabon. A distributed approximation algorithm for the minimum degree minimum weight spanning tree. *Journal of Parallel and Distributed Computing*, 68(2) :200–208, 2008.

[ocad :LMR07] Christian Lavault, Jean-Franois Marckert et Vlady Ravelomanana. Quasi-optimal energy-efficient leader election algorithms in radio networks. *Information and Computation*, 205(5) :679–693, 2007.

[ocad :R07] Vlady Ravelomanana. Optimal gossiping and initialization protocols for random wireless multihop network. *IEEE Transactions on Parallel and Distributed Systems*, 18(1) :246–257, 2007.

[ocad :LaLo06] Christian Lavault et Guy Louchard. Asymptotic analysis of a leader election algorithm. *Theoretical Computer Science*, 359(2) :239–254, 2006.

[ocad :R06a] Vlady Ravelomanana. The average size of giant components between the double-jump. *Algorithmica*, 46(3) :529–555, 2006.

[ocad :ABFM05] Foto Afrati, Euripidis Bampis, Lucian Finta et Ioannis Milis. Scheduling trees with large communication delays on two identical processors. *Journal of Scheduling*, 8(2) :179–190, 2005.

[ocad :BS05] Cyril Banderier et Sylviane Schwer. Why Delannoy numbers? *J. Statist. Plann. Inference*, 135(1) :40–54, 2005.

[ocad :BuBuLa04] Marc Bui, Franck Butelle et Christian Lavault. A distributed algorithm for the minimum diameter spanning tree problem. *Journal of Parallel and Distributed Computing*, 64(5) :571–577, 2004.

[ocad :R04] Vlady Ravelomanana. Extremal properties of three dimensional sensor networks with applications. *IEEE Transactions on Mobile Computing*, 3 :17–28, 2004.

Communications invitées dans une conférence internationale

[ocad :LaOlRa08] Christian Lavault, Stephan Olariu et Vlady Ravelomanana. Random geometric graphs and wireless network. *Proc. of the 1st Advanced Research Workshop on Information Security in Wireless Networks (ARW 2006)*, volume 13 de *NATO Security through Science Series D*, pages 18–51, 2008.

Communications dans des conférences internationales avec comité de lecture

[ocad :BuFiHa07] Franck Butelle, Lucian Finta et Mourad Hakem. Real-time distributed scheduling of precedence graphs on arbitrary wide networks. *Proceedings of the 21st IEEE International Parallel and Distributed Processing Symposium International Parallel and Distributed Processing Symposium*, page 220 (article sur CDROM). IEEE Computer Society, 2007.

[ocad :CaRaVa07] Jean Cardinal, Vlady Ravelomanana et Mario Valencia-Pabon. Chromatic edge strength of some multigraphs. *Proceedings of The IV Latin-American Algorithms, Graphs and Optimization Symposium – LAGOS 2007*, 2007. to appear in Electronic Notes in Discrete Mathematics.

[ocad :BG06] Cyril Banderier et Bernhard Gittenberger. Analytic combinatorics of lattice paths : Enumeration and asymptotics for the average area. *Mathematics and computer science, IV (Nancy, 2006)*, Discrete Mathematics & Theoretical Computer Science Proceedings, pages 345–356. 2006.

[ocad :CDR06] Christophe Cérin, Jean-Christophe Dubacq et Jean-Louis Roch. Methods for partitioning data to improve parallel execution time for sorting on heterogeneous clusters. Yeh-Ching Chung et José E. Moreira, éditeurs, *Advances in Grid and Pervasive Computing*, volume 3947 de *Lecture Notes in Computer Science*, pages 175–186, Taichung, Taiwan, mai 2006. Springer.

[ocad :LPF05] Christian Lavault, Gabriel Paillard et Felipe Franca. A distributed prime sieving algorithm based on smer. *4th International Symposium on Parallel and Distributed Computing*, pages 139–146, ISPDC'2005, 2005.

[ocad :R06b] Vlady Ravelomanana. Distributed k-clustering algorithms for random wireless multihop networks. *Proceedings of the IEEE ICN 2005*, pages 109–116, 2005.

[ocad :BBR04] Cyril Banderier, Jean-Marie Le Bars et Vlady Ravelomanana. Generating functions for kernels of digraphs (enumeration & asymptotics for a constraint from game theory). *Proceedings FPSAC'04 (Vancouver, 2004)*, Formal Power Series and Algebraic Combinatorics, pages 91–105. (electronic), 2004.

[ocad :DLPR04] F. Djerourou, C. Lavault, G. Paillard et V. Ravelomanana. Assigning codes in a random wireless network. *International Conference on Telecommunications*, numéro 3124 Lecture Notes in Computer Science, pages 348–353, Fortaleza, Brasil, 2004.

[ocad :HaBu04a] Mourad Hakem et Franck Butelle. A new on-line scheduling algorithm for distributed real-time system. *Proc. of the 3rd Int. Symp. and School on Advanced Distributed Systems (ISSADS)*, Lecture Notes in Computer Science, pages 241–251. Springer, 2004.

[ocad :S04] Sidi Mohammed Sedjelmaci. The accelerated euclidean algorithm. *Proc. of the Meeting on Computer Algebra and Applications (EACA)*, pages 283–287, 2004.

[ocad :BBM03] Cyril Banderier, René Beier et Kurt Mehlhorn. Smoothed analysis of three combinatorial problems. *Mathematical foundations of computer science 2003*, volume 2747 de *Lecture Notes in Comput. Sci.*, pages 198–207. Springer, Berlin, 2003.

Rapports

[ocad :ACDJ07] Heithem Abbes, Christophe Cérin, Jean-Christophe Dubacq et Mohamed Jemni. Performance analysis of publish/subscribe systems. RR 2007-XX, LIPN - CNRS UMR 7030 - Université Paris Nord, 2007.

3.5 Séminaire

Le séminaire de l'équipe OCAD se déroule à raison d'une séance toutes les deux semaines en moyenne. Les interventions d'invités extérieurs de tous horizons alternent avec les présentations de travaux de recherche récents de membres d'OCAD ou, le cas échéant, d'autres équipes du LIPN. De même, tout chercheur extérieur (à l'équipe ou au LIPN) intéressé par un exposé du séminaire est invité à y participer. Les doctorants doivent assister à ces séances de séminaire, qui s'inscrivent pleinement dans leur formation et participent de leur culture scientifique.

Chapitre 4
RCLN (Représentation des Connaissances et Langage Naturel)

Sommaire

4.1 Membres de l'équipe 150
4.1.1 Evolution de l'équipe 152
4.2 Orientation scientifique générale 153
4.3 Thèmes de recherche 154
4.3.1 Compréhension fine des textes 154
4.3.2 Sémantique de corpus 157
4.3.3 Ingénierie des connaissances textuelles 162
4.4 Bilan et perspectives de recherche 164
4.4.1 Une équipe renforcée 164
4.4.2 Un dialogue renouvelé 165
4.4.3 De nouveaux défis 165
4.5 Synthèse 166
4.6 Animation de la recherche 167
4.6.1 Evaluation de la recherche 167
4.6.2 Organisation de la recherche 169
4.6.3 Contrats 170
4.6.4 Dissémination 173
4.6.5 Collaborations et Invitations 175
4.7 Formation à la recherche 176

4.7.1 Habilitations à Diriger les Recherches 176
4.7.2 Thèses soutenues 176
4.7.3 Thèses en cours 177
4.7.4 Encadrement de stages de recherche 178
4.7.5 Cours de DEA ou de Master 2 recherche 178
4.8 Publications principales 179

4.1 Membres de l'équipe

(*membres au 11/09/2007*, en italique membres quittant ou ayant quitté le LIPN)

Permanents

Nom	*Prénom*	*Situation*	*Institution*
 Audibert	Laurent	MCF	IUTV
 Després	Sylvie	MCF	IG
 Gayral	Françoise	MCF	IUTV
 Hamon	Thierry	MCF	IG
 Kayser	Daniel	PU	IG
 Lévy	François	PU	IUTV
 Nazarenko	Adeline	PU	IG
 Poibeau	Thierry	CR	CNRS
 Recanati	Catherine	MCF	IG
 Rozenknop	Antoine	MCF	IG
 Salotti	Sylvie	MCF	IUTB
 Szulman	Sylvie	MCF	IUTV

RCLN

Doctorants

Nom	*Prénom*	*Situation*	*Institution*
 Bossard	Aurélien	AM	Université Paris 13
 Bouffier	Amanda	ATER	Université Paris 13
 Boutouhami	Sara	Projet MICRAC	Université Paris 13
 Cailliau	Frederik	CIFRE	Sinequa
 Hamon	Olivier	CIFRE	ELDA
 Messiant	Cédric	AM-DGA	Université Paris 13
 Mondary	Thibault	AM	Université Paris 13
 Weissenbacher	Davy	ATER	Université Paris 13

Impétrants

Nom	*Prénom*	*Situation actuelle*	*Institution*
 *Bourgois*	*Laure*	*Enseignante*	*Ecole d'ingénieur (Dakkar)*
 *Nouioua*	*Farid*	*Post-doctorant*	*LRI*

Personnel temporaire du 1/9/2003 au 1/10/2007

Nom	*Prénom*	*Situation*	*Début*	*Fin*	*Financement*
 *Aubin*	*Sophie*	*Ingénieur de recherche*	*2/2004*	*12/2006*	*Contrat de recherche*
 Ceausu-Dragos	Valentina	ATER	2/2007		Contrat de recherche
 *Derivière*	*Julien*	*Ing. de recherche*	*5/2004*	*12/2006*	*Contrat de recherche*
 *Jacques*	*Marie-Paule*	*Post-doctorante*	*10/2006*	*08/2007*	*Contrat de recherche*
 *Vauvert*	*Guillaume*	*Post-doctorant*	*3/2004*	*8/2004*	*Contrat de recherche*
 *Zargayouna*	*Haifa*	*Post-doctorante*	*5/2006*	*11/2006*	*Contrat de recherche*
 *Zimina*	*Maria*	*Post-doctorante*	*5/2006*	*11/2006*	*Contrat de recherche*

Personnels invités du 1/9/2002 au 1/9/2006 (au moins 1 semaine)

Prénom	*Nom*	*Origine*	*Période*	*Financement*
Sofia	Ananiadou	Manchester Univ. & NaCTem (UK)	5-28 juin 2006	professseur invité
Anna	Korhonen	Cambridge Univ., Cambridge (UK)	15-30 mai puis 1-15 juillet 2007	rofessseur invité
Aïcha	Mokhtari	USTHB, Alger (DZ)	5-20 juillet 2004	Projet CMEP-CNRS
Aïcha	Mokhtari	USTHB, Alger (DZ)	10-30 juillet 2005	Projet CMEP-CNRS

RCLN

4.1.1 Evolution de l'équipe

- **Permanents**

Départs

- B. Biébow, MCF (thème ingénierie des connaissances), départ à la retraite (2005)

Promotion

- A. Nazarenko, promue de MCF à PU (thème traitement automatique des langues) (2005)

Arrivées

- L. Audibert, MCF IUTV (thème traitement automatique des langues), 2004 (thèse Aix-Marseille)
- A. Rozenknop, MCF IG (thème traitement automatique des langues), 2005 (thèse EPFL, Suisse)
- S. Després, MCF IG (ingénierie des connaissances textuelles), 2007 (mutation)

- **Doctorants**

Trois thèses ont été soutenues sur la période :
- Touria Aït El Mekki a terminé sa thèse de doctorat sur la « Construction semi-automatique d'index de fin de livres », soutenue le 21/12/2004 avec mention Très Honorable. Elle est maître de conférences à l'Université d'Angers depuis sept. 2005.
- Farid Nouioua (Dir. D. Kayser) a soutenu en avril 2007 sa thèse de doctorat intitulée « Extraction et Utilisation des Normes pour un Raisonnement Causal dans un Corpus Textuel »). Il est actuellement post-doctorant au LRI (Université Paris-Sud).
- Laure Bourgois (Dir. F. Lévy) a soutenu en juin 2007 sa thèse de doctorat sur la « Représentation et comparaison de Web services complexes avec des logiques dynamiques ». Elle est actuellement en poste dans une école d'ingénieurs à Dakkar.

Six étudiants sont en cours de thèse au sein de l'équipe RCLN :
- Davy Weissenbacher s'intéresse au problème de l'incomplétude et de l'imperfection des informations entrant dans la résolution de l'anaphore pronominale depuis 2003 (Dir. C. Fouqueré et A. Nazarenko). La soutenance est prévue fin 2007.
- Amanda Bouffier a commencé fin 2004 une thèse sur l'analyse de structures discursives au sein de documents procéduraux, afin d'aider leur modélisation (Dir. D. Kayser et T. Poibeau). Elle est actuellement ATER.
- Aurélien Bossard a commencé fin 2006 une thèse comme allocataire-moniteur. Il s'intéresse à l'évolution et la fusion d'informations extraites de données textuelles (Dir. D. Kayser et T. Poibeau).
- Cédric Messiant a débuté une thèse fin 2006 sur l'acquisition automatique d'informations linguistiques sur corpus. Il est financé par une bourse DGA (Dir. A. Nazarenko et T. Poibeau).
- Thibault Mondary, qui est en thèse depuis fin 2006 comme allocataire-moniteur, s'intéresse à la construction d'ontologies à partir de textes (Dir. A. Nazarenko et S. Després).
- Sara Boutouhami a commencé fin 2006 une thèse sur l'argumentation dans la langue naturelle (Dir. D. Kayser) qui est financée dans le cadre d'un projet ANR.

Deux doctorants de l'équipe font actuellement leur thèse dans l'industrie :
- Frederik Cailliau est en thèse chez Sinequa depuis 2004 (Bourse Cifre - Dir. A. Nazarenko). Il travaille à la conception et à la mise en oeuvre d'un système de gestion de ressources linguistiques.
- Olivier Hamon est en thèse à ELDA depuis 2005 (Bourse Cifre - Dir. A. Nazarenko). Il définit des méthologies et des protocoles d'évaluation des technologies de la langue et de la communication homme-machine.

- **Bilan**

Sur la période, l'équipe a augmenté son potentiel de recherche en traitement automatique des langues sur le thème « sémantique de corpus » (arrivée de Laurent Audibert en 2004 après celle de Thierry Poibeau en 2003), ce qui a permis de créer un pôle fort sur ce thème et de participer à des projets ambitieux, tout en renforçant le lien avec le thème « compréhension fine

de textes ». L'arrivée d'Antoine Rozenknop en 2005 est venu compléter ce pôle en lui apportant des compétences en analyse probabiliste, qui faisaient défaut.

Sylvie Després (mutation 2007) est venue renforcer le thème « ingénierie des connaissances textuelles » qui reste cependant faible numériquement.

Après quelques années creuses en 2003-2004, le nombre de thèses encadrées par l'équipe a sensiblement augmenté.

A signaler également le recrutement sur la période 2004-2007 de six ingénieurs de recherche ou postdoctorants financés sur des contrats de recherche (S. Aubin, J. Derivière, G. Vauvert, M. Zimina, M.-P. Jacques, H. Zargayouna), ce qui a permis à l'équipe de poursuivre ses travaux de recherche théorique tout en se lançant dans des tâches de développement. Ces derniers permettent de valoriser les résultats des recherches antérieures et de se doter d'outils pour mener de nouvelles expérimentations.

4.2 Orientation scientifique générale

Les connaissances sont très généralement transmises par l'intermédiaire du langage, ce qui justifie que la Représentation des Connaissances, enjeu majeur de l'Intelligence Artificielle et des Sciences Cognitives, ait pour objectif de traiter les connaissances exprimées en Langage Naturel. De manière complémentaire, la compréhension de textes met en oeuvre des connaissances et des raisonnements qu'il est fondamental de décrire et modéliser.

L'équipe s'intéresse donc au langage, non pour ses propriétés en tant que système formel de signes, mais pour son pouvoir expressif, et à la Représentation des Connaissances, notamment en tant qu'outil mis au service du traitement du Langage Naturel.

Les travaux de l'équipe comportent à la fois des recherches à caractère fondamental et des recherches plus appliquées. L'évolution des thèmes de recherche de l'équipe a mené à travailler dans trois directions principales :

Compréhension fine de textes Il s'agit étudier finement, pour le modéliser, le calcul sémantique entrant dans certains phénomènes interprétatifs (liés au pluriel, au temps, aux normes).

Sémantique de corpus Ce thème vise à développer des outils destinés à faciliter l'accès au contenu de documents techniques ou de vastes bases textuelles, comme le Web,

Ingénierie des connaissances textuelles Ce troisième thème a pour objectif de développer des outils d'ingénierie des connaissances à partir de textes.

Ces activités sont détaillées ci-dessous comme trois thèmes de recherche distincts. En réalité, elles sont fortement liées. L'objectif commun est de développer des méthodes permettant de modéliser, à des degrés divers, le contenu des textes et documents, soit pour y faciliter la navigation et la recherche d'information, soit pour en restituer l'information factuelle, soit encore pour construire un modèle du domaine. Les recherches plus fondamentales du thème « compréhension fine de textes » nourrissent la réflexion du thème « sémantique de corpus », même si la nécessaire robustesse des traitements sur corpus imposent des simplifications. Le thème « ingénierie des connaissances à partir de textes » exploite les résultats des deux premiers thèmes pour élaborer des ontologies fondées sur les connaissances textuelles. Ces thèmes sont également liés à ceux de l'équipe A3 pour tout ce qui concerne l'apprentissage à partir de données textuelles.

Ces questions sont abordées à travers l'étude de différents problèmes applicatifs et dans le cadre de projets collaboratifs variés. La diversité des domaines et des corpus étudiés permet de tester nos hypothèses de travail et de mettre en oeuvre nos méthodes sur des problèmes concrets tout en testant leur robustesse. Nous avons étudié plus particulièrement les problématiques suivantes :

Corpus de constats d'accident L'équipe RCLN poursuit ses recherches sur la compréhension des constats d'accident, notamment pour ce qui concerne l'analyse des normes qui sont implicites dans les textes mais qui en conditionnent l'interprétation (projet MICRAC).

Corpus de génomique Nous poursuivons également notre travail d'analyse de la bibliographie en génomique : il s'agit à la fois d'extraire des informations sur les interactions géniques (projet ExtraPloDocs) et et de développer un moteur de recherche spécialisé (projet ALVIS).

Documents médicaux L'ensemble de l'équipe s'est engagé dans la conception de méthodes et d'outils d'accès au contenu des documents médicaux, depuis deux ans et à la faveur du projet ACTMed. Ce projet est un Programme Pluriformations commun à trois laboratoires de l'Université Paris 13[1]. Il a été lancé en 2005 pour quatre ans. L'ambition de ce projet est de créer un pôle d'expertise rassemblant des compétences en traitement automatique des langues, ingénierie des connaissances, analyse de l'information médicale et linguistique. Ce projet permet de fédérer et de concrétiser des résultats obtenus à travers les trois thèmes de recherche de l'équipe RCLN.

Corpus de directives européennes Depuis deux ans, nous avons commencé à travailler sur le domaine juridique pour construction des ressources juridique, à la faveur de notre participation au projet LOIS[2]. Nous souhaitons développer ce nouveau champ de recherche dans les prochaines années, notamment au travers de collaborations avec des partenaires européens.

4.3 Thèmes de recherche

4.3.1 Compréhension fine des textes

Ce thème s'attache à rendre compte d'une compréhension fine des textes qui s'oppose aux analyses nécessairement plus simples requises pour un traitement de corpus et qui se définit pour nous comme la capacité à tirer du texte des inférences adéquates.

Dans cet objectif, et à travers l'étude de certains phénomènes interprétatifs complexes, comme la référence aux événements, le pluriel ou les normes, nous examinons la dynamique du processus interprétatif et les principes sur lesquels il repose. Nous en dégageons plusieurs idées fortes :

- L'hypothèse de compositionnalité, généralement retenue pour le calcul du sens des expressions complexes, se heurte à des difficultés majeures et il est préférable d'envisager d'autres formes de calcul sémantique plus plausibles cognitivement [CI-47].
- Ce processus interprétatif mobilise des connaissances de différents niveaux (sémantiques : lexicales et grammaticales, connaissances liées au domaine, au type du textes) qu'il faut représenter et dont il faut examiner les interactions.
- La prise en compte de l'incertain, la non monotonie du raisonnement et son caractère analogique sont essentiels.

RCLN

• Comprendre les relations temporelles

Nous nous intéressons à la compréhension des relations temporelles entre événements, qu'elles soient marquées syntaxiquement ou implicites. On peut à l'heure actuelle contraster la finesse des analyses linguistiques (sémantique lexicale et sémantique des événements, notamment) qui distinguent la position de l'événement sur l'axe des temps (passé, présent, futur), sa dynamique interne (état, événement, processus), son aspect (duratif ou ponctuel, borné ou indéfini, continu ou itératif) et une variété de relations de succession (succession immédiate ou distante, transitions progressives ou brutales), avec le caractère plus rudimentaire des représentations calculables (essentiellement l'axe des réels ou des entiers, ou les relations d'une algèbre d'intervalles).

C'est à cette complexité de la représentation des événements que s'est intéressé le projet dirigé par François Lévy dans le programme interdisciplinaire TCAN du CNRS. L'accent à été mis sur la reconnaissance d'événements itératifs et l'analyse/représentation de leur aspect. S'agissant d'événements simples (non itératifs), l'extension proposée par L. Gosselin de la théorie de Reichenbach à des intervalles permet de rendre compte assez finement de l'aspect avec des règles automatisables : à chaque événement est attaché un intervalle d'énonciation, un intervalle de procès et un intervalle de référence (celui sur lequel l'attention est attirée, et qui peut ne recouvrir que partiellement ou même pas du tout le procès). L'étude menée dans le groupe a

[1] Le LIPN, le Laboratoire de Linguistique-Informatique de Paris 13 (LLI) et le laboratoire d'Informatique Médicale et Bioinformatique (Lim&Bio)

[2] Le LIPN à été associé à ce projet, mais sans financement.

montré que l'itération met l'accent diversement sur les séries d'intervalles que l'on peut construire, en focalisant soit sur la série en tant que telle, soit sur l'intervalle global la contenant. On peut en rendre compte en combinant des constructeurs d'intervalles complexes. Les difficultés rencontrées sont liées à la délimitation des cadres de discours. Un corpus de 18 000 exemples de phrases itératives extraites du journal *Le Monde* a été réalisé.

Par ailleurs, une étude sur le temps et l'aspect dans le corpus des constats d'accident se poursuit. Catherine Recanati, en collaboration avec l'équipe A^3, explore la possibilité de faire apparaître des séquences de motifs verbaux, participant d'un niveau discursif de granularité frontière entre syntaxe et sémantique (les étiquettes sont par exemple le temps grammatical doublé de la catégorie aspectuelle lexicale du verbe). Les algorithmes utilisés relèvent de l'apprentissage numérique non supervisé (connexioniste et stochastique).

• Réexaminer le concept de cause, extraire des normes

Nous avons entrepris depuis quelques années de réexaminer le concept de cause. En effet, il nous semble que mettre à la base d'une théorie sémantique les notions de référence et de conditions de vérité n'est pas satisfaisant pour rendre compte du processus de compréhension. En revanche, beaucoup plus que la dichotomie du vrai et du faux, la compréhension semble utiliser le contraste entre le normal et l'anormal, et ceci est particulièrement visible lorsqu'on analyse l'usage courant de la causalité.

Daniel Kayser et Farid Nouioua ont repris le travail sur un corpus de constats d'accidents qui avait fait l'objet d'études antérieures au sein de l'équipe. Il apparaît que, parmi les causes d'accident qui peuvent être données (p.ex. la présence d'essence dans le réservoir est une condition nécessaire pour qu'une voiture roule, donc pour qu'elle dérape), seules celles qui sont perçues comme des anomalies (la présence d'une flaque d'huile sur la chaussée) sont effectivement pertinentes [CI-40]. D. Kayser et F. Nouioua essayent donc de rendre compte du raisonnement tenu par le lecteur pour trouver la cause de l'accident, au moyen de règles d'inférence non-monotones (voir ci-dessous) [RI-28]. La méthode sera validée si, sur un échantillon différent de celui ayant servi à faire l'analyse, le système identifie la cause de l'accident comme l'aurait fait un lecteur humain.

Cette méthode d'extraction des normes peut avoir des conséquences importantes en sémantico-pragmatique, car les inférences tirées de normes sont beaucoup plus riches que les inférences sémantiques habituelles, fondées sur la préservation de la vérité.

Une première synthèse de cette approche de la causalité, ainsi qu'une étude concernant les réseaux bayésiens, a été publiée [RE-10] dans un numéro spécial de revue sur le raisonnement causal, numéro dont le LIPN a pris l'initiative et qu'il a piloté.

Ce travail se prolonge actuellement dans le cadre d'un projet inter-disciplinaire soutenu par l'Agence Nationale de la Recherche (projet MICRAC : Modèles Informatiques et Cognitifs du RAisonnement Causal) qui associe un Laboratoire de Psychologie (Toulouse Le Mirail) et trois Laboratoires d'Informatique (IRIT, CRIL, LIPN) pour recenser différentes approches philosophiques, logiques, et informatiques de la causalité et tester celles qui correspondent le mieux au raisonnement causal humain.

• Interpréter le pluriel

L'étude de l'interprétation sémantique du pluriel que nous avons réalisée à partir d'exemples extraits d'articles de journaux nous a conduits au constat suivant : les problèmes sémantiques soulevés sont beaucoup plus divers et riches que ceux évoqués par la littérature traditionnelle traitant de ce sujet, loin de la seule dichotomie entre lecture distributive et lecture collective généralement relevée.

Nous avons mis en évidence le rôle du temps qui intervient de façon essentielle à la fois dans la construction des collections associées aux expressions plurielles et dans leur évolution. En effet, la compréhension de nombreux énoncés oblige les expressions plurielles à être considérées à différents instants. Françoise Gayral, Daniel Kayser et François Lévy ont montré que deux interprétations sont possibles [CO-51] : une interprétation que nous appelons *de re* où la collection est composée d'éléments fixes même si chaque élément a des propriétés qui évoluent dans le temps (énoncé 1 ci-dessous) et une interprétation *de dicto* où la collection persiste alors que ses

éléments changent. Dans l'énoncé 2 en effet, l'expression "les bacheliers" convoque deux réalisations distinctes de la même collection *de dicto* (les bacheliers de l'an dernier et les bacheliers de cette année). F. Gayral, D.Kayser et F. Lévy ont également étudié les différents cas où une partition de la collection en sous-ensembles était indispensable à l'interprétation de la phrase [CO-17].

1. *Les bacheliers de ma classe sont maintenant inscrits à l'université.*
2. *Les bacheliers sont moins nombreux cette année que l'an dernier.*
3. *Moins nombreux, les étudiants proviennent de plus en plus d'Europe et de moins en moins d'Afrique.*

F. Gayral, D.Kayser et F. Lévy ont par ailleurs montré comment les trois questions ci dessus (caractère *de re/de dicto* des collections, existence de partitions et critère sur lequel s'effectue le partitionnement, caractère distributif ou collectif de la (des) prédication(s) s'appliquant à l'expression plurielle) sont extrêmement liées et mobilisent des connaissances de différents niveaux. L'étude de nombreux énoncés attestés a montré qu'il est impossible de ne s'intéresser qu'à une de ces questions, de l'isoler et d'y répondre sans faire intervenir la réponse aux autres. Dans l'énoncé 3 par exemple, l'interprétation du seul pluriel "les étudiants" donne lieu à la création d'au moins six ensembles et à différentes prédications concernant ces étudiants, prédication collective pour leur dénombrement, distributive pour leur provenance.

• Représenter les procédures interprétatives

Dans beaucoup d'approches en linguistique computationnelle ou en sémantique formelle, interpréter une phrase est vu comme l'application d'un mécanisme récursif combinant des informations lexicales et syntaxiques. L'hypothèse sous-jacente est l'hypothèse de compositionnalité qui postule que le sens d'une expression complexe est déterminé par sa structure et par le sens de ses constituants.

Françoise Gayral, Daniel Kayser et François Lévy développent trois critiques majeures pour montrer que cette hypothèse est inapte à rendre compte de l'activité de compréhension [CL-11].

- En considérant la polysémie comme une énumération de sens, elle oblige à considérer le processus interprétatif comme un processus de désambiguisation parmi des listes prédéfinies de sens alors que de nombreux glissements de sens difficilement prévisibles peuvent intervenir en contexte.
- par le « localisme » et l'incrémentalité du processus qu'elle implique, elle ne fait dépendre le sens d'une expression complexe que du sens de ses constituants et de sa structure au détriment d'autres facteurs dont nous montrons au contraire l'importance.
- par le but qu'elle fixe au processus interprétatif (construire une représentation unique du sens de l'expression et donc de chacun de ses composants), elle empêche de rendre compte du phénomène de « coprésence » dans lequel un même mot peut prendre des sens différents suivant la partie de la phrase auquel on le rattache alors que cette coprésence est courante dans les textes.

A partir de cette critique, nous proposons une autre approche basée sur une logique non-monotone dans laquelle nous tentons d'exprimer les règles d'interprétation [CI-73]. Nous avons choisi la logique des défauts de Reiter par les avantages qu'elle procure. Par la capacité à exprimer des règles de défaut, ce formalisme permet de rétablir la notion d'exception dans la position importante qu'elle a en grammaire, position d'où les théories purement compositionnelles l'avaient évincée. Il autorise la non-monotonie du raisonnement, c'est-à-dire la remise en cause de conclusions jusque-là acquises. Grâce à la notion d'extension définie comme la solution d'une équation de point fixe, il rend compte de la circularité liée à l'influence du contexte sur l'interprétation d'un mot, sachant que le mot influe lui aussi sur l'interprétation de son contexte. En contrepartie, le principal problème de ce type d'approches est celui du contrôle : il est absolument nécessaire de bloquer la dérivation de certaines valeurs sémantiques si celles-ci s'avèrent inadéquates et, à l'inverse, d'en privilégier d'autres [RI-3].

En marge de ces travaux sur la logique des défauts, nous explorons d'autres voies pour modéliser la prise en compte de l'incertain dans les processus interprétatifs. Dans le cadre de sa thèse sur la résolution d'anaphore qui est un problème de traitement automatique épineux, Davy Weissenbacher montre que les modèles graphiques probabilistes comme les réseaux bayésiens

[RE-22*] sont adaptés pour modéliser un faisceau de connaissances à la fois incertaines, partielles et interdépendantes et les exploiter pour identifier l'antécédent d'un pronom anaphorique.

Nous nous intéressons par ailleurs aux capacités « analogiques » mises en oeuvre dans la compréhension du discours et dans le raisonnement humain. Plusieurs psychologues et philosophes ont souligné en effet de manière convaincante le caractère analogique de nos représentations cognitives. Pour élucider cette notion, C. Recanati s'est attachée à caractériser formellement la distinction texte/diagramme dans les raisonnements [RE-17].

4.3.2 Sémantique de corpus

En parallèle et en complément avec les analyses fines développées dans le cadre du thème précédent, nous explorons des méthodes d'analyses sémantiques adaptées au traitement de grandes quantités de texte.

Au-delà des techniques de recherche d'information qui associent un ensemble de documents à une requête, un enjeu majeur aujourd'hui concerne l'accès au contenu même des documents textuels. Il s'agit à la fois d'exploiter des documentations scientifiques et techniques et de maîtriser le contenu sémantique des pages web par l'ajout d'+s sémantiques interprétables par une machine (web sémantique) [CL-6]. Dans tous les cas, accéder au contenu des documents repose sur une analyse sémantique partielle et robuste du corpus.

Ces travaux s'inscrivent dans un courant de recherche important qui se focalise sur les moteurs de recherche sémantiques, les techniques d'extraction d'information et de question-réponse. Notre approche de ces problèmes est originale à plusieurs égards : nous mettons l'accent sur l'analyse de corpus spécialisés plutôt que sur des textes tout-venant ; nous articulons une analyse linguistique riche avec des techniques d'apprentissage et des méthodes d'acquisition de ressources ; nous cherchons à dépasser les frontières entre les techniques traditionnelles (recherche d'information, extraction d'information, question-réponse, navigation hypertextuelle) pour élaborer des approches plus riches de compréhension de textes.

- **Concevoir de nouvelles méthodes d'accès au contenu textuel**

En nous appuyant sur des outils d'analyse de corpus (voir plus loin), nous développons des méthodes facilitant l'accès au contenu textuel par le lecteur. Nous explorons en parallèle différentes approches, parfois sur les mêmes données, l'idée étant à terme d'arriver à combiner dynamiquement les différentes méthodes d'accès au contenu textuel [CL-8] :

Extraction d'information en biologie Cette tâche consiste, dans un domaine restreint, à extraire des éléments d'information précis à partir d'un ensemble de textes homogènes et à remplir des formulaires prédéfinis avec ces éléments d'information. Elle vise soit à alimenter des bases de données à partir d'informations extraites des textes, soit à ajouter des annotations sémantiques explicites aux documents. L'équipe RCLN travaille sur ces problème d'extraction depuis 1998 [LI-2], et plus particulièrement dans le domaine de la biologie depuis 2001. Notre participation à différents projets d'extraction d'information (ex. ExtraPloDocs) a montré le coût prohibitif de la tâche d'acquisition des ressources et notamment des règles d'extraction qui guident le processus d'extraction [CL-7]. Thierry Hamon, Adeline Nazarenko, Thierry Poibeau ont mis au point une méthode pour apprendre automatiquement ces ressources à partir de corpus (voir figure **??**) en collaboration avec Erick Alphonse (équipe A3) et des spécialistes de l'unité MIG de l'INRA [CO-60]. Les résultats obtenus montrent l'intérêt des méthodes d'apprentissage relationnel et l'importance de la normalisation linguistique préalable pour l'apprentissage des règles d'extraction. Cette normalisation permet de représenter les données linguistiques en faisant abstraction des variations paraphrastiques de surface (voir figure ci-dessous), ce qui met en évidence des régularités sémantiques et permet l'apprentissage de règles d'extraction.

Recherche d'information spécialisée Nos travaux dans cette direction sont plus récents. Dans le cadre du projet ALVIS, le moteur de recherche spécialisé en biologie exploite les résultats de la plate-forme d'annotation présentée ci-dessous, adaptée à la biologie [CO-110]. Nous cherchons tout d'abord à mesurer l'intérêt des annotations linguistiques et

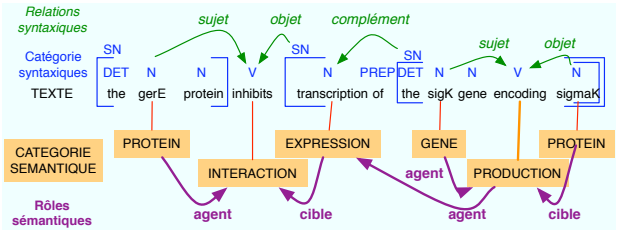


FIG. 4.1: Exemple de texte enrichi d'annotations linguistiques. Les différents niveaux d'annotation correspondent à des description de plus en plus abstraites et normalisées du matériau textuel.

sémantiques pour l'indexation des documents et pour la visualisation des informations dans l'interface utilisateur (si 'cat' est un nom de gène, il est important qu'il soit reconnu comme tel par un étiqueteur d'entités nommées pour éviter qu'il ne soit confondu avec l'animal 'chat' par un moteur de recherche dédié à la biologie). Les expériences correspondantes sont en cours. Sylvie Salotti et Haïfa Zargayouna étudient par ailleurs ce que la prise en compte de la structure même du document (structure XML) apporte à la recherche d'information. Dans le système qui a été développé par Haïfa Zargayouna, les termes sont désambiguïsés en fonction de leur contexte d'apparition [CO-36], lequel est défini par les éléments de structure des documents. Ce système permet d'effectuer une recherche ciblée. Au-delà du document, il donne accès directement aux points d'entrée pertinents dans le document.

Interrogation de documents Poser des questions et avoir directement la réponse sans avoir à lire des documents est une ambition ancienne du traitement automatique des langues. C'est ce que visent les systèmes de question-réponse. Nous avons mis l'accent sur les systèmes spécialisés. En collaboration avec Pierre Zweigenbaum (équipe INSERM U729, désormais LIMSI), Thierry Poibeau a développé un système de question-réponse médical et a participé à la campagne d'évaluation Technolangue EqueR (2004). Celle-ci a montré l'intérêt de la problématique des questions-réponses dans un domaine de spécialité. Nous pensons par ailleurs qu'exploiter la structure des documents permet de mieux guider l'interrogation. Sylvie Salotti et Haïfa Zargayouna ont défini un langage de requête par l'exemple où l'utilisateur peut spécifier les éléments de structure ainsi que le contenu dans un format XML et Victor Rosas montre comment traduire des requêtes en langue naturelle dans ce langage structuré. Ces travaux s'inscrivent dans cadre de la campagne 2006 d'Evaluation INEX (Initiative for the evaluation of XML Retrieval) mais ces travaux sont également testés en interne dans le cadre du projet ACTmed pour l'interrogation par les médecins des guides de bonne pratique médicale. La coopérativité des systèmes de question-réponse est aussi important pour leur meilleure utilisation et conviviabilité. Si la réponse exacte à une question n'est pas présente dans les textes, il peut être intéressant de proposer des réponses relâchant certaines contraintes de la question. Cette question de la coopérativité dans les systèmes de question-réponse a été abordée par Françoise Gayral en collaboration avec Patrick Saint-Dizier.

RCLN

Navigation documentaire Les méthodes d'analyse de corpus développées au sein de l'équipe RCLN, notamment d'analyse terminologique, nous ont permis de mettre au point des techniques originales d'accès au contenu dans les documents qui s'inspirent des approches traditionnelles en matière de consultation et de navigation documentaire. Le travail de thèse de Touria Aït El Mekki a conduit au développement d'IndDoc, un outil qui aide l'indexeur à produire un index de fin de livre [RE-15] [RI-26]. Ce travail se poursuit aujourd'hui en collaboration avec Laurent Audibert et Haïfa Zargayouna. L'originalité de IndDoc par rapport à d'autres outils d'indexation tient à la qualité de l'analyse du contenu documentaire qui apporte une aide substantielle à l'indexeur pour la sélection de l'information. Cette approche a fait ses preuves dans la construction de l'index de la version électronique d'un ouvrage scientifique (voir figure ci-dessous)et elle suscite aujourd'hui l'intérêt des indexeurs professionnels [RE-21]. Dans le prolongement de ces premiers

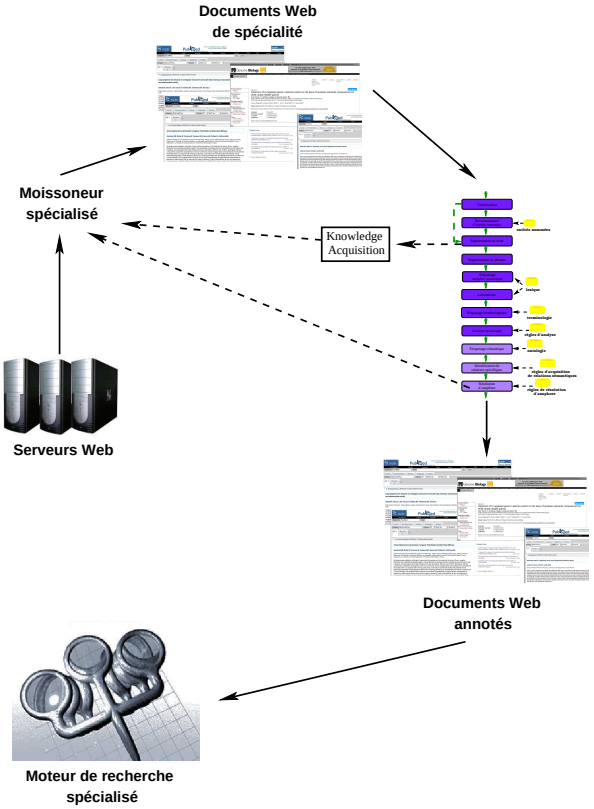


FIG. 4.2: Architecture de la plate-forme d'annotation de documents

RCLN

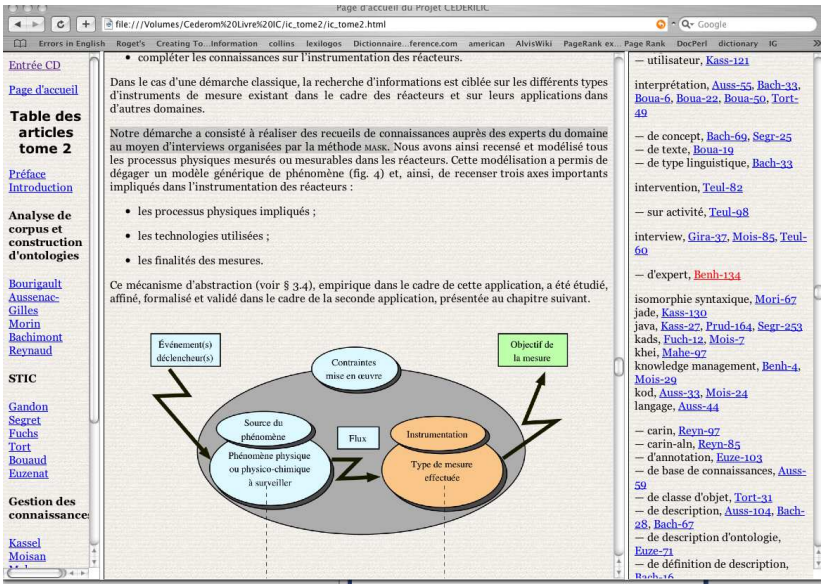


FIG. 4.3: Version électronique de l'index de l'ouvrage *Ingénierie des connaissances*, coordonné par R. Teulier *et al.*, L'Harmattan, 2005. L'écran est divisé en trois zones : la table des matières à gauche, l'index à droite et une fenêtre de navigation dans le document. Le lecteur a consulté l'index et à l'entrée, « interview d'expert », il a cliqué sur le lien « Benh-134» (indiqué en rouge), ce qui provoqué l'affichage du passage correspondant dans la fenêtre de navigation. Cette interface repose sur une indexation sous-jacente du document qui a été réalisée par IndDoc.

travaux, Amanda Bouffier et Adeline Nazarenko ont proposé une méthode de navigation qui s'appuie sur une analyse terminologique des documents eux-mêmes, sur leur table des matières et sur un modèle du domaine. L'outil de navigation interactif qui a été réalisé par Thibault Mondary est testé dans le cadre du projet ACTMed pour faciliter la consultation des Guides de Bonne Pratique Clinique (GBP) par les médecins.

RCLN

• Développer des outils d'analyse de corpus

Ces méthodes d'accès au contenu textuel s'appuient sur des outils d'analyse de corpus performants. Un effort important a été entrepris au sein de l'équipe RCLN pour développer des outils d'analyse de corpus qui permettent d'explorer de nouvelles approches sémantiques, d'analyser le contenu textuel ou d'amorcer l'acquisition de connaissances. Différents outils ont été réalisés à la faveur des projets dans lequels l'équipe RCLN a été impliquée depuis 2002 :

- une plate-forme d'annotation linguistique a été conçue et développée par Thierry Hamon, secondé par Julien Derivière (ingénieur de recherche contractuel) [Lo-12][3]. Cette plate-forme d'annotation linguistique peut produire différents types d'annotation, sur le français et l'anglais. Elle intègre les modules ci-dessous.
- divers modules d'analyse ont été réalisés : étiquetage des entités nommées (TagEN, Jean-François Berroyer et Thierry Poibeau), étiquetage terminologique (Thierry Hamon), analyse syntaxique spécialisée pour la biologie (BioLG, Sophie Aubin) [CI-95], désambiguïsation sémantique (Laurent Audibert), résolution d'anaphores (Davy Weissenbacher).

Nous avons mis l'accent sur la robustesse des traitements ainsi que sur la rapidité de l'annotation en proposant une architecture modulaire et distribuée. Thierry Hamon et Julien Derivière

[3]Cette plateforme d'annotation linguistique, développée en Perl, est disponible sous forme de module CPAN (Comprehensive Perl Archive Network – répertoire des modules Perl développés dans le monde) à l'adresse suivante : http ://search.cpan.org/ thhamon/Alvis-NLPPlatform-0.1/.

ont montré l'efficacité de cette approche lors d'une première expérience menée sur une collection de 55 329 documents en rapport avec la biologie et issus du web (soit plus de 80 millions de mots) [CI-117]. Cette expérience a prouvé que, même s'il reste coûteux, ce processus d'annotation permet de gérer le flux de documents qu'un moteur de recherche spécialisé doit pouvoir indexer. Elle a également montré que notre plate-forme est plus robuste (par rapport au nombre de documents à traiter, à leur taille, à la qualité du nettoyage préalable des données) que les plates-forme concurrentes comme celle de GATE.

Cette plate-forme est conçue pour être adaptée, sans chute de performance, à différents domaines ou à de nouveaux corpus par un processus de spécialisation. Cette spécialisation s'opère de deux manières :

- Par l'exploitation de connaissances spécialisées (dictionnaires d'entités nommées, terminologie, ontologie, notamment) : nous avons montré que l'ambiguïté de l'analyse syntaxique est réduite de 40% [4] quand les connaissances terminologiques peuvent être prises en compte [CI-89] ;
- Par le recours à des méthodes d'apprentissage à partir de corpus. Pour la désambiguïsation sémantique, Laurent Audibert a montré que l'impact de la qualité des indices fournis aux algorithmes d'apprentissage est bien plus important que la qualité intrinsèque des algorithmes employés. Les paramètres du module de résolution des anaphores peuvent eux aussi être adaptés pour un nouveau corpus par apprentissage : Davy Weissenbacher a modélisé le problème de la résolution des anaphores sous la forme d'un réseau bayésien dont les différents paramètres sont appris à partir des fréquences observées dans un corpus étiqueté [CI-91].

En marge de ces travaux qui reposent sur l'analyse linguistique et pour répondre aux besoins internes de nos travaux de recherche, L. Audibert a également développé des outils de manipulation et d'exploration de corpus : Concordanta, XExplora et Explora.

• Analyser la structure des documents textuels

Nos travaux sur l'accès au contenu textuel ont fait émerger une nouvelle thématique au sein de l'équipe RCLN. Il nous est apparu que la structure des documents est de première importance pour aider le lecteur à accéder à la bonne information, alors que l'accès à la documentation électronique s'est développé autour de méthodes fondées sur l'analyse de mots clés ou de segments isolés de textes, qui ignorent largement la structure des documents.

Partant de ce constat, nous avons commencé à travailler sur l'analyse de la structure des documents. Il s'agit à la fois d'exploiter les documents semi-structurés (en HTML ou en XML par exemple) pour compléter l'information portée par le contenu textuel et, inversement, d'analyser le texte lui-même pour découvrir des séquences porteuses de sens (titres, blocs informationnels cohérents, etc) et les mettre en relation.

A travers l'étude des Guides de Bonnes Pratiques médicales, choisis au sein du projet ActMed comme corpus de référence, Amanda Bouffier et Thierry Poibeau ont mis en évidence des structures sémantiques qui s'étendent au-delà de la phrase, comme des enchaînement « conditions-actions" et qui sont essentielles pour la lecture et la compréhension de ces documents. L'analyse de ces structures pose des problèmes linguistiques fondamentaux : comment repérer ces séquences de textes ? A partir de quelle combinaison d'indices ? Quelle est la portée d'une condition ? Le travail de thèse d'A. Bouffier vise à répondre à ces questions, à partir d'une analyse fine de documents médicaux. Il sera étendu afin de s'assurer de la généricité de l'approche [CO-130].

Ces recherches sont complétées par l'analyse fine d'autres éléments pertinents comme les titres. En fonction de leur contenu, nous cherchons à les typer selon une ontologie du domaine. Il s'agit d'aider à l'interprétation du document afin d'offrir des parcours de lecture pertinents et non linéaires, lors du passage au document électronique.

La prise en compte de la structure des documents doit permettre d'offrir de nouveaux moyens d'accès aux textes et d'enrichir les outils d'accès au contenu présentés ci-dessus (outils de navigation, notamment).

[4] L'ambiguïté est mesurée à partir du nombre d'analyses différentes produites par phrase.

4.3.3 Ingénierie des connaissances textuelles

Le troisième volet des recherches de l'équipe RCLN concerne l'acquisition de connaissances à partir de textes. Le développement de l'analyse de corpus a montré en effet qu'on peut s'appuyer sur le matériau textuel pour amorcer et guider le processus complexe d'acquisition qui vise à élaborer les ressources nécessaires aux systèmes à base de connaissances.

Nos travaux sur ce thème s'inscrivent dans un courant de recherche actif en France depuis une dizaine d'années, qui est en partie animé par le groupe de travail « Terminologie et Intelligence Artificielle » dont Brigitte Biébow, Adeline Nazarenko et Sylvie Szulman sont membres. Ces travaux rejoignent les recherches plus récentes qui sont menées au niveau européen dans la perspective de la construction d'un web sémantique. L'ancienneté de nos recherches et la coopération étroite entre le traitement automatique des langues et l'ingénierie des connaissances, donne une place originale à l'équipe RCLN dans ce paysage scientifique.

Nous mettons ici l'accent sur la constitution des bases de connaissances terminologiques et conceptuelles (ontologies).

• Construire des ressources terminologiques à partir de corpus

La construction des ressources lexicales constitue souvent le goulot d'étranglement des systèmes de traitement de corpus : les dictionnaires de langue sont généralement peu adaptés, les ressources lexicales spécialisées sont rarement disponibles et toujours difficiles à réutiliser. Nous développons des méthodes et des outils d'aide à la construction de ces ressources terminologiques pour un corpus et en fonction d'une tâche donnée. Nous travaillons dans deux directions spécifiques : l'extraction des termes et celle des relations qu'ils entretiennent.

Extraction de termes Sophie Aubin et Thierry Hamon ont proposé une méthode hybride qui combine l'acquisition de connaissances en corpus et l'exploitation de ressources terminologiques existantes. Les résultats montrent que la combinaison des deux approches permet d'obtenir de meilleurs résultats que l'une ou l'autre des deux approches prise isolément. Cette méthode a été implémentée dans l'extracteur de termes YaTeA (pour Yet Another Term ExtrActor) [CI-130]. Le système effectue une analyse syntaxique superficielle d'un corpus pour identifier des termes candidats et utilise en parallèle une ressource terminologique du domaine pour adapter le processus d'extraction de termes à ce domaine en reprenant en compte certaines de ses caractéristiques. L'exploitation de ressources lexicales ne va cependant pas de soi : il est important d'identifier leur adéquation au corpus. Adeline Nazarenko et Goritsa Ninova ont proposé une définition formelle de la couverture à partir de quatre métriques qui donnent une vue globale de la pertinence d'une ressource pour un corpus [CO-71].

Extraction de relations terminologiques Nos premiers travaux avaient porté sur le repérage de liens de synonymie à l'aide d'un dictionnaire de langue générale, sur le calcul de similarité sémantique, et sur l'extraction de relation d'hyperonymie à l'aide de patrons lexico-syntaxiques [RE-13]. Touria Aït El Mekki et Adeline Nazarenko ont poursuivi dans cette voie en montrant comment ces différentes méthodes peuvent collaborer pour construire un réseau sémantique de termes. On peut ainsi construire des index de documents structurés où les différentes entrées de l'index sont organisées en niveaux hiérarchiques [CI-90].

• Construire des ontologies à partir de corpus

Les méthodes de construction d'ontologies restent un thème de recherche important de l'équipe RCLN : il s'agit d'utiliser les textes commme sources de connaissances. Brigitte Biébow et Sylvie Szulman ont développé Terminae, une méthode et un outil d'aide à la construction d'ontologies à partir de textes, qui ont été présentés en 1999. Cette méthode s'appuie sur une analyse préalable des corpus qui repose sur les principes linguistiques et des logiciels de TAL présentés ci-dessus.

La méthode Terminae continue d'évoluer, d'une part dans la recherche d'éléments linguistiques permettant de modéliser les connaissances ontologiques, d'autre part dans l'intégration de ressources existante. B. Biébow et S. Szulman ont présenté dans [CL-12] un guide pratique de modélisation pour construire des ressources ontologiques à partir de textes. Le guide expose

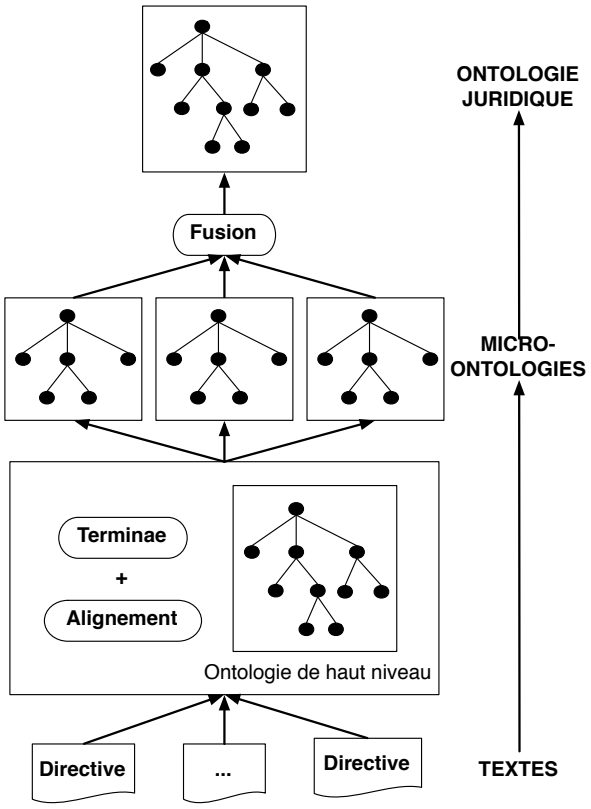


FIG. 4.4: Méthode de construction d'ontologie : application au domaine juridique.

en détail comment l'utilisation de logiciels de traitement automatique des langues et de modélisation peut aider un analyste lors du dépouillement et de la fouille dans les données textuelles. Il précise pas à pas les différentes étapes du travail en mettant en évidence la navigation entre les nombreux éléments de modélisation. La méthode devrait prochainement intégrer les relations linguistiques trouvées dans le corpus et utiles à la description d'un domaine [?].

L'outil Terminae a également évolué. Il permet désormais la visualisation des résultats des analyseurs syntaxiques tels Syntex[5] et Yatea []. Il intègre les résultats de Synoterm[6]. Le langage de description d'ontologie de Terminae autorise l'importation et l'exportation en OWL [CO-39]. La technologie de plugin a été étudiée et une première implantation a été créée.

Ces différents travaux ont été effectuées dans le cadre de projets collaboratifs (ATONANT, LOIS) ou partenariat bilatéraux (TENNAXIA).

RCLN

- **Réutiliser des ressources termino-ontologiques existantes**

Une autre évolution de la méthode Terminae devrait permettre d'intégrer des ontologies génériques préexistantes pour soulager le travail de modélisation. En collaboration avec Sylvie Després (CRIPS 5 - équipe IAD), Sylvie Szulman étudie différents aspects du problème : l'alignement des ontologies et l'intégration d'ontologies de haut-niveau comme (Dolce) [?]. La figure ci-après présente la méthode Terminae résultante qui a été testée essentiellement dans le domaine juridique, plus particulièrement l'étude de directives européennes [?].

Lorsque les ontologies produites sont utilisées pour l'analyse ou l'annotation de textes, au-delà de la fusion des ontologies acquises à partir de corpus et des ontologies préexistantes, se pose la question de l'articulation des connaissances qui servent à l'interprétation. Elles sont de

[5]Développé par D. Bourigault, ERSS.
[6]Développé par T. Hamon.

différents niveaux : connaissances terminologiques et ontologiques ressources morphologiques et syntaxiques utilisées pour l'analyse ou règles de désambiguïsation et d'extraction mobilisées dans les tâches d'étiquetage et d'extraction. L'articulation de ces différents niveaux de connaissances représente un défi important pour tout ce qui relève de l'annotation sémantique de documents et partant, du web sémantique. Cette question, que nous avons abordée de manière éparse à travers différents projets, constitue le sujet central de la thèse de Frederik Cailliau. Dans un contexte industriel, celui-ci développe un système complet de gestion des ressources linguistiques exploitées par le moteur de recherche de Sinequa, Intuition [CO-128]. Cette réflexion de fond, essentielle pour développer de véritables systèmes de compréhension de textes devrait être prolongée dans les prochaines années.

4.4 Bilan et perspectives de recherche

Cette section dresse le bilan des recherches menées au sein de l'équipe RCLN sur la période 2004-2007 et montre dans quelles directions ces travaux vont se prolonger dans les prochaines années.

4.4.1 Une équipe renforcée

La période considérée a permis à l'équipe RCLN de renforcer ses compétences et son potentiel de recherche.

- **Les moyens d'une recherche expérimentale**

L'équipe RCLN a conduit des recherches assez fondamentales en TAL depuis longtemps. La fin des années 1990 et le début des années 2000 ont été l'occasion d'ouvrir un nouvel axe autour de l'analyse de corpus qui comporte une part significative de travail expérimental. Un important effort de développement a donc été réalisé au cours de la période écoulée pour se doter d'un ensemble d'outils de traitement et de manipulation de corpus textuels ainsi qu'une plate-forme qui les intègre [CI-117]. Cet investissement facilite aujourd'hui la réalisation de prototypes de recherche, il permet de faire des expériences d'envergure sur corpus (analyse sémantique, construction d'ontologie) et de mieux valoriser nos travaux de recherche sous la forme d'outils opérationnels (système IndDoc, par ex. [RE-21]). Cet effort doit être poursuivi dans les prochaines années pour maintenir et faire évoluer les outils existants en fonction de nos besoins de recherche. L'équipe RCLN souhaite recruter un ingénieur de recherche permanent pour assurer la pérennité de ces développements et coordonner le travail des ingénieurs temporaires recrutés sur différents contrats.

RCLN

- **Un ensemble élargi de compétences du TAL à l'apprentissage**

Depuis quelques années, le LIPN a mis en avant un axe inter-équipes pour le développement de travaux combinant les techniques du TAL et de l'apprentissage. Cette collaboration s'est mise en place sur la période considérée, grâce au recrutement de maîtres de conférences avec des compétences mixtes (2 dans l'équipe RCLN, 1 dans l'équipe A^3) et dans le cadre de différents projets (apprentissage de règles d'extraction, par ex.). Depuis 2006, les deux équipes RCLN et A^3 sont engagées dans un projet commun (Infom@gic) pour le développement d'outils de traitement de l'information textuelle.

- **Une forte implication dans les projets collaboratifs**

Composée de 12 membres permanents auxquels viennent s'ajouter des doctorants, des chercheurs post-doctorants, des ingénieurs de recherche et des stagiaires, l'équipe RCLN a désormais atteint une taille critique. Sur la période considérée, elle a participé (parfois comme coordinateur) à de nombreux projets qui portaient aussi bien sur des axes de recherche fondamentaux que sur des thématiques plus appliquées. Cette politique doit se poursuivre dans les prochaines années : certains projets ont été lancés en 2006 (projet RNTL TextCoop autour de l'analyse de la typologie

des textes, projet ANR MICRAC sur l'analyse du raisonnement causal, projet Infom@gic pour la fouille de données textuelles), d'autres en 2007 (projet ANR RNTL DAFOe de développement d'une plateforme de construction d'ontologies) et d'autres encore doivent démarrer en 2008 (projet Quaero pour le développement de moteurs de recherche multimedia, projet ANR Masse de données CROTAL). La diversité de ces projets et, pour certains, leur envergure montrent que l'équipe RCLN est aujourd'hui reconnue sur l'ensemble de ses axes de recherche, aussi bien au plan national qu'au plan européen et international.

4.4.2 Un dialogue renouvelé

Si les années antérieures avaient permis à un nouvel axe de voir le jour, la période 2004-2007 a permis de renforcer le dialogue interne de l'équipe.

- **Un projet fédérateur concernant l'accès à l'information médicale**

Le lancement du projet ACTMed a permis à l'ensemble de l'équipe RCLN de collaborer sur un même projet en lien avec le LLI et le LIM&Bio. Ce travail a mis en évidence des synergies entre les différentes directions de recherche de l'équipe RCLN : les outils de recherche d'information et de navigation que nous avons proposés exploitent des techniques de TAL et d'ingénierie des connaissances mais ils reposent aussi sur une analyse fine de la sémantique des textes et du raisonnement médical. Ce projet a également fait émerger de nouvelles thématiques : le problème de la structuration des documents est ainsi apparu comme central dans la plupart de nos travaux sur les Guides de Bonne Pratique ; la question de la complémentarité des diverses techniques classiques d'accès à l'information (recherche et extraction d'information, question-réponse, navigation) est également apparue comme importante, l'objectif étant désormais de permettre à l'utilisateur de passer dynamiquement et naturellement d'une activité de recherche documentaire à la consultation de résumés ou à la navigation dans le document, voire à l'interrogation de documents [CL-8], dans une forme de « compréhension à profondeur variable ».

- **L'émergence de préoccupations transversales**

L'évolution des travaux de l'équipe a fait apparaître des questions communes aux différents thèmes de recherche.

A la suite des travaux sur le contenu des documents (analyse de la terminologie, des entités nommées, de la causalité, etc.), un travail récent a porté sur des structures discursives plus larges, comme les séquences conditions-actions au sein de documents procéduraux (voir section 4.3.2). Des besoins récurrents sont apparus autour de ce thème encore récent au sein de l'équipe. Cet axe de recherche va donc être approfondi. Ces travaux reposent sur une analyse précise des textes et devraient permettent à terme d'améliorer l'accès au document sous forme électronique.

Cet axe de recherche implique également l'identification de genres textuels, afin d'obtenir des ensembles de textes homogènes sur lesquels les outils d'analyse linguistique peuvent s'appliquer de façon pertinente. Cette problématique est étudiée au sein du projet TextCoop. Pour mener à bien ces recherches, il est apparu nécessaire de faire appel à la fois à une analyse fine des marqueurs textuels et à des méthodes plus robustes d'analyse de corpus.

Nos recherches sur la sémantique des verbes se situent également à la croisée de nos différents thèmes de recherche puisqu'il s'agit aujourd'hui d'apprendre des schémas prédicatifs à partir de corpus pour améliorer l'extraction d'information [CI-33] (voir ci-dessous).

4.4.3 De nouveaux défis

- **La compréhension fine, des questions anciennes toujours d'actualité**

Nos recherches antérieures en sémantique lexicale concernaient notamment le traitement de la polysémie des verbes. Elles s'attachaient à mettre au jour les influences croisées entre le verbe et ses arguments souvent nécessaires pour faire émerger l'interprétation de chaque composant. Les projets actuels de l'équipe RCLN amènent à relancer ces recherches dans une

perspective plus appliquée qu'initialement : on a besoin de prendre en compte la sémantique des verbes pour mettre en relation des entités nommées qui auraient été extraites des textes (extraction d'information) ou pour établir des relations ontologiques « transversales » (ingénierie des connaissances textuelles). Au-delà des indices de surface, nous proposons de prendre en compte une analyse sémantique locale qui repose sur l'acquisition et la reconnaissance de schémas prédicatifs en corpus. Une thèse est en cours sur ce sujet (Cédric Messiant).

Différentes recherches ont été menées au cours de la période autour de la problématique du raisonnement et de sa modélisation. Sur les différentes thèmes considérés (calcul des événements, identification des normes qui structurent le raisonnement, analyse du raisonnement médical), nous avons montré la puissance explicative de nos modèles et proposé des pistes pour articuler ces modélisations avec l'analyse linguistique. Ces travaux sont bien reconnus au plan international. Ils se poursuivent dans le cadre du projet MICRAC et avec la thèse de Sara Boutouhami.

- **L'ingénierie des connaissances, un rendez-vous à ne pas manquer**

L'équipe RCLN a investi depuis une dizaine d'années dans la problématique de la construction d'ontologies à partir de textes. Ces travaux sont aujourd'hui reconnus au niveau européen. Au-delà de la seule construction d'ontologies, nos recherches s'étendent désormais à l'ensemble des opérations de l'ingénierie ontologique (fusion, alignement d'ontologies) dont l'enjeu devient central avec le développement du web sémantique et des outils de recherche multimedia qui présupposent une annotation sémantique des documents. En dépit de la faiblesse numérique de cette thématique au LIPN, les collaborations extérieures ont permis de maintenir une activité importante (participation au projet européen LOIS, , projet RNTL ATONANT, partenariat industriel). Cette activité se développe encore à la faveur de projets où le LIPN apporte son expertise dans la construction de ressources termino-ontologiques (projet RNTL DAFOE et projet Quaero, prévu pour 2008). Les principaux enjeux concernent désormais l'intégration de ressources existantes et la coopération de méthodes hétérogènes dans la construction d'ontologies.

- **La sémantique de corpus à la rencontre de vrais enjeux**

L'émergence de moteurs de recherche sémantiques distribués lance de nouveaux défis au TAL. Dans la lignée du projet ALVIS et en lien avec les recherches actuelles sur les architectures distribuées de moteurs de recherche, nous cherchons à définir des modèles sémantiques permettant à la fois de caractériser un document au sein d'une base documentaire et un moteur de recherche dans son réseau de pairs, de manière à permettre un traitement intelligent de la requête de l'utilisateur qui ne sait pas nécessairement à quel moteur de recherche s'adresser. Ce projet fait appel aux techniques d'acquisition de ressources termino-ontologiques de l'équipe RCLN et ouvre de nouvelles pistes de travail sur la catégorisation de documents et l'alignement d'ontologies. Une étude préliminaire est en cours entre le LIPN (Antoine Rozenknop et Adeline Nazarenko) et l'EPFL (Martin Rajman) mais ce projet a vocation à être élargi à d'autres partenaires, si les premières pistes de travail se révèlent fructueuses.

RCLN

La maturité de nos travaux d'analyse de corpus nous invite également à repenser nos méthodes d'évaluation, l'objectif étant de mieux mesurer l'avancée du TAL et leur adéquation aux besoins applicatifs. Sur la période considérée, nous avons commencé à nous intéresser plus directement à ces problèmes d'évaluation que posent les techniques de TAL et d'ingénierie des connaissances textuelles [CL-3]. Cette réflexion a été initiée par notre participation à quelques campagnes d'évaluation et par l'organisation d'ateliers. Une thèse sur ce thème a été lancée en coopération avec ELDA. Cet effort va s'intensifier dans les années à venir. Une réflexion de fond a été initiée dans le cadre du groupe de travail TIA et l'équipe RCLN doit jouer un rôle actif dans l'évaluation au sein du projet Quaero sur certaines tâches de TAL (élaboration des protocoles d'évaluation et organisation de campagnes).

4.5 Synthèse

Très soudée autour des problématiques de la représentation des connaissances et du traitement automatique des langues, l'équipe RCLN présente une orientation scientifique originale dans

le paysage français. Même si elle a encore besoin d'être consolidée par le recrutement d'un maître de conférence ou d'un chargé de recherche en ingénierie des connaissances et par le recrutement d'un ingénieur de recherche permanent, l'évolution de l'équipe RCLN sur les dernières années (recrutement de nouveaux maîtres de conférences en TAL, financement sur contrat de personnel temporaire) lui a permis d'atteindre une taille critique : elle mène de front des recherches théoriques, des recherches plus appliquées et des travaux de développement quand nécessaire. Cette équipe jouit désormais d'une bonne reconnaissance au niveau local, national et européen, comme en témoignent, la diversité de ses collaborations, le nombre de ses publications et la participation à des projets d'envergure.

4.6 Animation de la recherche

4.6.1 Evaluation de la recherche

- **Comités de rédaction**

  - *Cognitive Science Quarterly*, 2000-2004 ; D. Kayser
  - *Intellectica*, 1996-2004 ; D. Kayser
  - *Traitement Automatique des Langues*, numéros spéciaux, 2006, 2007 ; A.Nazarenko, T.Poibeau

  Daniel Kayser fait en outre partie du comité éditorial d'une collection co-édité par Ophrys et la Maison des Sciences de l'Homme.

- **Comités de programmes**

  - EuroCogSci 07, *European Conference on Cognitive Science*, mai 2007, Delphes, Grèce ; D. Kayser (co-président)
  - RFIA, *Reconnaissance des Formes et Intelligence Artificielle*, 2004, 2006 ; D. Kayser (comité restreint)
  - TALN *Traitement Automatique des Langues Naturelles*, 2005 ; D. Kayser
  - IWCS *International Workshop on Computational Semantics*, 2005, 2007 ; D. Kayser
  - SARA *Symposium on Abstraction, Reformulation and Approximation*, 2005, Edinburgh, UK ; D. Kayser
  - ISPS *International Symposium on Programming and Systems*, 2005 ; D. Kayser
  - TIA *Terminologie et Intelligence Artificielle*, 2005, 2006, 2007 ; D. Kayser, A. Nazarenko, S. Szulman
  - RECITAL *Rencontre de Etudiants-Chercheurs en Traitement Automatique des Langues*, 2004, 2005, 2006, 2007 ; T. Hamon
  - Texte et Connaissances, colloque organisé dans le cadre de la semaine de la Connaissance, 2006 ; S. Després, S. Szulman
  - SemAnnot *International Workshop on Knowledge Markup and Semantic Annotation*, 2004-2006 ; S. Szulman
  - EKAW *Knowledge Engineering and Knowledge Management Managing Knowledge*, 2004-2005 ; B. Biébow
  - IC *Ingénierie des Connaissances*, 2004-2007 ; B. Biébow, S. Després, S. Szulman
  - OntoTexte *Atelier scientifique "Ontologies et Textes"*, 2007 ; A. Nazarenko
  - OLP *ECAI Workshop on Ontology Learning/Population*, 2004 ; A. Nazarenko
  - SMBM *2nd International Symposium on Semantic Mining in Biomedicine*, 2006 ; A. Nazarenko
  - PSB *New frontiers in biomedical text mining, PSB workshop*, 2007 ; A. Nazarenko
  - MEMURA *LREC workshop on Methodologies and Evaluation of Multiword Units in Real-world Applications*, 2004 ; A. Nazarenko
  - LREC *Language Resources and Evaluation Conference*, 2006 ; A. Nazarenko
  - KDLL *PKDD workshop on Knowledge Discovery in Life Science Literature*, 2006 ; A. Nazarenko
  - CIFT, *Conférence Internationale sur la fouille de textes*2004 ; A. Nazarenko
  - FinTAL *5th International Conference on Natural Language Processing*, 2006 ; A. Nazarenko

- RJCRI 2007, *Rencontres Jeunes Chercheurs en Recherche d'Information*, 2007 ; H. Zargayouna
- SAAKM *Semantic Authoring, Annotation and Knowledge Markup Workshop*, 2007 ; S. Szulman
- KW4ED Workshop *Knowledge Management and Semantic Web for Engineering Design Workshop*, 2007 ; S. Szulman
- RFIA *Reconnaissance des Formes et Intelligence Artificielle* 2008 ; D. Kayser (comité restreint)
- ISPS *International Symposium on Programming and Systems* 2007 ; D. Kayser
- SARA *Symposium on Abstraction, Reformulation and Approximation* 2007 ; D. Kayser
- KRAQ*Knowledge and Reasoning for Answering Questions* 2005 (satellite de IJCAI), 2006 (satellite de EACL) ; D. Kayser
- EGC *Extraction et Gestion des Connaissances*, 2005 et 2008 ; S. Després, S. Szulman
- Toth *Conférence Terminologie et Ontologie : Théorie et Applications*, 2007-2008 ; S. Després
- Atelier RAPC *Atelier Raisonnement à partir de cas*, 2005-2008 ; S. Després
- OTJ *Journée Ontologies et Textes Juridiques, organisée dans le cadre de la semaine de la connaissance*, 2006 ; S. Després, S. Szulman

• Comités de lecture

Les membres de l'équipe assurent régulièrement des relectures d'articles pour des conférences et des revues parmi lesquelles

- *International Joint Conference on Artificial Intelligence* (conférence)
- *European Conference on Artificial Intelligence* (conférence)
- *Data & Knowledge Engineering* (revue)
- *Human Language Technology* (revue)
- *Techniques et Sciences Informatiques* (revue)
- *Revue des Nouvelles Technologies de l'Information* (revue)
- *Traitement Automatique des langues* (revue)
- *Linguisticae Investigationes* (revue)
- *Information - Interaction - Intelligence* (revue)

RCLN

• Comités d'évaluation

- (membre) Jury d'attribution des prix de thèses européennes; Daniel Kayser, comme ECCAI Fellow.
- (membre) Comité national du CNRS, section 45 ; 2003 et 2004 ; Daniel Kayser.
- (président) Jury d'attribution du prix de thèse de l'ASTI, 2006 ; Daniel Kayser.
- (membre) Jury d'attribution du prix de thèse de l'ASTI, 2007 ; Daniel Kayser.
- (président) Jury d'admissibilité des directeurs de recherche (CNRS, section 45, 2003-04) ; Daniel Kayser .
- (président) Comité d'évaluation du LORIA (Univ. Nancy 1 et 2, CNRS & INRIA), 2004 ; Daniel Kayser.
- (président) Comité scientifique disciplinaire n°1 de l'ANR, 2007 ; Daniel Kayser
- (membre) Comité scientifique du Pôle "Intelligence Logicielle" ; Daniel Kayser (contrat Etat-Région Lorraine).
- (membre) Comité scientifique du Crip5 (Univ. Paris 5) entre 2001 et 2005 ; Daniel Kayser.
- (membre) Commission technique "Ingénierie des Connaissances" du Pôle de compétitivité CapDigital depuis 2005 ; Adeline Nazarenko
- (membre) Divers comités d'évaluation de laboratoires et de formations pour le MENESR ; Daniel Kayser.
- (membre et président en 2006) Jury d'attribution de Primes d'encadrement doctoral depuis plusieurs années; Daniel Kayser.
- (expert) Comité de préparation du 7ème PCRD, 2006 ; Daniel Kayser .

- **Expertises de projets**

  - ANR Programme *Blanc*, 2005, 2006 (D. Kayser), 2007 (A. Nazarenko)
  - ANR Programmes *RNTL* et *Technologies logicielles*, 2005, 2006, 2007 (A. Nazarenko)
  - ANR Programme *Masse de données*, 2006, 2007 (A. Nazarenko, T. Poibeau)
  - ANR Programme *Jeunes Chercheurs*, 2007 (A. Nazarenko)
  - ANR Programme *Corpus*, 2007 (A. Nazarenko, F. Gayral)
  - ANR Programme *TECSAN*, 2007 (A. Nazarenko)
  - Projets-contrats de la région de la Basse-Normandie (A. Nazarenko)
  - Projets de recherche canadien, 2005-2006 (A. Nazarenko)
  - Projets de recherche de la Swiss National Science Foundation, 2006 (A. Nazarenko)

- **Jurys**

  - Jurys de thèse (hors Université Paris 13) : D. Kayser (3 dont 2 comme rapporteur), F. Lévy (2 comme rapporteur), A. Nazarenko (10 dont 7 comme rapporteur et 1 à l'étranger), S. Szulman (2)
  - Jury d'Habilitation à diriger des Recherches (hors Université Paris 13) : D. Kayser (2 comme rapporteur)
  - Jurys de thèse (hors Université Paris 13) : S. Després (2 dont 1 comme rapporteur)

4.6.2 Organisation de la recherche

- **Organisation d'événements scientifiques et conférences**

Les membres de l'équipe RCLN ont organisé ou participé à l'organisation de divers événements scientifiques et conférences.

  - Participation à l'organisation du *Genic Interaction Extraction Challenge* associé au workshop LLL-ICML 2005, 6 équipes officiellement participantes
  - Co-organisation de l'atelier *Evaluation des méthodes de terminologie computationnelle*, associé à la conférence "Terminologie et Intelligence Artificielle (TIA)", Rouen, 2005, 50 participants.
  - Co-organisation de l' *International Joint Workshop on Natural Language Processing in Biomedicine and its Applications* associé à la "20th International Conference on Computational Linguistics (COLING)", 2004, Geneve, 70 participants.
  - Organisation de l'*International Workshop on Terminology design : quality criteria and evaluation methods (TermEval)*, may 2006, in conjunction with the *Language Resources and Evaluation Conference (LREC)*, 2006, Gènes, Italie, 30 participants.

- **Organisation de séminaires**

L'équipe RCLN organise depuis 2001 des séminaires « Connaissances, Langue et Informatique » à raison de trois séminaires par an en moyenne. Cette formule de séminaire « en duo » fait intervenir deux orateurs sur une demi journée et permet de confronter deux visions différentes sur un même sujet. L'objectif est que ces présentations donnent lieu à un dialogue non seulement entre les orateurs, mais également avec l'assistance. Le séminaire est destiné aussi bien aux chercheurs confirmés qu'aux étudiants s'initiant au domaine abordé. Ce séminaire est largement ouvert à l'extérieur. Selon les thèmes abordés, l'audience varie entre 20 et 50 participants.

L'équipe RCLN accueille également régulièrement des chercheurs qui viennent présenter leurs travaux. Ces séminaires, à visée plus interne, contribue à la vie scientifique de l'équipe.

4.6.3 Contrats

Type	Coordinateur	Resp. scient.	Nom	Financeur	Période	Durée (mois)	Budget global (Keuros)	Part LIPN (Keuros)
Bioinfo.		A. Nazarenko	Caderige	inter-EPST	11/02-10/03	12		50
RNTL	•	A. Nazarenko	ExtraPloDocs	MENESR	12/02-09/05	33	390	170
RNTL		S. Szulman	ATONANT	MENESR	09/02-11/05	38	671	31
RNTL		T. Hamon	CESART	MENESR	01/04-01/06	36	142	8
CMEP-CNRS		D. Kayser	Modèles de l'action	EGIDE	01/02-12/05	48	80	15
TCAN-CNRS	•	F. Lévy	OGRE	CNRS	07/03-12/05	30	40	10
STREP		A. Nazarenko	ALVIS	UE	10/04-12/06	36	2 258	210
PPF	•	A. Nazarenko	ACTmed	MENESR	10/06-12/08	48	200	85
ANR		D. Kayser	MICRAC	ANR	01/06-12/08	36	325	64
RNTL		F. Gayral	TextCoop	ANR	01/06-12/08	36	424	128
Cap Digital		A. Nazarenko	Infom@gic	DGE	12/05-11/06	12		48
RNTL		S. Szulman	DAFOe	ANR	01/07-12/09	36	800	95

Caderige *Autres partenaires : AÏDA-IRISA, Leibniz-IMAG, LRI, INRA-MIG, INRA-ENSAR.*

Caderige - *Catégorisation Automatique de Documents pour l'Extraction de Réseaux d'Interactions GEniques* - Ce projet visait au développement d'une application d'interrogation en langage naturel de la base Medline pour les biologistes, l'objectif étant dans un premier temps d'exploiter les notices bibliographiques portant sur la transcription des gènes. Une méthode fondée sur l'extraction d'information et l'acquisition de connaissances a été définie [CO-60]. Cette première étude s'est poursuivie dans le projet ExtraPloDocs (http://caderige.imag.fr/).

RCLN

ExtraPloDocs *Autres partenaires : MIG (INRA), Isoft*

ExtraPloDocs - *EXTRAction d'information pour l'exPLOration de la DOCumentation Scientifique* - Dans le cadre de ce projet, ont été réalisés des outils permettant de 1) extraire des informations en génomique à partir de grandes bases textuelles comme Medline ; 2) acquérir, à partir d'un échantillon de corpus, les ressources nécessaires à l'analyse des textes (dictionnaires d'entités nommées, terminologies, ontologies) ; 3) apprendre les règles d'extraction par des méthodes d'apprentissage relationnel. Le LIPN qui a assuré la coordination du projet et a principalement contribué aux deux premières tâches (http://www-lipn.univ-paris13.fr/RCLN/projets/ExtraPloDocs/).

ATONANT *Autres partenaires : INSA, LaRIA, LIP6 (Univ. Paris 6 & CNRS), EADS, CEA*

ATONANT - *ATONANT* (RNTL-Technolangue) - L'objectif du projet était de développer des outils d'aide à l'enrichissement semi-automatique d'ontologies destinés à une plate-forme modulaire. Le domaine d'application était la médecine, plus précisement l'enrichissement de ressources du CISMeF (Catalogue et Index des Sites Médicaux Francophones). Le LIPN a participé au couplage de l'outil Terminae avec un éditeur mettant en oeuvre la méthode OntoSpec développée au LaRIA. Nous avons aussi adapté Terminae pour traiter les données fournies par le CHU de Rouen et créé dans le langage de description de Terminae une ontologie relative à la

virologie. L'importation et l'exportation en OWL d'ontologies ont été implémentés afin de comuniquer nos résultats à nos partenaires dans un langage standard (http://atonant.insa-rouen.fr/).

CESART *Autres partenaires : ELDA, Cersates (Univ. Lille 3), EDF, Temis, RALI, CEA*

CESART - *Campagne d'Evaluation de Systèmes d'Acquisition de Ressources terminologiques* (RNTL-Technolangue) - Ce projet a permis de mener une campagne d'évaluation des outils d'acquisition de ressources terminologiques à partir de corpus écrits (http://www.technolangue.net/article.php3?id_article=200/). Le LIPN est intervenu dans la définition du protocole d'évaluation et comme fournisseur de logiciel évalué, SynoTerm, pour l'acquisition de relations sémantiques (synonymie) entre termes. La campagne d'évaluation s'est déroulée en juin 2005 sur deux corpus du domaine de la médecine (pages web de Santé Canada) et de l'éducation (numéros de la revue SPIRALE).

Modèles de l'action *Autres partenaires : Université des Sciences et Techniques Houari Boumediène, Alger*

Modèles de l'action - Ce contrat, conclu initialement pour prolonger le travail sur les modèles de l'action commencé dans les thèses d'Université (française) et d'État (algérienne) d'Aïcha Mokhtari sous la direction de Daniel Kayser a permis à la partie algérienne à reprendre des activités de recherche en liaison avec l'équipe RCLN mais aussi avec d'autres équipes du LIPN (des thèses seront bientôt soutenues sous la direction de Laure Petrucci, équipe LCR). De son côté, la partie française a repris ses travaux antérieurs sur la causalité avec notamment l'apport de Sara Boutouhami (étudiante de magister d'Aïcha Mokhtari, qui va commencer sa thèse dans l'équipe RCLN).

OGRE *Autres partenaires : Limsi (CNRS), GREYC (Univ. Caen & CNRS), Dyalang (Univ. Rouen & CNRS)*

OGRE - *Intervalles temporels et applications à la linguistique textuelle* - Ce projet a réuni des linguistes et des informaticiens et a jeté les bases d'une analyse et d'une représentation des phrases itératives (ex : « le lundi, il mange à la cantine »). Avec ses partenaires, le LIPN a montré la nécessité de prendre en compte types d'intervalles pour représenter des événements itératifs selon que l'accent est mis sur la continuité du procès global, sur son éclatement ou sur certains de ses composants. Deux formalismes ont été étudiés pour représenter ces différentes valeurs aspectuelles. A des fins d'expérimentation, le LIPN a constitué un corpus de 18000 phrases itératives annotées extraites du journal *Le Monde*. (http://www-lipn.univ-paris13.fr/~levy/TCAN/).

ALVIS *Autres partenaires : Helsinki Institute for Information Technology, Unité Mathématique, Informatique et Génome (INRA), Ecole Polytechnique Fédérale de Lausanne, Lund University, Technical University of Denmark, Index Data Aps, Exalead SA, ALMA Bioinformatica S.L., Jozef Stefan Institute, Tsinghua University*

ALVIS - *Superpeer Semantic Search Engine* - Ce projet vise au développement d'un moteur de recherche sémantique et distribué. Les aspects sémantiques assurent une meilleure adéquation des résultats aux requêtes formulées par les utilisateurs. La recherche distribuée repose sur la coopération de différents serveurs spécialisés dans des domaines différents. Le LIPN est responsable de l'enrichissement sémantique des documents et a réalisé la plateforme d'annotation de corpus pour l'anglais et le français. Le LIPN participe aussi à la tâche d'asquisition des ressources spécialisées qui, pour un domaine donné, sont utilisées par le module d'annotation : ont été développés un extracteur de termes, une méthode de spécialisation pour un parseur, un système d'apprentissage de règles de résolution des anaphores (http://www.alvis.info).

ACTMED *Autres partenaires : Lim&Bio (Univ. Paris 13), LLI (Univ. Paris 13 & CNRS*

ACTmed - *Accès au Contenu des Textes Médicaux : Analyse des besoins et développement d'outils* - Ce projet réunit trois laboratoires de l'Université Paris-Nord qui collaborent pour développer des méthodes d'accès à l'information textuelle médicale. Il s'agit en priorité de concevoir les outils qui permettent aux médecins de consulter les Guides de Bonne Pratique médicale produits par l'Agence Française de Sécurité Sanitaire des Produits de Santé (AFSSAPS). En complément de l'expertise des autres laboratoires dans l'ingénierie des systèmes de santé et la

linguistique, le LIPN analyse la structure des documents et le raisonnement médical sous-jacent et développe des prototypes intelligents d'interrogation et de consultation des guides.

MICRAC *Autres partenaires : IRIT (Univ. Toulouse 3 & CNRS), laboratoire Dynamiques Sociocognitives et Vie Politique (Univ. Toulouse 2), Centre de Recherche en Informatique de Lens (Univ. Artois & CNRS)*

MICRAC - *Modèles Informatique et Cognitifs du raisonnement Causal* - Ce projet pluridisciplinaire (http://www.irit.fr/MICRAC/), qui repose sur une étroite collaboration entre chercheurs en intelligence artificielle et en psychologie cognitive, a pour ambition la construction d'un système informatique capable de représenter des relations causales, de reconnaître ces relations dans un ensemble d'événements ordonnés, de fournir des explications quant au déroulement de ces événements, et d'attribuer des responsabilités à des agents parties prenantes de ces événements. Dans ce projet, l'équipe RCLN s'intéresse aux constats d'accident. Elle cherche à faire valider par des protocoles appropriés les conjectures effectuées jusqu'ici sur ce qu'un lecteur standard considère comme cause de l'accident, à comparer différentes théories formelles du raisonnement causal pour voir celles qui correspondent le mieux aux résultats observés, et enfin à implémenter un système de raisonnement causal opérant selon ces principes (thèse de Farid Nouioua à soutenir début 2007, et thèse de Sara Boutouhami commençant en octobre 2006).

TextCoop *Autres partenaires : IRIT(Univ Toulouse 3 & CNRS, SINEQUA*

TextCoop - *Typologies et grammaires de textes afin de rendre un TEXTe plus «COOPératif» en gestion de l'information* - Le projet TextCoop a pour but l'utilisation de typologies textuelles et de grammaires de textes dans des systèmes d'accès à l'information textuelle (moteurs de recherche, systèmes de question/réponse, extraction d'informations). Dans ce projet, le rôle du LIPN concerne l'analyse des types de texte (manuel utilisateur, recommandations médicales, documentation technique, inventaire, article scientifique, discours, etc.) et la conception des méthodes qui permettant de prendre en compte le type des textes dans des applications de recherche d'information. On sait en effet que la notion de type de texte influe sur certaines caractéristiques textuelles, que celles-ci concernent le lexique, la morphosyntaxe ou des modes d'organisation plus globaux.

Infom@gic *Autres partenaires : une vingtaine de partenaires industriels (Thales, EADS, ONERA, Bertin, Vecsys, Sinequa...) et institutionels (Universités Paris 6, Paris 8, Marne La vallée, GET, Limsi-CNRS.*

RCLN

Infom@gic - Ce projet s'inscrit dans le cadre du pôle de compétitivité francilien Cap Digital - IMVN (Image, multimédia et vie numérique) qui est consacré aux technologies de l'information et de la communication. Infom@gic vise à mettre en place, sur une période de trois ans, un laboratoire industriel de sélection, de tests, d'intégration et de validation d'applications opérationnelles des meilleures technologies franciliennes dans le domaine du traitement de l'information. Ce laboratoire s'appuie sur une plate-forme commune qui doit couvrir les grands domaines de l'analyse d'information quelles que soient les sources (données structurées, texte, images et sons) : la recherche et l'indexation, l'extraction de connaissances et la fusion d'informations multimédias. Le LIPN apporte à ce projet son expertise à la fois en apprentissage artificiel (équipe A^3) et en fouille de données textuelles (équipe RCLN).

DAFOe *Autres partenaires : INSERM UMR_S 872, ENST GET, IRIT, LIPN, LISI, Mondeca, Supelec, UTC.*

DAFOe - *Differential and Formal Ontologies Editor For Aapplications* L'objectif de ce projet est de proposer une méthode complète associée à une plateforme technique pour concevoir des ontologies, de la modélisation à partir du domaine à leur évolution en passant par leur formalisation et exploitation. S'appuyant sur les acquis de travaux antérieurs, à la fois issus des partenaires et de la littérature du domaine, le projet a pour but de prendre en charge la modélisation sémantique des concepts ontologiques (différentiels et formels) pour motiver et justifier les représentations formelles qui seront utilisées et en faciliter la révision. La plateforme technique DAFOe est un ensemble d'outils dont un éditeur d'ontologies qui prend en charge toute la question de la sémantique de ces ontologies, à travers des questions épistémologiques liées aux concepts

formels de haut niveau ? *i.e.* la top ontology ou ontologie formelle ? et, vis-à-vis de la composante métier, à travers des travaux sur les corpus textuels. Le but est ainsi de rendre disponible, à travers la plateforme, les méthodologies qui vont permettre de construire l'ontologie. On obtient ainsi une ontologie formalisée qui pourra être traitée dans un éditeur dontologie respectant les standards des langages d'ontologies du W3C (OWL).

4.6.4 Dissémination

- **Distinctions**

A. Nazarenko s'est vu décerner la médaille de bronze du CNRS (section 45, 2004).

- **Logiciels**

Terminae est une plate-forme d'aide à la construction de terminologies et d'ontologies à partir de l'étude de textes. Il apporte une aide à la modélisation d'un domaine à partir de textes. Il intègre un concordancier et les résultats d'outils de traitement de corpus[7]. La version actuelle permet l'importation d'une ontologie en RDFS et OWL et l'exportation d'une ontologie Terminae en OIL et OWL.

Distribution : contacter Sylvie Szulman.

IndDoc est un outil d'aide à la construction d'index de fin de livre. Il suit une méthodologie bien définie qui exploite l'apport de la terminologie computationnelle et intégre des outils liguistiques[8]. L'objectif de IndDoc est de mettre à la disposition des indexeurs (documentalistes, éditeurs, auteurs de livre, etc.) une nomenclature riche ainsi qu'une liste de renvois. Les indexeurs peuvent valider ou valider les entrées de l'index produit ainsi que spécifiifier un paramétrage (contraintes éditoriales, etc.) pour la mise en forme de l'index final (sous forme papier ou numérique).

Une présentation détaillée de l'outil ainsi qu'une démo sont disponibles sur le site Web IndDoc : http ://www-lipn.univ-paris13.fr/RCLN/outils/IndDoc/

TagEN est un logiciel de reconnaissance d'entités nommées à base d'automates à nombre fini d'états (reposant sur le logiciel Unitex de l'Université de Marne-la-Vallée). Des ressources (lexiques et grammaires) existent pour le français et l'anglais et des systèmes noyaux pour d'autres langues (ressources à étendre pour les rendre utilisables). Le logiciel est open source et fonctionne essentiellement sous Linux (les ressources sont utilisables sous Windows, linux et Mac). Voir http ://www-lipn.univ-paris13.fr/ poibeau/tagen.html.

Distribution : contacter Thierry Poibeau.

YaTeA (Yet another Term extrActor) identifie et extrait des groupes nominaux pouvant être des termes, *i.e.* des termes candidats. Chaque terme candidat est analysé syntaxiquement pour faire apparaître sa structure sous la forme de têtes et modifieurs. L'extraction de termes candidats repose sur une stratégie hybride par laquelle l'extraction à partir de patrons syntaxiques construits manuellement peut être guidée et corrigée à l'aide de ressources terminologiques existantes (désambiguisation exogène) [CI-130]. Ces termes attestés aident au repérage des groupes, à leur analyse syntaxique et participent également à l'extraction des termes candidats. YaTeA sera prochainement disponible sous licence GPL sous la forme d'un module CPAN. D'autres informations sont disponibles à l'adresse suivante : http ://www-lipn.univ-paris13.fr/ aubin/yatea.html

BiolG est une version modifiée de l'analyseur syntaxique Link Grammar Parser de Sleator et Temperley, adapté au domaine biomédical telle que décrite dans [CI-95]. BioLG présente deux extensions par rapport à la version originale de Link Grammar : il intègre un système de spécification de règles de "morpho-guessing" par expressions régulières et prend en compte un pré-étiquetage morpho-syntaxique du corpus à analyser.

[7] Lexter et/ou Syntex développés par D. Bourigault (didier.bourigault@univ-tlse2.fr), SynoTerm et Yatea développés au LIPN par T. Hamon et S. Aubin.

[8] Tels que YaTeA développé par S. Aubin et T. Hamon ou Faster développé par C. Jacquemin

BioLG est accessible et téléchargeable sous licence GPL à cette adresse : http://www.it.utu.fi/biolg/. BiolG a été réalisé en collaboration avec Sampo Pyysalo, Département de Technologie de l'Information de l'université de Turku, Finlande.

Ogmios La platforme d'annotation linguistique développé dans le cadre du projet ALVIS, offre la possibilité d'enrichir linguistiquement des documents textuels, rédigés en français ou en anglais. Ce processus d'annotation est à la fois robuste et rapide, du fait le l'architecture modulaire et distribuée.

La platforme propose plusieurs types de traitements linguistiques : reconnaissance d'entités nommées, segmentation en mots et en phrases, lemmatisation, étiquetage morpho-syntaxique, étiquetage terminologique, et analyse syntaxique (uniquement sur l'anglais, pour l'instant).

La plateforme d'annotation linguistique est développée en Perl. Elle est disponible en open-source, sous forme de module CPAN (http://search.cpan.org/~thhamon/Alvis-NLPPlatform-0.1/), ainsi que sous paquetage Debian[9].

SynoTerm SynoTerm est un logiciel de détection de relations de synonymie entre termes. Il peut être utilisé sur le français ou l'anglais, soit pour l'indexation libre d'un corpus de textes (calcul des relations de synonymie entre les termes extraits d'un corpus), soit pour l'indexation contrôlée d'un corpus de textes : il identifie les relations de synonymie entre des termes extraits d'un corpus et ceux issus d'une liste de termes contrôlés (terminologie, thesaurus).

L'interface de SynoTerm permet d'assister le travail de constitution de terminologie structurée : les relations proposées par SynoTerm doivent être validées par un utilisateur.

SynoTerm est développé en C et en Perl (utilisation de Perl/Tk pour l'interface graphique). Il sera bientôt disponible en open-source et sa mise à disposition sur SourceForge (http ://sourceforge.net/) est à l'étude.

http://www-lipn.univ-paris13.fr/~hamon/SynoTerm/SynoTerm.html

Concordanta Concordanta est un concordancier qui permet d'effectuer des recherches dans des fichiers préalablement enrichis d'annotations linguistiques. Cet outil repose sur la librairie LOX (C++). Il fonctionne sur les systèmes les plus courants (Linux/Macintosh/Windows).

Disponibilité : contacter L. Audibert.

XExplora et Explora Ces outils permettent de rechercher, d'observer des phénomènes linguistiques dans des corpus écrits et de formater librement le résultat des recherches. Ces applications exploitent la librairie C++, LOX, qui implémente le concept de méta-expressions régulières et son application sur des corpus. Ces logiciels fonctionnent sur les systèmes les plus courants (Linux/Macintosh/Windows).

Disponibilité : contacter L. Audibert.

RCLN

• Actions de vulgarisation

- T. Aït El Mekki : Fête de la science/Savantes banlieues, Université Paris 13, 2004 :démonstration ("Construction automatique d'index de fin de livres")
- A. Nazarenko : Fête de la science/Savantes banlieues, Université Paris 13, 2004, 2006, 2007 : mini-conférences ("La fouille de texte ou comment trouver une aiguille dans une botte de textes")
- A. Nazarenko : participation à un déjeuner de la technologie organisé par organisé par différents acteurs de la Seine Saint Denis sur le thème "Convergences technologiques - GESTIOgestion de documents électroniques : défis et innovation", 23 mars 2007 "

• Invitations, tutoriels, école d'été

- D. Kayser : conférence invitée au '6th International Symposium on Programming and Systems (ISPS)', 2003

[9] http://www-lipn.univ-paris13.fr/~hamon/ALVIS/Debian/testing/libalvis-nlpplatform-perl_0.1-1_all.deb

- T. Poibeau : conférence invitée au 'Workshop of the network of excellence on text mining and statistics (Nemis)', 2004
- D. Kayser : conférence invitée au '1st Indian Conference on Logic and Applications', 2005
- D. Kayser : conférence invitée à la journée scientifique 'Langage, origines et développement', 2005
- T. Poibeau : co-organisation de la table-ronde « Future trends in Question Answering » avec M. Minock, dans le cadre du workshop « knowledge and Reasoning for Question Answering » (Worlshop IJCAI KRAQ), 2005
- D. Kayser : participation à la table ronde 'Langage et Cognition' lors de l'atelier de préparation du 7ème PCRD, 2006
- A. Nazarenko : participation à la table ronde 'TAL, document et compréhension' animée par B. Victorri dans le cadre l'école de printemps 'Langue, connaissances, information', 2005
- A. Nazarenko : participation à la table ronde 'Evaluation of terminological tools and products' dans le cadre du 'Workshop lors de la journée 'Texte et connaissances' de la semaine de la connaissance, 2006
- A. Nazarenko et S. Szulman : animation et participation à la table ronde 'Marqueurs linguistiques et extraction de connaissances à partir de textes' lors de la journée 'Texte et connaissances' de la semaine de la connaissance, 2006

- **Sociétés savantes**

Les membres de l'équipe RCLN sont adhérents de différentes sociétés savantes. Ils ont joué un rôle actif dans l'ATALA (Association pour le Traitement Automatique des LAngues) et l'ARCo (Association pour la Recherche Cognitive).

4.6.5 Collaborations et Invitations

- **Collaborations internationales (hors contrats de recherche)**

- Coopération franco-algérienne : cette coopération a permis le séjour au LIPN pour des durées d'une à six semaines et des séjours d'une semaine à Alger pour 4 chercheurs du LIPN. Plusieurs publications ont été effectuées dans ce cadre [] et deux thèses sont en cours d'achèvement.
- Partenariat avec les Universités roumaines (Politehnica de Bucarest, Univ. Stefan Cel Mare de Suceava) dans le cadre de l'Université Virtuelle Francophone : ce partenariat a permis la réalisation d'un cours sur le web, la co-rédaction d'un livre [], la participation à des écoles d'été et diverses missions bilatérales.
- Protocole de coopération scientifique dans le cadre du projet européen eContent LOIS (Lexical Ontologies for legal Information Sharing), avec l'Italian National Research Council (IT), les Universités de Vienne (AT), Sheffield, Evora (PT), Leiden (NL), l'Académie des Sciences de la République Tchèque (CS) (http ://www.loisproject.org/modules/addresses), 2005-2006.
- Participation au projet STIC-Asie (coordonné par P. Saint-Dizier) : T. Poibeau a visité le NAIST (Nara, Japon) en mai 2006 et présenté les recherches du LIPN en matière de Traitement des Langues ; il donnera une série de cours lors de l'école d'été Language, Artificial Intelligence and Computer Science for Natural Language Processing applications (LAICS-NLP, Bangkok, 16-22 oct. 2006) sur les thèmes suivants.

- **Collaborations universitaires (hors contrats de recherche)**

- Coopération avec l'INSERM (U729) : participation conjointe à une campagne d'évaluation des systèmes de question-réponse (T. Poibeau)
- Participation à l'action spécifique RIP-WEB (recherche d'informations précises sur le web) coordonnée par le Limsi-CNRS (T. Poibeau)
- Participation au GDR 2521 "Sémantique et Modélisation" (C. Recanati)

- Participation au groupe de travail "Terminologie et Intelligence artificielle (TIA, AFIA, GDR I3) (B. Biébow, A. Nazarenko, S. Szulman)
- Membre du bureau du réseau thématique pluridisciplinaire « Langage naturel, logique et philosophie du langage » (section 34 et 35 du CNRS, 2005-2006) (A. Nazarenko)
- Participation au Groupe de travail Léonard (Acquisition du langage et grammaticalisation, A. morgenstern) : conseil pour l'informatisation des corpus – programme Childes et base de données Chat (T. Poibeau)

- **Collaborations industrielles (hors contrats de recherche)**

  - Contrat d'encadrement de master recherche avec la société TENNAXIA et en collaboration avec le LaRIA (Univ. Amiens), 2006 (S. Szulman)
  - Convention Cifre avec la société Sinequa (2004-2007, thèse de F. Cailliau)
  - Convention Cifre avec la société ELDA (2005-2008, thèse de O. Hamon)
  - Collaboration avec Isoft : expertise du LIPN pour l'intégration d'outils linguistiques au sein de la plate-forme de traitement de données Amadea (T. Poibeau)
  - Collaboration avec Exalead et encadrement d'un stagiaire de Master Recherche pour l'enrichissement automatique de thesaurus à partir de données issues du web (T. Poibeau)

- **Autres échanges universitaires**

  - Séjour d'un mois de T. Poibeau au Japon à l'Université de Tokyo et au National Institute of Informatics, oct-nov. 2004
  - Séjour de trois semaines de Sampo Pyysalo (Université de Turku, FI) au LIPN, juin 2006 : collaboration autour du développement d'un analyseur syntaxique adapté à l'analyse des textes de biologie [CI-95]
  - Séjour de 3 semaines en juin 2006 de Sophia Ananiadou (U. de Manchester, NaCTeM National Center on Text Mining) au LIPN comme professeure invitée

4.7 Formation à la recherche

4.7.1 Habilitations à Diriger les Recherches

RCLN

Brigitte Biébow HDR en informatique, Université Paris 13, « De Dasert à Terminae : voyage au pays de la modélisation des connaissances à partir de textes », soutenue le 9 juillet 2004
Coordinateur : D. Kayser
Situation actuelle du diplômé : retraite

Adeline Nazarenko HDR en informatique, Université Paris 13, « Donner accès au contenu des documents textuels Acquisition de connaissances et analyse de corpus spécialisés », soutenue le 14 décembre 2004
Coordinateur : D. Kayser
Situation actuelle du diplômé : Professeur à l'Université Paris 13

4.7.2 Thèses soutenues

Touria Aït El Mekki Doctorat en informatique, Université Paris 13, « Construction semi-automatique d'index de fin de livre », soutenue le 21 décembre 2004
Directeurs : D. Kayser et A. Nazarenko
Financement : allocataire-monitrice
Situation actuelle du diplômé : Maître de Conférences à l'Université d'Angers

Laure Bourgois Doctorat en informatique, Université Paris 13, « Représentation et comparaison de Web services complexes avec des logiques dynamiques », soutenue le 19 juin 2007 avec mention Très Honorable.
Directeur : F. Levy
Financement : 3 ans de contrat France-Télécom R&D, un an Ater
Situation actuelle de la diplômée : enseignante en école d'ingénieur à Dakkar.

Valentina Ceausu Doctorat en informatique, Université Paris 5, « Définition d'un cadre sémantique pour la catégorisation de données textuelles. Application à l'accidentologie », soutenue le 18 juin 2007
Directeurs : S. Despres
Financement : allocataire-monitrice - ATER
Situation actuelle du diplômé : ATER à l'IUT de l'Université Paris 13

Farid Nouioua Doctorat en informatique, Université Paris 13, « Extraction et Utilisation des Normes pour un Raisonnement Causal dans un Corpus Textuel », soutenue le 25 Avril 2007 avec la mention très honorable avec félicitations
Directeur : D. Kayser
Financement : allocataire-moniteur
Situation actuelle du diplômé : un post-doctorat à l'Université Paris 11-Orsay.

4.7.3 Thèses en cours

Rim Bentebibel inscrite en thèse d'informatique à l'Université Paris 5, depuis 10/2004
Sujet : « Adaptation dans le contexte du raisonnement à partir de cas textuel »
Directeur : S. Despres
Financement : allocataire-moniteur - ATER à l'université Paris 5

Aurélien Bossard inscrit en thèse d'informatique à l'Université Paris 13, depuis 10/2006
Sujet : « Evolution et fusion d'informations extraites de corpus en langage naturel »
Directeurs : D.Kayser et T. Poibeau
Financement : allocataire-moniteur

Amanda Bouffier inscrit en thèse d'informatique à l'Université Paris 13, depuis 11/2004
Sujet : « Structuration de textes pour l'aide à la modélisation »
Directeurs : D. Kayser et T. Poibeau
Financement : allocataire-moniteur
Situation actuelle : ATER

Sara Boutouhami-Nouioua inscrit en thèse d'informatique à l'Université Paris 13, depuis fin 2006
Sujet : « Etude de l'aspect argumentatif dans les textes d'accidents de la route »
Directeur : D. Kayser
Financement : projet ANR

Frederik Cailliau inscrit en thèse d'informatique à l'Université Paris 13, depuis 11/2004
Sujet : « Conception et mise en oeuvre d'un système de gestion de ressources linguistiques »
Directrice : A. Nazarenko
Financement : Bourse Cifre (Sinequa)

Olivier Hamon inscrit en thèse d'informatique à l'Université Paris 13, depuis 11/2005
Sujet : « Méthodes et les protocoles l'évaluation des technologies de traitement automatique des langues et de la communication Homme-Machine »
Directrice : A. Nazarenko
Financement : Bourse Cifre (ELDA)

Cédric Messiant inscrit en thèse d'informatique à l'Université Paris 13, depuis 10/2006
Sujet : « Acquisition automatique d'informations linguistiques sur corpus »
Directeurs : A. Nazarenko et T. Poibeau
Financement : bourse DGA et monitorat

Sonia Mhiri inscrit en thèse d'informatique à l'Université Paris 5, co-tutelle avec la faculté des sciences de Tunis, depuis 10/2005
Sujet : « Intégration d'ontologies »
Directeurs : S. Despres & E. Zagrouba
Financement : bourse de l'état tunisien

Thibault Mondary inscrit en thèse d'informatique à l'Université Paris 13, depuis 10/2006
Sujet : « Construction d'ontologies à partir de textes »
Directeurs : S. Després et A. Nazarenko
Financement : allocataire-moniteur

Davy Weissenbacher inscrit en thèse d'informatique à l'Université Paris 13, depuis 10/2003
Sujet : « Incomplétude et Imperfection des informations pour l'Extraction d'Information, un cadre applicatif : la résolution de l'anaphore pronominale »
Directeurs : C. Fouqueré et A. Nazarenko
Financement : allocataire-moniteur
Situation actuelle : ATER

4.7.4 Encadrement de stages de recherche

- 2003 (A. Nazarenko) : Z. Qian (DEA IA-OC, P13&P8), *Etude de la contribution de la terminologie au filtrage de fragments de textes*
- 2004 (A. Nazarenko) : G. Ninova (Master 2 recherche, P7), *Mesures de l'adéquation des ressources lexicales et ontologiques en corpus*
- 2004 (A. Nazarenko) : F. Cailliau (DEA IA-OC, P13&P8), *Vers une gestion intégrée des ressources linguistiques*
- 2004 (T. Poibeau) : J.-F. Berroyer (DEA IA-OC, P13&P8), *TagEN : un système de repérage des entités nommées*
- 2005 (S. Salotti) : I. Mansour (Master 2 recherche, P9), *Etiquetage sémantique et désambiguïsation d'un corpus médical*
- 2005 (S. Szulman) : F.-E. Calvier (Master 2 recherche, P13), *Alignement et fusion d'ontologies*
- 2005 (T. Poibeau) : A.-M. Renoux (Master 2 professionel, INaLCO), *Apport de la textométrie pour l'analyse de textes littéraires*
- 2006 (D. Kayser) : A. Gheerbrant (Master 2 recherche, P13), *Vers une formalisation non monotone des guides de bonne pratique médicale*
- 2006 (H. Zargayouna et S. Salotti) : V. Rosas (Master 2 professionnel, Univ. Aix-Marseille 2, *Analyse et structuration de requêtes en langage naturel pour l'interrogation de documents médicaux*
- 2006 (A. Bouffier et A. Nazarenko) : T. Mondary (Master 2 recherche, P13), *Etude de la navigation dans un document structuré. Application aux guides de bonne pratique médicale*
- 2006 (T. Poibeau) : A. Bossard (Master 2 recherche, P13), *Utilisation et acquisition de schémas prédicatifs pour l'extraction d'information*
- 2006 (C. Recanati et Y. Bennani) : N. Rogovschi (Master 2 recherche, P13), *Systèmes d'apprentissage non supervisé connexionnistes et stochastiques pour la fouille de données structurées en séquences*
- 2006 (T. Hamon) : J. Derivière (Master 2 recherche, P13), *Etude de l'impact du traitement automatique des langues en recherche d'information*
- 2006 (T. Poibeau) : T. Moal (Master 2 recherche, P13), *Utilisation de listes et de structures de pages web pour l'enrichissement de thesaurus*
- 2007 (D. Kayser) : N. Kamenoff (Master 2 recherche, P13), *Analyse sémantique, Modélisation et Visualisation :application aux constats d'accidents de la route*
- 2007 (S. Després et S. Szulman) : M. Drine (Master 2 recherche, P13), *Construction d'une ressource termino-ontologique à partir de propositions de scénarios d'aménagements urbains*
- 2007 (D. Weissenbacher et A. Nazarenko) : A.A. Bouchouareb (Master 2 recherche, P13), *Fouille de textes pour l'enrichissement d'une base de données de noms de gènes*
- 2007 (H. Zargayouna et A. Nazarenko) : M. Rakho (Master 1, P7), *Génération et visualisation d'index de fin de livre*

4.7.5 Cours de DEA ou de Master 2 recherche

A. Nazarenko est responsable du Master 2 Recherche en Informatique de l'Université Paris 13 (spécialité « Modélisation Informatique des Connaissances et du raisonnement »).

- B. Biébow, D. Kayser, F. Levy, A. Nazarenko, S. Szulman : cours de DEA IA-OC (P13) : *Ingénierie des connaissances textuelles*, 2003-2005.

- B. Biébow, S. Szulman : cours de DEA 127 (P9) : *Ontologies et Web sémantique*, 2003-2004.
- D. Kayser : cours de Master 2 recherche MICR (P13) : *Traitement de l'incertain et de la causalité*, 2005-07.
- F. Lévy : cours de Master 2 recherche MICR (P13) : *Formalismes de représentation des connaissances*, 2005-06.
- F. Gayral, D. Kayser, A. Nazarenko, A. Rozenknop : cours de Master 2 recherche MICR (P13) : *Traitement automatique des langues* [CL-17], 2005-06.
- T. Hamon, A. Nazarenko, S. Szulman : cours de Master 2 recherche MICR (P13) : *Ingénierie des connaissances textuelles*, 2005-07.
- R. Besançon, A. Nazarenko, T. Poibeau : cours de Master 2 recherche MICR (P13) : *Recherche et extraction d'information*, 2005-06.
- A. Nazarenko : cours de Master 2 recherche MICR (P13) *Méthodologie de la recherche*, 2005-06.
- S. Szulman : cours de DEA 127 (P9) : *Ontologies et Web sémantique*, 2004-2005.
- S. Szulman : cours de Master 2 (P9) : *Ingénierie des ontologies et Terminae*, 2005-2006.
- H. Zargayouna et A. Nazarenko : cours de Master 2 pro MÉTIERS DU LIVRE (P13) *Traitement du document numérique*, 2006, 2007.

4.8 Publications principales

Livres

[POIBEAU :2003 :CCSD-00005506 :1] Thierry Poibeau. *Extraction automatique d'information : Du texte brut au Web sémantique*. Lavoisier, 2003, ISBN 2-7462-0610-2.

Chapitres de livres

[NAZARENKO :2006 :CCSD-00097092 :1] Adeline Nazarenko. *Le point sur l'état actuel des connaissances en Traitement Automatique des Langues*, Gérard Sabah, éditeur, *Compréhension des langues et interaction*, pages 31–70. Hermès, 2006.

[BIEBOW :2005 :CCSD-00089085 :1] Brigitte Biébow, Sylvie Szulman et Nathalie Aussenac-Gilles. *Modélisation du domaine par une méthode fondée sur l'analyse de corpus*, R. Teulier ; J. Charlet ; P. Tchounikine, éditeur, *Ingénierie des connaissances*, pages 49–71. L'Harmattan, 2005.

[GAYRAL :2005 :CCSD-00084948 :1] Françoise Gayral, Daniel Kayser et François Lévy. *Challenging the Principle of Compositionality in Interpreting Natural Language Texts*, E. Machery, M. Werning et G. Schurz, éditeurs, *The Compositionality of Meaning and Content*, volume 2, pages 83–105. Ontos Verlag, 2005.

[NAZARENKO :2005 :CCSD-00098028 :1] Adeline Nazarenko. *Sur quelle sémantique reposent les méthodes automatiques d'accès au contenu textuel ?*, Anne Condamines, éditeur, *Sémantique et corpus*, pages 211–244. Traité IC2, série Cognition et traitement de l'information. Lavoisier, 2005.

[NEDELLEC :2005 :CCSD-00097093 :1] Claire Nédellec et Adeline Nazarenko. *Ontology and Information Extraction : a necessary symbiosis*, Bernardo Magnini Paul Buitelaar, Philipp Cimiano, éditeur, *Ontology Design and Population*, pages 155–170. IOS Press, 2005.

[POIBEAU :2005 :CCSD-00005507 :1] Thierry Poibeau. *Content annotation for the semantic Web*, Spiros Sirmakessis, éditeur, *Knowledge Mining*, pages 250–262. Springer, 2005. ISBN 3-540-25070-0.

[POIBEAU :2004 :CCSD-00085173 :1] Adeline Nazarenko et Thierry Poibeau. *L'évaluation des systèmes d'analyse et de compréhension de textes*, Stéphane Chaudiron, éditeur, *L'évaluation des systèmes de traitement de l'information*, pages 127–148. Hermès, 2004. ISBN 2-7462-0862-8.

RCLN

Articles dans des revues internationales avec comité de lecture sélectif

[KAYSER :2005 :CCSD-00091643 :1] Daniel Kayser et Farid Nouioua. About norms and causes. *International Journal on Artificial Intelligence Tools*, 14 :7–23, 2005.

[NAZARENKO :2005 :CCSD-00098030 :1] Adeline Nazarenko et Touria Aït El Mekki. Building back-of-the-book indexes. *Terminology*, 11(1) :193–218, 2005.

[KAYSER :2003 :CCSD-00091608 :1] Daniel Kayser. Abstraction and natural language semantics. *Philosophical Transactions of the Royal Society of London*, 358 :1261–1268, 2003.

Articles dans des revues nationales ou internationales

[RE-6*] Davy Weissenbacher. Les réseaux bayésiens : un formalisme adapté au traitement automatique des langues ? *Revue d'Intelligence Artificielle, Numéro spécial : "Modèles Graphiques Probabilistes"*, 2006. À paraître.

[ZARGAYOUNA :2006 :CCSD-00092307 :1] Haifa Zargayouna, Touria Ait El Mekki, Laurent Audibert et Adeline Nazarenko. IndDoc : An aid for the back-of-the-book indexer. *The Indexer*, 25(2) :122–125, 2006.

[RECANATI :2005 :CCSD-00085004 :1] Catherine Recanati. Raisonner avec des diagrammes : perspectives cognitives et computationnelles. *Intellectica*, 40 :9–42, 2005.

[AITELMEKKI :2004 :CCSD-00098041 :1] Touria Aït El Mekki et Adeline Nazarenko. L'index de fin de livre, une forme de résumé indicatif ? *Traitement Automatique des Langues*, 45(1) :121–150, 2004.

[GRABAR :2004 :CCSD-00084817 :1] Natalia Grabar et Thierry Hamon. Les relations dans les terminologies structurées : de la théorie à la pratique. *Revue d'Intelligence Artificielle*, 18/1 :57–85, 2004.

[LEVY :2004 :CCSD-00088705 :1] François Lévy. Statistiques et contrefactuels, quelques explications. *Intellectica*, 38 :187–191, 2004. Introduction à "Expérimentation, inférence statistique et analyse causale" de B. Lecoutre dans le même numéro.

Communications dans des conférences internationales avec comité de lecture

[AUBIN :2006 :CCSD-00090074 :1] Sophie Aubin et Thierry Hamon. Improving Term Extraction with terminological resources. Tapio Salakoski, Filip Ginter, Sampo Pyysalo et Tapio Pahikkala, éditeurs, *Advances in Natural Language Processing (5th International Conference on NLP, FinTAL 2006)*, LNAI 4139, pages 380–387. Springer, 2006.

[DERIVIERE :2006 :CCSD-00090079 :1] Julien Derivière, Thierry Hamon et Adeline Nazarenko. A scalable and distributed NLP architecture for Web document annotation. Tapio Salakoski, Filip Ginter, Sampo Pyysalo et Tapio Pahikkala, éditeurs, *Advances in Natural Language Processing (5th International Conference on NLP, FinTAL 2006)*, LNAI 4139, pages 56–67. Springer, 2006.

[PYYSALO :2006 :CCSD-00082533 :1] Sampo Pyysalo, Tapio Salakoski, Sophie Aubin et Adeline Nazarenko. Lexical adaptation of Link Grammar to the biomedical sublanguage : a comparative evaluation of three approaches. *Proceedings of the Second International Symposium on Semantic Mining in Biomedicine (SMBM 2006)*, pages 60–67. Sophia Ananiadou, Juliane Fluck, 2006.

[WEISSENBACHER :2006 :CCSD-00095189 :1] Davy Weissenbacher. Bayesian Network, a model for NLP ? *European Conference of the Association for Computational Linguistics (EACL 2006)*, pages 195–198, 2006.

[AITELMEKKI :2005 :CCSD-00098036 :1] Touria Aït El Mekki et Adeline Nazarenko. Using NLP to build the hypertextuel network of a back-of-the-book index. *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP)*, pages 316–320. Institute for Parallel Processing, Bulgarian Academy of Sciences, 2005.

[AUBIN :2005 :CCSD-00082542 :1] Sophie Aubin, Adeline Nazarenko et Claire Nédellec. Adapting a general parser to a sublanguage. G. Angelova, K. Bontcheva, R. Mitkov, N. Nicolov et N. Nikolov, éditeurs, *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP'05)*, pages 89–93. INCOMA Ltd., 2005.

RCLN

[KAYSER :2005 :CCSD-00091645 :1] Daniel Kayser. What logic can and cannot do in natural language semantics. *First Indian Conference on Logic and its relationship with other disciplines*, Actes en cours d'édition, Bombay, Inde, 2005.

[GAYRAL :2004 :CCSD-00092218 :1] Françoise Gayral, Daniel Kayser et François Lévy. Challenging the Principle of Compositionality in Interpreting Natural Language Texts. *Conference on Compositionality, Concepts and Cognition*, pages 111–127, Düsseldorf, Allemagne, 2004.

[KAYSER :2004 :CCSD-00091633 :1] Daniel Kayser et Farid Nouioua. Representing knowledge about norms. *Proc. 16th European Conference on Artificial Intelligence*, Valencia, Spain, pages 363–367. IOS Press, 2004.

[POIBEAU :2004 :CCSD-00009168 :1] Thierry Poibeau. Automatic extraction of paraphrastic phrases from medium size corpora. Sergei Nirenburg, éditeur, *Actes de la conférence Computational Linguistics (COLING'2004)*, pages 638–644. Association for Computational Linguistics, 2004.

Workshops et autres conférences avec comité de lecture

[BOUFFIER :2006 :CCSD-00085170 :1] Amanda Bouffier et Thierry Poibeau. Segmentation et structuration de textes procéduraux pour l'aide à la modélisation de connaissances : le rôle de la structure visuelle. *Semaine de la Connaissance*, pages 195–205, Nantes, France, 2006.

[CAILLIAU :2006 :CCSD-00098149 :1] Frederik Cailliau. Unifier la gestion de ressources linguistiques en contexte multilingue. *Actes de la 13ème conférence sur le Traitement Automatique des Langues Naturelles TALN'2006*, pages 455–461. Presses universitaires de Louvain, 2006.

[NAZARENKO :2006 :CCSD-00094950 :1] Adeline Nazarenko, Claire Nédellec, Erik Alphonse, Sophie Aubin, Thierry Hamon et Alain-Pierre Manine. Semantic annotation in the ALVIS project. *Proceeding of IIIA-2006 : International Workshop on Intelligent Information Access*, Helsinki, Finlande, 2006.

[NINOVA :2005 :CCSD-00084841 :1] Goritsa Ninova, Adeline Nazarenko, Thierry Hamon et Sylvie Szulman. Comment mesurer la couverture d'une ressource terminologique pour un corpus ? *Actes de la 12ème conférence sur le Traitement Automatique des Langues Naturelles TALN'2005*, pages 293–302. ATALA, 2005.

[ALPHONSE :2004 :CCSD-00098040 :1] Erick Alphonse, Sophie Aubin, Philippe Bessières, Gilles Bisson, Thierry Hamon, Sandrine Lagarrigue, Adeline Nazarenko, Alain-Pierre Manine, Claire Nédellec, Mohamed Ould Abdel Vetah, Thierry Poibeau et Davy Weissenbacher. Event-based information extraction for the biomedical domain : the Caderige project. Nigel Collier, Patrick Ruch et Adeline Nazarenko, éditeurs, *Proceedings of the International Joint Workshop on Natural Language Processing in Biomedicine and Its Applications (COLING'04)*, pages 43–49, Geneva, Suisse, 2004.

[GAYRAL :2004 :CCSD-00092223 :1] Françoise Gayral, Daniel Kayser et François Lévy. The role of time in considering collections. *Journées de Sémantique et Modélisation*, pages 30–34, Lyon, France, 2004.

[SZULMAN :2004 :CCSD-00087557 :1] Sylvie Szulman et Brigitte Biébow. OWL et Terminae. *IC'2004 - 15ème journées francophones d'Ingénierie des Connaissances*, pages 41–52. Presses Universitaires de Grenoble, 2004.

[ZARGAYOUNA :2004 :CCSD-00094619 :1] Haifa Zargayouna et Sylvie Salotti. Mesure de similarité dans une ontologie pour l'indexation sémantique de documents XML. *15ème journées francophones d'Ingénierie des Connaissances (IC'2004)*, pages 249–260. Presses Universitaires de Grenoble, 2004.

[GAYRAL :2003 :CCSD-00092208 :1] Françoise Gayral, Daniel Kayser et François Lévy. Collections évolutives et partitions. *Colloque "La partition en langue et en discours"*, France, 2003.

Outils

[DERIVIERE :2006 :CCSD-00090753 :1] Julien Derivière et Thierry Hamon. Alvis NLP platform. A scalable architecture using existing NLP tools to annotate large collections of web documents, 2006.

Partie III
Activités du laboratoire

Chapitre 1
Formation par la recherche

1.1 Master Informatique

Le Master (Sciences, Technologie, Santé) **mention Informatique** offre deux spécialités de formation qui s'appuient notamment sur des axes forts et des compétences bien développées au LIPN :

une spécialité à finalité professionnelle : Exploration Informatique des Données (EID)

une spécialité à finalité recherche : Modélisation Informatique des Connaissances et du Raisonnement (MICR)

La spécialité EID a pour objet l'acquisition de connaissances approfondies dans des domaines particuliers complémentaires : l'informatique et l'extraction de connaissances à partir de données. Cette spécialité est fondée sur l'acquisition de techniques destinées à favoriser l'exercice d'un métier déterminé dans le domaine de l'exploration des données.

La spécialité MICR a pour objectif de fournir aux étudiants les éléments fondamentaux leur permettant d'entamer une recherche tant en intelligence artificielle qu'en traitement automatique du langage, en spécifications de systèmes complexes ou en théorie de la programmation. Dans cette spécialité, plusieurs parcours-types sont possibles, on peut cependant en préfigurer quelques uns :

– Parcours « Traitement Automatique des Langues Naturelles » : Traitement automatique

des langues, Ingénierie des connaissances textuelles, Recherche et extraction d 'information, Ressources lexicales ;
- Parcours « Spécifications et Programmation » : Systèmes parallèles asynchrones, Systèmes infinis, temporisés et hybrides, Programmes et preuves, Unité d'approfondissement au choix ;
- Parcours « Logique et Programmation » : Calculs de processus, Programmes et preuves, Concurrence, Unité d'approfondissement au choix ;
- Parcours « Apprentissage Artificiel et Fouille de Données » : Apprentissage numérique, Apprentissage symbolique et RaPC, Apprentissage statistique, Fouille de données relationnelles.

Ce master est ouvert à la pluridisciplinarité qui est un des atouts de l'université Paris 13, avec par exemple une UE de linguistique sur les ressources lexicales et une UE sur la bio-informatique.

1.2 Master Mathématique Informatique

Le Master (Sciences, Technologie, Santé) **mention Mathématique et Informatique** a pour objectif de fournir aux étudiants les éléments fondamentaux leur permettant d'acquérir une véritable double compétence en Informatique et en Mathématiques et leurs applications. Elle offre l'accès à des cours très complémentaires dans les domaines de l'informatique, des mathématiques discrètes et continues, des phénomènes aléatoires et de la combinatoire. Elle permet, outre les aspects théoriques, de toucher aux problèmes applicatifs sources de modèles complexes. La formation s'appuie à la fois sur les laboratoire LIPN et LAGA de l'université Paris 13.

Outre les thèses possibles dans les laboratoires, en particulier au LAGA ou au LIPN, les étudiants peuvent se tourner vers les activités de recherche et développement en entreprise. Cette seconde voie n'est pas à négliger, les entreprises devenant un endroit où les compétences solides en mathématiques et informatique appliquées, attestées par une formation par la recherche, sont de plus en plus recherchées. Le master mathématiques et informatique vise à donner au travers de la formation le bagage indispensable d'un bon Ingénieur en Recherche et Développement ou d'un Consultant auprès d'entreprises de divers secteurs de pointe.

La première année est consacrée aux apprentissages des outils de base en mathématiques appliquées et en informatique. En fin d'année, un travail personnel d'étude et de recherche mené en coordination avec l'équipe pédagogique doit permettre à l'étudiant de mettre par lui-même en application les connaissances acquises pendant l'année. La seconde année est l'occasion d'approfondissement des connaissances en choisissant parmi de nombreux sujets : résolution numérique d'équations aux dérivées partielles, recherche opérationnelle, combinatoire, mathématiques financières.

1.3 Thèses et Habilitations

Thèses en cours

Formation par la recherche

Nom Prénom *Titre*	1ère insc. *Directeur*	Situation *Co-encadrant*
BENCHETTARA Hamimed *Prédiction de liens dans les réseaux sociaux*	12/01/07 *C. Rouveirol*	sans *R. Kanawati*
RODRIGUES BENTO Christophe *Apprentissage par renforcement relationnel*	10/01/07 *C. Rouveirol*	sans *P. Gérard*
CABANES Guénaël *Apprentissage non supervisé à base de noyaux pour la classification topographique évolutive*	10/01/07 *Y. Bennani*	Alloc (07-10)

CHEBALLAH Hayat	10/01/07	Alloc (07-10)
KAMENNOFF Nicolas	10/01/07	bourse Epitech
Utilisation d'une simulation 3D pour la compréhension de textes	*D. Kayser*	
SAIU Luca	10/01/07	Alloc (07-10)
Définition et compilation d'un langage de programmation logique et fonctionnel inspiré de la logique linéaire et de la théorie des jeux combinatoires	*C. Fouqueré*	*J.V. Loddo*
BOSSARD Aurélien	10/01/06	Alloc moniteur (06-09)
Evolution et fusion d'informations extraites de données en langage naturel	*D. Kayser*	*T. Poibeau*
BOUTOUHAMI Sarah	11/01/06	vacation ANR
Argumentation	*D. Kayser*	
RASONAIVO Andry	11/01/06	bourse AUF
Conception d'algorithmes efficaces dans les réseaux de télécommunication	*C. Lavault*	*S. Toulouse*
GROZAVU Nistor	11/01/06	contrat Cap Digital
Réduction de dimensions en apprentissage non supervisé	*Y. Bennani*	
ROGOVSCHI Nicoleta	11/01/06	contrat Septia
Apprentissage numérique et réseaux sociaux	*Y. Bennani*	*E. Viennet*
KENNICHE Hichem	11/01/06	Alloc moniteur (06-09)
	C. Lavault	*G. Duchamp*
SADKI Jalila	11/01/06	Alloc moniteur (06-09)
Méthodes de décomposition pour la résolution de problèmes de planification optimale de grande taille en transport	*A. Nagih*	*L. Alfandari*
MONDARY Thibault	10/01/06	Alloc moniteur (06-09)
Conception d'une méthode coopérative de mise à jour de règles d'extraction à base de corpus textuels	*A. Nazarenko*	
MESSIANT Cédric	10/01/06	Alloc DGA moniteur (06-09)
Acquisition semi-automatique de patrons d'extraction	*A. Nazarenko*	*T. Poibeau*
FKAIER Hazem	10/01/06	co-tutelle France-Tunisie
	C. Cérin	
ABBES Heithem	10/01/06	co-tutelle France-Tunisie
	C. Cérin	

Formation par la recherche

BERTRAND Olivier	02/01/06	CDI ONERA
Simulations distribuées de type HLA et reconnaissance d'activités à base de chroniques	*C. Choppy*	*P. Carle*
HAMON Olivier	12/01/05	CIFRE (Elda)
Méthodologies et protocoles d'évaluation des tyechnologies de la langue et de la communication homme-machine	*A. Nazarenko*	
ATASSI Vincent	10/01/05	Alloc moniteur (05-08)
Programmation fonctionnelle, typage en logique linéaire et réduction optimale	*J. Vauzeilles*	*P. Baillot*
CAILLIAU Frédérick	03/01/05	CIFRE Sinequa
Conception et mise en oeuvre d'un système de gestion de resources linguistiques	*A. Nazarenko*	*Claude de Loupy*
LAVAL Olivier	10/01/05	Alloc moniteur (05-08)
Approximation du problème de plus court chemin avec contraintes de ressources au sein d'un schéma de génération de colonnes	*A. Nagih*	*S. Toulouse*
TOUATI Nora	10/01/05	Alloc moniteur (05-08)
Ré-optimisation au sein d'un schéma de génération de colonnes	*A. Nagih*	*L. Létocart*
BOUFFIER Amanda	10/01/04	Alloc (04-07)
Evolution et fusion des informations extraites de corpus en langage naturel	*D. Kayser*	*T. Poibeau*
DEHNI Lahcène	10/01/04	(salarié m-temps) ; ANR
Apprentissage connexionniste non supervisé sous contraintes et réseaux de capteurs	*Y. Bennani*	
ANDRIAMAMPIANINA Tsiriniaina	10/01/03	$\frac{1}{2}$ ATER(05-06)
Modélisations algorithmiques des hypergraphes aléatoires	*C. Lavault*	*V. Ravelomanana*
WEISSENBACHER Davy	10/01/03	Alloc moniteur (03-06)
Normalisation syntaxico-sémantique de textes pour faciliter l'extraction d'informations : collaboration entre les techniques de TAL et d'apprentissage	*C. Fouqueré*	*A. Nazarenko*
LEMALADE Jean-Luc	11/01/02	Bourse franco-malgache (02-05)
Optimisation de la gestion des ressources agricoles en milieu forestier tropical	*G. Plateau*	*A. Nagih*

Formation par la recherche

Thèses soutenues

Nom Prénom	Situation ultérieure	Soutenance	Mention
Titre		*Directeur*	*Co-directeur*
Jury : Jury			
ELATI Mohamed		12/18/07	

Apprentissage de réseaux de régulation génétique à partir de données d'expression — *C. Rouveirol*

Jury : Jean-François Boulicaut (rapp), Jean-Daniel Zucker (rapp), Florence d'Alché-Buc, Younès Bennani, Amedio Napoli, François Radvanyi (invité)

ASEERVATHAM Sujeevan — 12/12/07 — THF

Apprentissage à base de Noyaux Sémantiques pour le Traitement de données textuelles — *Y. Bennani* — *E. Viennet*

Jury : Massih-Reza Amini, Christophe Fouqueré (pdt), Cyril Goutte, Jean-François Marcotorchino, Alain Rakotomamonjy, Michèle Sebag

KAROUI Hager — 12/11/07 — TH

Système coopératif de type égal-à-égal pour la recommandation : Application à la gestion et la recommandation de références bibliographiques — *L. Petrucci* — *R. Kanawati*

Jury : Jean-Paul Barthès, Catherine Berrut, Sylvie Després, Alain Mille, Enric Plaza, Laurence Vignollet

DJEROUROU Fayçal — ing R&D Intelligence Power — 06/15/07 — TH

Etude des méthodes de recherche locale pour la satisfaction de contraintes — *H. Bennaceur*

Jury : Gérard Plateau (Pdt), Jin-Kao Hao (rapp), Lakhdar Sais (rapp), Christophe Lecoutre, Christian Lavault, Monique Guignard-Spielberg

BOURGOIS Laure — CDD Sénégal — 06/19/07 — TH

Exploitation d'ontologies pour la combinaison automatique de composants logiciels — *F. Lévy* — *Alexandre Delteil*

Jury : Philippe Dague (rapp), Christophe Fouqueré (pdt), Laurent Hénocque, Farouk Toumani (rap)

NOUIOUA Farid — Post-doc LRI — 04/25/07

Extraction et utilisation de normes pour le raisonnement causal dans un corpus textuel — *D. Kayser*

Jury : Gérard Sabah (rapp), Pierre Marquis (rapp), Jean-Paul Haton, Pascal Nicolas, A. Nazarenko

KHEMOUDJ Mohand Ou Idir — Contrat EDF (05/06-02/08) puis post-doc — 04/04/07 — TH

Contribution de la programmation par contraintes à la résolution du problème de placement des arrêts des réacteurs nucléaires d'EDF — *H. Bennaceur* — *Marc Porcheron*

Jury : M. Porcheron (rapp), M. Schiex (rapp), C. Fouqueré, G. Plateau, M. Aggoun (invité)

HAKEM Mourad — Post-doc Lyon — 12/13/06

Ordonnancement Multiprocesseurs et Distribué Temps Réel — *C. Lavault* — *F. Butelle*

Jury : J. Carlier (rapp), C. Hanen (rapp), C. Laforest, E. Jeannot, C. Cérin

GUERIF Sébastien — MCF P13 — 12/11/06

Approches connexionnistes pour l'exploration et l'extraction de connaissances à partir de données spatio-temporelles — *Y. Bennani*

Jury : C. Bertelle (rapp), G. Venturini (rapp), P. Kuntz, M. Magnusson, JD Zucker, E. Ecosse (invité), E. Janvier (invité)

HAMDOUN Sana — 12/20/06

Formation par la recherche

Modélisation, maintenance et integration des données hétérogènes — *K. Lellahi* — *F. Boufarès*

Jury : Y. Bennani, C.F. Ducateau, D. Laurent

AMARA - HACHMI Nejla — post-doc FT — 12/06/06

GAMA : une architecture générique pour le développement d'agents mobiles sensibles au contexte et auto-adaptatifs — *A. El Fallah*

Jury : Olivier Boissier, Salima Hassas, Adeline Nazarenko, Henry Soldano

MARCOS Nicolas — ing SNCF — 11/23/06

Modélisation et intégration de la gestion tactique des ressources traction — *G. Plateau* — *A. Nagih*

Jury : Stéphane DAUZERE-PERES (rapp), Frédéric SEMET (rapp), Laurent ALFANDARI, , David DE ALMEIDA, Jacquelines VAUZEILLES, Monique GUIGNARD-SPIELBERG, Christian WEBER

MOUNGLA Hassine — Post-doc INRIA — 11/20/06

Supervision distribuée de la QoS dans les réseaux IP contrôlés par politique des opérateurs de telecommunications — *F. Krief*

Jury : A. Karmouch, N Agoulmine, M. Salaün, Y Bennani, K Chen

PAUCHET Alexandre — post-doc FT R&D Lannion ; MDC Rouen — 09/22/06 — TH

Modélisation cognitive d'interactions humaines dans un cadre de planification multi-agents — *A. El Fallah* — *Jean-Paul Sansonnet (Rapp)*

Jury : Sylvie Pesty (Rapp), D. Kayser, H. Soldano, N. Chaignaud

EL FERKH - JRAD Zeina — post-doc Inria (Axis) — 05/15/06 — TH

Apport des techniques de l'Intelligence Artificielle dans la négociation dynamique des niveaux de service : proposition d'une interface utilisateur pour l'Internet de nouvelle génération — *F. Krief*

Jury :

EL ZANT - KADHI Nahla — 12/01/05 — TH

Apprentissage de motifs relationnels dans des objets partiellement ordonnés et application à la caractérisation de motifs dans l'ADN — *D. Kayser* — *H. Soldano*

Jury :

KADRI-DAHMANI Hakima — post-doc IGN Belgique — 11/22/05 — TH

Méthodes de modélisation pour la gestion de la notion de cycle de vie dans les systèmes d'information géographique — *P. Dague* — *A. Osmani*

Jury : Younès Bennani , Cyrille Bertelle, Philippe De Mayer, Gérard Duchamp, Anne Ruas, Jan Beyen

Formation par la recherche

LOISEAU Irène — Pr Univ Buenos Aires — 06/29/05 — TH

Sur la génération de colonnes en nombres entiers — *G. Plateau*

Jury : Abdel Lisser (rapp), Nelson Maculan (Pdt), Anass Nagih, Catherine Roucairol (rapp), Jacqueline Vauzeilles

PAILLARD Gabriel — MCF brésil — 06/28/05 — TH

Quelques contributions à l'algorithmique distribuée. Réseaux de capteurs et génération de nombres premiers — *C. Lavault*

Jury : Michel Cosnard, Gérard Duchamp, Felipe Franca (rapp), Michel Habib (rapp), Colette Johnen (invitée), Nelson Maculan

AIT EL MEKKI Touria — MCF Angers (LERIA) — 12/21/04 — TH

Construction semi-automatique d'index de fin de livre — *D. Kayser* — *A. Nazarenko*

Jury : Béatrice Daille (rapp), Jean-Pierre Desclés (rapp), Jean Véronis

VAUVERT Guillaume — Post-doc INRIA-LIP6 — 01/15/04 — TH

Atteinte de consensus basé sur l'échange de préférences entre agents rationnels et autonomes : application à la formation d'alliances — *A. El Fallah-Seghrouchni*

Jury : Philippe Dague, Richard Emilion (rapp), Suzanne Pinson (rapp), Henry Soldano (invité), Patrick Taillibert, Jean-Daniel Zucker

ZEHRAOUI Farida — MCF Evry (LAMI) — 10/12/04 — THF

Systèmes d'apprentissage connexionniste et raisonnement à partir de cas pour la classification et le classement des séquences — *Y. Bennani*

Jury : Philippe Dague, Maurice Milgram, Alain Mille (rapp), Nicolas Nicoloyannis (rapp), Sylvie Salotti, Rushed Kanawati (invité), Eric Janvier (invité), Françoise Fessant (invité)

BENABDESLEM Khalid — Post-doc CNRS — 12/17/03 — TH

Approches connexionnistes pour la visualisation et la classification des données issues d'usages d'Internet — *Y. Bennani*

Jury : Sylvie Thiria (Pdt), Philippe Dague, Eric Janvier (invité), Dominique Laurent (rapp), Mohamed Quafafou (rapp), Adeline Nazarenko, Catherine Recanati (invitée)

GUERRA-HERNANDEZ Alejandro — Pr Univ Veracruzana — 12/06/03 — THF

Apprentissage d'agents rationnels BDI dans un univers multi-agents — *A. El Fallah-Seghrouchni*

Jury : Philippe Dague, Christian Lemaitre Leon (rapp), Yves Demazeau (Pdt), Joël Quinqueton (rapp), Henry Soldano (invité), Jean-Daniel Zucker

MALLAT-DESMORTIERS Didier — Ing — 01/24/03 — TH

Raisonnement automatique sur les croyances et les incertitudes d'un agent formalisé au sein de la théorie de l'interaction rationnelle — *Ph. Dague* — *Ph. Bretier*

Jury : Andréas Herzig (rapp), Daniel Kayser, Henri Prade (rapp), David Sadek

THIONGANE Babacar — Post-doc montréal — 02/24/03 — THF

Réoptimisation dans le dual lagrangien du biknapsack en variables 0-1 — *G. Plateau* — *A. Nagih*

Jury : Alain Billionnet (rapp), Monique Guignard, Nelson Maculan (rapp), Catherine Roucairol

CASTRO DE ANDRADE Rafaël — MCF Brésil COPPE Rio — 12/13/02 — THF

Synthèse de réseau à demande incertaine — *G. Plateau*

Jury : Pierre Fraignaud, Alexei Gaivoronski (rapp), Eric Gourdin, Abdel Lisser, Nelson Maculan, Michel Minoux (rapp), Anass Nagih, Jacqueline Vauzeilles

FOLCH Helka — 09/24/02 — TH

Articuler les classifications sémantiques induites d'un domaine — *D. Kayser* — *B. Habert*

Jury : Saadi Lahlou, Adeline Nazarenko, Bernard Victorri (rapp), Pierre Zweigenbaum (rapp)

Formation par la recherche

JUNGJARIYANONN Setthachaï — 03/20/02 — TH

Algorithmes de calculs sur les chaînes d'intervalles et leurs applications — *S. Schwer*

Jury : Younès Benanni, Françoise Gire (rapp), Mohand-Saïd Hacid (rapp), Elisabeth Metais, Jean-Marie Rifflet

POIBEAU Thierry — CR CNRS (LIPN) — 03/08/02 — TH

Extraction d'information à base de connaissances hybrides — *D. Kayser* — *A. Nazarenko*

Jury : Christian Fluhr, Christian Jacquemin (rapp), Patrick Saint-Dizier (rapp), Célestin Sedogbo, Yorick Wilks

FAURE Pierre-Philippe Henri — 06/18/01 — TH

Une approche à base de modèles fondés sur les intervalles pour la génération automatique d'arbres de diagnostic optimaux. Application au domaine de l'automobile — *Ph. Dague* — *L. Travé-Massuyès*

Jury : Jean-Marc David, François Lévy, Krishna Pattipati (rapp), Hervé Poulard, José Ragot (rapp)

BAUDOT Rémi — Ing Bea — 12/14/00 — TH

Programmation logique : non commutativité et polarisation — *J. Vauzeilles* — *C. Fouqueré*

Jury : Vito Michele Abrusci, Jean-Marc Andréoli , Andrea Asperti (rapp), Jean-Yves Girard (Pdt), Dale Miller (rapp)

BOUCHER Patrick — Ing — 07/05/00 — TH

Métaheuristiques pour l'optimisation combinatoire : thèse, antithèse, synthèse — *G. Plateau*

Jury : Arnaud Fréville (rapp), Jin-Kao Hao (rapp), Nelson Maculan, Vassilis Zissimopoulos

HAMON Thierry — MCF Paris 13 (LIPN) — 12/19/00 — TH

Variation sémantique en corpus spécialisé : Acquisition de relations de synonymie à partir de ressources lexicales — *C. Fouqueré* — *A. Nazarenko*

Jury : Henry Boccon-Gibod, Béatrice Daille , Benoît Habert (rapp), Jean Véronis (rapp)

RADEVSKI Vladimir — Ass. Pr univ Galatasaray — 06/21/00 — TH

Fusion de caractéristiques et de décisions dans les systèmes d'apprentissage connexionnistes — *Y. Bennani*

Jury : Fouad Badran, Dusan Cakmakov, Jacqueline Castaing, Bernadette Dorizzi (rapp), Srdjan Stankovic (rapp), Jacqueline Vauzeilles

Formation par la recherche

AFFANE Mohamed Salah — 02/03/99 — TH

Réseaux de contraintes : satisfaction et optimisation — *H. Bennaceur*

Jury : Jean-Jacques Chabrier (rapp), Mehmet Dincbas, Boi Faltings (rapp), Philippe Michelon (rapp), Gérard Plateau, Vassilis Zissimopoulos

BOSSAERT Fabrice — MCF Paris 13 (Lim&Bio) — 10/22/99 — TH

Approches connexionnistes pour le diagnostic de systèmes complexes : Application au réseau téléphonique — *Y. Bennani*

Jury : Béchir Ayeb (rapp), Philippe Dague, Elisabeth Didelet, Daniel Stern, Sylvie Thiria (rapp)

BOURREAU Eric — Ing Bouygues — 03/30/99 — THF

Traitement de contraintes sur les graphes en programmation par contraintes — *G. Plateau*

Jury : Alexander Bockmayr (rapp), Mehmet Dincbas , Odile Favaron, Jean-Pierre Lepape (rapp), Laurent Trilling (rapp), Michel Minoux , Vassilis Zissimopoulos

CHARNOIS Thierry — MCF Caen (Greyc) — 09/30/99 — TH

Détection d'anomalies sémantiques dans les textes de spécifications Application au problème des interactions de services de télécommunications — *D. Kayser* — *S. Szulman*

Jury : Pierre Combes, Philippe Trigano, Bernard Victorri (rapp), Pierre Zweigenbaum (rapp)

FERNANDEZ Rodrigo — Ing france — 06/25/99 — THF

Machines à vecteurs de support pour la reconnaissance des formes : propriétés et applications — *G. Plateau* — *E. Viennet*

Jury : Younès Bennani , Stéphane Canu (rapp), Patrick Gallinari (rapp), Monique Guignard, Michèle Sebag

NOBÉCOURT Jérôme — MCF Paris 13 (Lim&Bio) — 10/13/99 — TH

De la modélisation à la représentation : un exemple de langage de modélisation multi-formalismes — *D. Kayser* — *B. Biebow*

Jury : Nathalie Aussenac-Gilles, Gilles Kassel (rapp), Amedeo Napoli, Marie-Christine Rousset (rapp)

OSMANI Aomar — MCF Paris 13 (LIPN) — 12/17/99 — TH

Diagnostic de pannes dans les réseaux de télécommunications : approche à base de modèles et raisonnement temporel — *F. Lévy*

Jury : Philippe Balbiani, Luca Chittaro (rapp), Marie-Odile Cordier, Christophe Dousson, Malik Ghallab, Gérard Ligozat (rapp)

SEDJELMACI Sidi Mohamed — Enseignant — 09/30/99 — TH

Contribution au calcul parallele du PGCD de deux entiers — *C. Lavault*

Jury : Christiane Frougny, Tudor Jebelean (rapp), Ivan Lavallée (absent), Gérard Plateau, Kenneth Weber (rapp, excusé)

VICARD Annie-Florence — 07/05/99 — TH

Formalisation et optimisation des systèmes informatiques temps réel embarqués — *C. Lavault* — *Y. Sorel*

Jury : Marc Bui (rapp), Denis Trystram (rapp), Vassilis Zissimopoulos

YACOUB Meziane — MCF CNAM — 01/20/99 — TH

Sélection de caractéristiques et optimisation d'architectures dans les systèmes d'apprentissage connexionistes — *Y. Bennani* — *[A. Saoudi]*

Jury : Bruce Denby, Patrick Gallinari (rapp), Maurice Nivat (absent), Gérard Plateau , Sarunas Raudys (rapp)

BELHADJI Saïd — 09/23/98 — TH

Réseaux de contraintes temporels. Satisfaction de contraintes. Ordonnancement de tâches. — *G. Plateau*

Jury : Gérard Ligozat (rapp), Marie-Claude Portmann, Philippe Dague (Pdt), Virginie Gabrel

BRANKI Mohamed Tarek — 09/29/98 — H

Un processus d'intégration de bases de données spatiales hétérogènes par logiques de description — *N. Bidoit* — *B. Defude*

Formation par la recherche

Jury : Omar Boucelma (rapp), Robert Laurini (rapp), Pascal Coupey, Kazem Lellahi

GOUACHI Idir — Ing Cosytec — 09/30/98 — TH

LIRESNE, Langage pour Interpréter et Résoudre des Systèmes en Nombres Entiers — *G. Plateau* — *H. Bennaceur*

Jury : Jean-Jacques Chabrier (rapp), Philippe Michelon (rapp), Vassilis Zissimopoulos (Pdt), Mehmet Dincbas

GUIGNOT Jacques — MCF Paris 12 — 10/01/98 — TH

Application des stratégies neuronales au dilemme des prisonniers — *G. Plateau* — *[A. Saoudi]*

Jury : Didier Fayard (rapp), Arnaud Fréville (rapp), Nathalie Mourgue, Vassilis Zissimopoulos

NAULLEAU Elie — Ing — 01/05/98 — TH

Apprentissage et filtrage syntaxico-sémantique de syntagmes nominaux pertinents pour la recherche documentaire — *D. Kayser* — *B. Habert*

Jury : Jean-Pierre Desclés (rapp), Bernard Victorri (rapp), Adeline Nazarenko, Jean-David Sta (Pdt)

NEDJARI Tayeb — Ing Canada — 09/17/98 — TH

Réseaux de neurones artificiels et connaissances symboliques : insertion, raffinement et extraction — *D. Kayser* — *Y. Bennani*

Jury : Fouad Badran (rapp), Thierry Den ?ux (rapp), Dusan Cakmakov (rapp), Gérard Plateau

PERNELLE Nathalie — MCF Paris 11 (LRI) — 02/24/98 — TH

Traitement automatique de polysémies relationnelles : utilisation et contrôle de règles d'extension de sens — *D. Kayser*

Jury : Patrick Saint-Dizier (rapp), Victorri Bernard (rapp), Bernard Fradin, Françoise Gayral

SAHRAOUI Abderrazak — 10/22/98 — H

Modélisation et inférence des relations causales dans les systèmes physiques — *Ph. Dague*

Jury : Joseph Aguilar Martin (rapp), Luca Chittaro (rapp), Michel Dumas, François Lévy, Patrick Taillibert, Louise Travé-Massuyès (rapp)

BENJAMIN Didier — Ing — 10/16/97 — TH

Régularisation appliquée au traitement d'images — *F. Fogelman-Soulié* — *E. Viennet*

Jury : Gérard Plateau (Pdt), Bernadette Dorizzi (rapp), Patrick Gallinari (rapp)

Formation par la recherche

BIBAS Serge — Ing — 01/09/97 — TH

GASPAR : Simulateur à base de modèles pour la Supervision de Réseaux de Télécommunication — *Ph. Dague*

Jury : Malik Ghallab (Pdt), Henri Farreny (rapp), Sylviane Gentil (rapp), Marie-Odile Cordier, François Lévy

ISSAC Fabrice — MCF Paris 13 (LLI) — 09/02/97 — TH

Analyse syntaxique et apprentissage des langues — *D. Kayser* — *C. Fouqueré*

Jury : Thierry Chanier (rapp), Benoît Habert, Aravind Joshi, Bernard Lang (rapp)

MAABOUT Sofian — 06/02/97 — THF

Règles de mise à jour dans les bases de données — *N. Bidoit*

Jury : Françoise Gire (rapp), Jan Paredaens (rapp), Christine Froidevaux, Nicolas Spyratos, Jacqueline Vauzeilles

MOUSSOUNI Fouzia — 06/24/97 — TH

Construction, spécification et analyse d'algorithmes parallèles de recherche de mots — *C. Lavault*

Jury : Maxime Crochemore (rapp), Leszek Gasieniec (rapp), Jean-Frederic Myoupo (rapp), Gérard Plateau

VEIGA Geraldo — 06/13/97 — TH

Sur l'implantation des méthodes de points intérieurs pour la programmation linéaire — *G. Plateau*

Jury : Philippe Bourgeois, Monique Guignard, Nelson Maculan, Philippe Michelon (rapp), Mauricio Resende, Pierre Tolla (rapp)

VENTOS Véronique — MCF Paris 11 (LRI) — 04/01/97 — TH

Définition et apprentissage de concepts décrits à l'aide de propriétés "défauts" et de type "exception" — *D. Kayser* — *P. Coupey*

Jury : Joachim Quantz, Marie-Christine Rousset (rapp), Jean Sallantin (rapp), Pierre Siegel, Henri Soldano

BELABBAS Azeddine — Ing — 11/25/96 — TH

Satisfaction de contraintes et validation des grammaires du langage naturel — *D. Kayser* — *C. Fouqueré*

Jury : Gérard Plateau (Pdt), Bernard Levrat (rapp), Marie-Catherine Vilarem (rapp), Hachemi Bennaceur, Philippe Blache

BÖTTCHER Claudia — Ing Allemagne — 06/20/96 — TH

Un cadre pour guider la localisation automatique de défauts en utilisant des préférences stratégiques locales — *Ph. Dague*

Jury : François Lévy (Pdt), Marie-Odile Cordier (rapp), Peter Struss (rapp), Patrick Taillibert

BOUTHINON Dominique — MCF Paris 13 (LIPN) — 12/17/96 — THF

Apprentissage à partir d'exemples ambigus : étude théorique et application à la découverte de structures communes à un ensemble de séquences d'ARN — *D. Kayser* — *H. Soldano*

Jury : Philippe Dague (Pdt), Christel Vrain (rapp), Joël Quinqueton (rapp), Jacques Haiech

CHAIGNAUD Nathalie — MCF INSA Rouen — 01/25/96 — TH

Etude cognitive et informatique de la résolution de problèmes : analyse, modélisation et implantation — *D. Kayser* — *F. Lévy*

Jury : Anh Nguyen Xuan, Marie-Odile Cordier (rapp), Monique Baron (rapp), Suzanne Pinson (rapp)

NAGIH Anass — MCF Paris 13 (LIPN) — 06/21/96 — THF

Sur la résolution des programmes fractionnaires en variables 0-1 — *G. Plateau*

Jury : Nelson Maculan (Pdt), Monique Guignard (rapp), Pierre Hansen (rapp), Philippe Bourgeois, Arnaud Fréville, Kurt Spielberg

BALDY Bernard — Ing MasterSolutions — 10/17/95 — TH

Vérification, détection et correction syntaxico-sémantique dans le Traitement Automatique des Langues à partir du formalisme des grammaires syntagmatiques généralisées — *D. Kayser* — *C. Fouqueré*

Formation par la recherche

Jury : Gaston Gross, Jean Véronis (rapp), Pierre Zweigenbaum (rapp)

BRETIER Philippe — 12/15/95 — THF

La communication orale coopérative : contribution à la modélisation logique et à la mise en oeuvre d'un agent rationnel dialoguant — *D. Kayser* — *D. Sadek*

Jury : Luis Farinas del Cerro (Pdt), Philippe Besnard (rapp), P.R. Cohen (rapp), D. Bachut, C. Fouqueré, B. Levrat

DE AMO - BOJARCZUK) Sandra Aparecida — 01/31/95 — THF

Contraintes dynamiques et schémas transactionnels — *N. Bidoit*

Jury : Jacqueline Vauzeilles (Pdt), Serge Abiteboul (rapp), Laurence Cholvy (rapp), Victor Vianu (rapp), Richard Lassaigne

ETTAOUIL Mohamed — 02/25/94 — TH

Satisfaction de contraintes non linéaires en variables 0-1 et outils de la programmation quadratique — *G. Plateau*

Jury : Didier Fayard (rapp), Arnaud Fréville (rapp), Pierre Tolla (Pdt), Hachemi Bennaceur

GRAÏNE Slimane — 04/06/94 — TH

Inférence grammaticale régulière par les réseaux de neurones — *G. Plateau* — *[A. Saoudi]*

Jury : Satoru Miyano (rapp), Maurice Nivat (rapp), Daniel Kayser, Stefano Crespi-Reghizzi (rapp), Younès Bennani

MADANI Nacéra — 05/25/94 — TH

Étude de l'héritage des propriétés dans les réseaux sémantiques. Notion de réseau d'héritage légal — *D. Kayser*

Jury : Guy Chaty (Pdt), Serge Garlatti (rapp), Éric Grégoire (rapp), François Rechenmann (rapp), François Lévy

MOKHTARI Aïcha — MCF USTHB Alger — 06/14/94 — TH

Apport des logiques non-classiques à la représentation de la causalité — *D. Kayser*

Jury : Erik Sandewall (Pdt), Daniel Coulon (rapp), Jean Sallantin (rapp), Daniel Andler, Philippe Dague, François Lévy

PERRIN - NAZARENKO Adeline — MCF Paris 13 (LIPN) — 01/31/94 — THF

Compréhension du langage naturel : le problème de la causalité — *D. Kayser* — *J. Fargues*

Jury : Mario Borillo (Pdt), Jacques Jayez, Gérard Ligozat (rapp), Gérard Sabah, Gabriel Bès (rapp)

Formation par la recherche

SENOUSSI Houcine — 07/05/94 — TH

Algorithmes parallèles pour le traitement d'images — *G. Plateau* — *[A. Saoudi]*

Jury : Maurice Nivat (Pdt), Maxime Crochemore (rapp), Steven Tanimoto (rapp), Jean-Marc Chassery, Christian Lavault

THERET Philippe François — 02/08/94 — THF

De l'efficacité des interpréteurs de systèmes de règles de production dans les systèmes à base de connaissances — *S. Schwer*

Jury : Jean-Paul Haton (Pdt et rapp), Norbert Cot (rapp), Pascal Grojean, Gérard Plateau

BENSIMON DARCEL Nelly — MCF Paris8 — 02/25/93 — TH

ObAdE, des objets pour une modélisation cognitive — *D. Kayser*

Jury : Jean-Paul Haton (Pdt), Danièle Dubois (rapp), Danielle Hérin-Aime (rapp), Gérard Ligozat (rapp), Jacques Ferber

CHAMPESME Marc — MCF Paris 13 (LIPN) — 02/12/93 — TH

Apprentissage par Détection de Similarités utilisant le formalisme des Graphes Conceptuels — *D. Kayser* — *H. Soldano*

Jury : Ahmed Saoudi (Pdt), Jean Sallantin (rapp), Michel Chein (rapp), Guy Chaty

HERNERT Patrice — Ens univ Montréal — 06/25/93 — TH

Un système d'acquisition de définitions basé sur le modèle des graphes conceptuels — *D. Kayser* — *J. Fargues*

Jury : Michel Chein (Pdt), Violaine Prince (rapp), Bernard Levrat, Daniel Coulon (rapp)

ISLI Amar — Post-doc — 11/08/93 — TH

Automates alternants et logiques temporelles, satisfaction de contraintes — *G. Plateau* — *[A. Saoudi]*

Jury : Paul Schupp (rapp), Maurice Nivat (rapp et Pdt), Daniel Kayser, Malik Ghallab, Marie-Catherine Vilarem

JOUBEL Catherine — 11/30/93 — H

Comment le temps change-t-il les hypothèses ? — *D. Kayser*

Jury : Hélène Bestougeff (Pdt), Olivier Raiman (rapp), Bernard Robinet (rapp), Claude Hans

THOMAS Isabelle — 09/28/93 — TH

Modèle de l'utilisateur, modèle de la tâche, et modèles d'explications pour expliquer une résolution de problème. Application en conception d'équipements électronique — *D. Kayser* — *B. Levrat*

Jury : Jacques Ferber (Pdt), Marie-Christine Rousset (rapp), Rose Dieng (rapp), Denis Carcagno

TOLLU Christophe — MCF Paris 13 (LIPN) — 03/01/93 — THF

Planification en Logique Linéaire et Langages de Requêtes Relationnelles avec Compteurs — *J. Vauzeilles*

Jury : Etienne Grandjean (rapp), Jean-Yves Girard, Michel de Rougemont (rapp), Serge Abiteboul, Jean-Pierre Jouannaud, Daniel Kayser

ANDRY François — 05/27/92 — TH

Mise en oeuvre de prédictions linguistiques dans un système de dialogue oral homme-machine coopératif — *D. Kayser*

Jury : Jean-Paul Haton (Pdt), Jean-Sylvain Liénard (rapp), Jacques Siroux (rapp), Hélène Bestougeff, Francis Charpentier, C. Fouqueré

BONTÉ Eric — MCF IUFM — 03/19/92 — TH

Calcul des extensions dans les théories de défaut en réseau et application au raisonnement à profondeur variable — *D. Kayser*

Jury : Michel Chein (Pdt rapp), Christine Froidevaux (rapp), Jacqueline Castaing, Maxime Crochemore, Jean-Paul Haton

BOUHADANA Henri — 06/11/92 — TH

Conception d'un ensemble d'outils pour la représentation, l'analyse et l'exploitation de connaissances en sûreté de fonctionnement. L'atelier FIGARO — *D. Kayser*

Formation par la recherche

Jury : Ahmed Saoudi (rapp), Michel Gondran (rapp), Jean-Louis Bon, Claudie Ancelin

BREZELLEC Pierre — MCF Evry — 10/05/92 — TH

Techniques d'apprentissage par explication et par détection de similarités — *D. Kayser* — *H. Soldano*

Jury : Jean-Gabriel Ganascia (Pdt), Maxime Crochemore (rapp), Joël Quinqueton (rapp), Alain Viari

DORÉ Laurent — 10/05/92 — TH

Intégration de récit : application à la compréhension de textes médicaux — *D. Kayser* — *P. Zweigenbaum*

Jury : Jean-Francois Boisvieux (Pdt), Alain Bonnet (rapp), Gérard Sabah (rapp)

GAYRAL Française — MCF Paris 13 (LIPN) — 01/30/92 — TH

Sémantique du langage naturel et profondeur variable : une première approche — *D. Kayser*

Jury : Philippe Jorrand (Pdt), Gérard Ligozat (rapp), Bernard Victorri (rapp), Gérard Plateau, Robert Martin

GRANDEMANGE Philippe — MCF Paris 13 — 02/12/92 — TH

Représentation des connaissances et profondeur variable : une implantation — *D. Kayser*

Jury : Philippe Besnard, Francois Lévy, Malik Ghallab (Pdt), Jean-Francois Richard, Jacqueline Vauzeilles (rapp), Bernard Victorri (rapp)

MELA Augusta — MCF Montpellier — 12/08/92 — TH

Traitement automatique de la coordination — *D. Kayser* — *C. Fouqueré*

Jury : Jean-Pierre Descles (Pdt), Laurence Danlos (rapp), Paul Sabatier (rapp), Gaston Gross

SZULMAN Sylvie — MCF Paris 13 (LIPN) — 12/13/90 — TH

Enrichissement d'une base de connaissances à partir de textes en langage naturel — *D. Kayser* — *B. Biebow*

Jury : Jacques Pitrat (Pdt), Guy Chaty (rapp), Geneviève Jomier (rapp), Michel Chein

BENNACEUR Hachemi — MCF-HDR Cergy (LIPN) — 12/18/89 — TH

Le problème de satisfaction de contraintes : synthèse et méthode exacte de résolution — *G. Plateau*

Jury : Gerd Finke (Pdt rapp), Catherine Roucairol, Daniel Kayser, François Thomasset

Formation par la recherche

CHANIER Thierry — MCF Paris 5 — 06/28/89 — TH

Compréhension de textes dans un domaine technique : le systeme ACTES. Application des grammaires d'unification et de la théorie du discours — *D. Kayser* — *C. Sedogbo*

Jury : Daniel Coulon (Pdt), Gérard Sabah (rapp), Patrick Saint-Dizier (rapp)

COUPEY Pascal — MCF Paris 13 (LIPN) — 01/27/89 — TH

Etude d'un réseau sémantique avec gestion des exceptions. Interprétation logique et implantation informatique — *D. Kayser*

Jury : Jean-François Perrot (Pdt), Guy Chaty (rapp), Christine Froidevaux (rapp), Marie-Odile Cordier, Michel Gilloux

LÉVY François — MCF Paris 13 (LIPN) — 02/08/89 — TH

Contribution à la realisation d'un raisonneur à profondeur variable : le système de maintenance — *D. Kayser*

Jury : Jean-Pierre Jouannaud (Pdt), Jacqueline Castaing (rapp), Patrice Enjalbert (rapp), Pierre Siegel (rapp), Claude Puech

MIHOUBI Douadi — 07/07/89 — TH

Modes de reconnaissance et équité dans les omega-automates à pic — *M. Crochemore*

Jury : Paul Schupp (rapp et Pdt), Ahmed Saoudi (rapp), Irène Guessarian, Daniel Kayser

FOUQUERÉ Christophe — MCF Paris 13 (LIPN) — 01/26/88 — TH

Systèmes d'analyse tolérante de textes en langage naturel — *D. Kayser*

Jury : Guy Perennou (Pdt), Daniel Coulon (rapp), Gérard Sabah (rapp), Christian Fluhr

Habilitations à Diriger des Recherches soutenues

Nom Prénom — Soutenance

Titre

Jury :

Viennet Emmanuel — 12/18/07

Contributions aux méthodes d'apprentissage numérique pour la fouille de données structurées

Jury : Stéphane Canu (rapp), Pascale Kuntz (rapp), Gilbert Ritschard (rapp), Younès Bennani, Françoise-Fogelman-Soulié, Christophe Fouqueré, Gilbert Saporta

Butelle Franck — 12/17/07

Jury :

Toumazet Frédéric — 12/13/07

Combinatorics of group characters

Jury : Gérard Duchamp, Christophe Fouqueré, Florent Hivert (rapp), Gleb Koshevoy (rapp), Alain Lascoux (rapp), Jean-Christophe Novelli, Peter Paule (rapp), Jean-Yves Thibon (pdt)

Roupin Frédéric — 11/24/06

Jury : M. Guignard (rapp), C. Lemaréchal (rapp), M. Minoux (rapp), A. Billionnet, M.C. Costa, C. Lavault, P. Michelon, G. Plateau

Biébow Brigitte — 07/09/04

De Dasert à Terminæ ; voyage au pays de la modélisation des connaissances à partir de textes

Jury : Jean Charlet (rapp), Rose Dieng-Kuntz (rapp), Gilles Kassel (rapp), Daniel Kayser, François Lévy

Nagih Anass — 12/14/04

Contributions à la résolution des problèmes en nombres entiers : réoptimisation, problèmes de grande taille et planification en transport

Jury : Alain Billionnet (rapp), Monique Guignard-Spielberg (rapp), Alain Quilliot (rapp), D. Gauyacq, Nelson Maculan, Gérard Plateau, François Soumis, Jacqueline Vauzeilles

Perrin-Nazarenko Adeline 12/10/04

Donner accès au contenu des documents textuels. Acquisition de connaissances et analyse de corpus spécialisés

Jury : Maria Teresa Pazienza (rapp), Jean-Marie Pierrel (rapp), Bernard Victorri (rapp), Benoît Habert, Daniel Kayser, François Lévy

Faye Alain 12/17/03

Contributions à la programmation quadratique en 0-1

Jury : Monique Guignard (rapp), R. Mahjoub (rapp), Pierre Michelon (rapp), Alain Billionnet, Marie-Christine Costa, Daniel Fayard, Gérard Plateau, Jacqueline Vauzeilles

Elloumi Sourour 06/26/02

Contributions à l'Optimisation Combinatoire

Jury : Pierre Hansen (rapp), Nelson Maculan (rapp), Philippe Michelon (rapp), Alain Billionnet, Marie-Christine Costa, Martine Labbé, Christian Lavault, Michel Minoux, Gérard Plateau

Yahia Hussein 03/13/02

Modèles d'évolution spatio-temporelle

Jury : Jacqueline Vauzeilles (Pdt), Henri Maitre (rapp), Laurent Younes (rapp), Rachid Deriche (rapp), Bruno Arnaldi, Younes Benanni, Eric Simon, Isabelle Herlin

Lellahi Kazem 12/20/01

Jury : Andrée C . Ehresmann (rapp), Dominqiue Laurent (rapp), George Loizou (rapp), Christine Choppy, Geneviève Jomier, Nicolas Spyratos, Jacqueline Vauzeilles

Porcheron Marc 09/26/01

Techniques d'IA pour le diagnostic automatique de systèmes industriels

Jury : Jacques Pitrat (Pdt), Luca Console (rapp), Marie-Odile Cordier (rapp), Philippe Dague (rapp), Jean-Paul Krivine, François Lévy, Bernard Monnier, Louise Travé-Massuyès

El Fallah-Seghrouchni Amal 01/25/00

Coordination d'Agents : Modèles, Algorithmes et Protocoles

Jury : Malik Ghallab (Pdt), Jean-Pierre Briot (rapp), Brahim Chaib-Draa (rapp), Patrice Enjalbert (rapp), Philippe Dague, Serge Haddad, Daniel Kayser, Christophe Sibertin-Blanc

Formation par la recherche

Lisser Abdel-Ilah 12/21/00

Jury : Bernard Lemaire (Pdt), Nelson Maculan (rapp), Michel Minoux (rapp), Pierre Tolla (rapp), Gérard Plateau, James Roberts, Vassilis Zissimopoulos

Balbiani Philippe 05/11/99

Représentation logique et traitement algorithmique de l'espace

Jury : Gérard Plateau (Pdt), Gérard Ligozat (rapp), John-Jules Meyer (rapp), Michael Rusinowitch (rapp), Luis Farinas del Cerro, Jacqueline Vauzeilles

Bennani Younès 03/16/98

Contributions au contrôle de la capacité de généralisation des systèmes d'apprentissage connexionistes

Jury : Françoise Fogelman (Pdt), Gérard Chollet (rapp), Yann Le Cun (rapp), Patrick Gallinari (rapp), Milan Milosavljevic (rapp), Daniel Kayser, Gérard Plateau, Daniel Stern

Zweigenbaum Pierre 02/06/98

Traitement automatique de la langue médicale

Jury : Pierre Dujols (rapp), Naomi Sager (rapp), Bernard Victorri (rapp), Jean-François Boisvieux, Mario Fieschi, Christophe Fouqueré, Daniel Kayser

Bennaceur Hachemi 01/08/97

Contributions à la résolution de problèmes de satisfaction de contraintes

Jury : Nelson Maculan (rapp), Pierre Siegel (rapp), Marie-Catherine Vilarem (rapp), Philippe Dague, Mehmet Dincbas, Gérard Plateau

Coupey Pascal 05/02/97

Propriété par défaut : du point de vue inférentiel au point de vue définitionnel

Jury : Michel Chein (rapp), Amedeo Napoli (rapp), Daniele Nardi (rapp), Daniel Kayser, Maurice Nivat, Jacqueline Vauzeilles

Michelon Philippe 06/10/97

Contributions à l'optimisation en variables binaires

Jury :

Fouqueré Christophe 01/24/95

Syntaxe, taxonomies et concepts : cherchez l'intrus

Jury : Jean-Pierre Jouannaud (Pdt), Patrice Enjalbert (rapp), Jose Meseguer (rapp), Jean-François Perrot (rapp), Philippe Besnard, Daniel Kayser, Jacqueline Vauzeilles

Costa Marie-Christine 11/29/94

Contribution à la résolution de problèmes d'optimisation en variables bivalentes. Application au placement dans les systèmes distribués et à la localisation dans les réseaux de télécommunications

Jury : Gérard Florin (Pdt), Pierre Hansen (rapp), Michel Minoux (rapp), Gérard Plateau (rapp), Philippe Chrétienne, Claude Kaiser, Jean-Claude Konig, Christian Lavault

Trigano Philippe 01/19/94

Des Bases de Données et de Connaissances aux Interfaces en Langue Naturelle : un Enrichissement Réciproque

Jury : Daniel Coulon (Pdt), Jean-Paul Haton (rapp), Daniel Kayser (rapp), Gérard Sabah (rapp), Jean-Paul Barthès, Marie-Odile Cordier, Bernard Dubuisson, Claudette Sayettat

El Bèze Marc 01/18/93

Jury : Gaston Gross (Pdt), Joseph Mariani (rapp), Bernard Mérialdo (rapp), Guy Chaty, Philippe Dague

Levrat Bernard 02/26/93

(thèse d'état) Le probleme du sens dans les systemes de traitement automatique du langage naturel : une approche alternative au travers de la paraphrase

Jury : Igor Mel'cuk (Pdt, rapp), Michel Chein (rapp), Jean-Marie Pierrel (rapp), Guy Chaty, Daniel Coulon, Catherine Fuchs, Daniel Kayser, Gerard Sabah

Lévy François 02/10/93

Études sur le raisonnement

Jury : Jean-Pierre Jouannaud (Pdt), Philippe Besnard (rapp), Mike Clarke (rapp), Henri Prade (rapp), Daniel Kayser, Jean-Francois Le Ny, Jacqueline Vauzeilles

Fréville Arnaud 01/10/91

Contribution à l'optimisation en nombres entiers

Jury : Gerd Finke (Pdt, rapp), F.Glover (rapp), Monique Guignard-Spielberg (rapp), J.C. Fiorot, J. Fonlupt, Daniel Kayser, Gérard Plateau, Catherine Roucairol

Ribeiro Celso 09/10/90

Contribution à l'optimisation combinatoire et ses applications

Jury : Gerd Finke (Pdt, rapp), D. De Werra (rapp), Gérard Plateau (rapp), Jean-Pierre Barthélémy, Pierre Hansen, Daniel Kayser, Michel Minoux

Chapitre 2
Organisation de conférences

Le LIPN a soutenu ou organisé les conférences et manifestations suivantes :

- *Implicit Computational Complexity and Logic*, septembre 2004, université Paris 13 (ACI GEOCAL), 30 participants, dont 8 étrangers.
- *Workshop on Implicit Computational Complexity*, 13-17/2/2006, (GEOCAL, *Geometry of Computation*, 2006), Marseille. 20 exposés et 50 participants (de 7 pays).
- *2es Journées thématiques "Apprentissage Artificiel & Fouille de Données"*, 27-28/4/2006, http://www-lipn.univ-paris13.fr/A3/AAFD06/
- *Journées ALEA*, 2005, (CIRM, SMF-CNRS, Luminy)
- *Marches aléatoires discrètes : théorie et applications en combinatoire, biologie algorithmique, informatique théorique, probabilités, physique statistique* (DRW-2003), Institut Henri Poincaré, 2003

Partie IV
Publications

Les références comportant « * » sont acceptées et à paraître.

Sommaire

Articles dans des revues avec comité de lecture (ACL) 207
Conférences invitées (INV) 214
Communications avec actes (ACT) 215
Direction d'ouvrages (DO) 239
Ouvrages scientifiques (OS) 242
Autres publications (AP) 243

Articles dans des revues avec comité de lecture (ACL)

Articles dans des revues internationales avec comité de lecture sélectif

2008 [RI-86*] Pierre Boudes. Non uniform (hyper/multi)coherence spaces. *Mathematical Structures in Computer Science*, 2008. 39 pages Accepté pour publication à MSCS F.3.2.

[RI-85*] L. Hillah, F. Kordon et L. Petrucci. Application des méthodes formelles à la robotique modulaire — méthodes formelles pour l'analyse des robots autonomes et modulaires. 2008.

[RI-84] Christian Lavault et Mario Valencia-Pabon. A distributed approximation algorithm for the minimum degree minimum weight spanning tree. *Journal of Parallel and Distributed Computing*, 68(2) :200–208, 2008.

2007 [RI-83] Laurent Alfandari. Improved approximation of the soft-capacitated facility location problem. *RAIRO Operations Research*, 41 :83–93, 2007.

[RI-82*] Erick Alphonse et Aomar Osmani. On the connection between the phase transition of the covering test and the learning success rate in ilp. *Machine Learning Journal*, 2007.

[RI-81] V. Atassi, P. Baillot et K. Terui. Verification of ptime reducibility for system f terms : Type inference in dual light affine logic. *Logical Methods in Computer Science*, 2007. à paraître. 32 pages.

[RI-80] P. Blasiak, K. A. Penson, A. Horzela, A. I Solomon et G. H. E. Duchamp. Combinatorics and boson normal ordering : A gentle introduction. *Am. J. Phys.*, 75 :639–646, 2007.

[RI-79] Sylvie Després et Sylvie Szulman. Merging of Legal Micro-Ontologies from European Directives. *Artificial Intelligence and Law*, 15(2) :187–200, Feb. 2007.

[RI-78] G. Duchamp et J.-G. Luque. Lazard's elimination (in traces) is finite-state recognizable. *International Journal of Algebra and Computation*, 17(1) :53–60, février 2007.

[RI-77*] Christophe Fouqueré et Virgile Mogbil. Rewritings for polarized multiplicative and exponential proof structures. *Electronic Notes in Theor. Comp. Sci.*, pages 54–65, 2007.

[RI-76] Danielo Gomes, Nazim Agoulmine, Younès Bennani et Neuman De Souza. Predictive connectionist approach for vod bandwidth management. *International Journal for Computer Communications*, 30(10) :2236–2247, 2007.

[RI-75*] S. Hamdoun, F. Boufarès et M. Badri. Construction et maintenance des entrepôts de données hétérogènes. *eTI journal : Electronic Journal of Information Technologies*, 2007. À paraître.

[RI-74*] Hakima Kadri-Dahmani, Cyrille Bertelle, Gérard Duchamp et Aomar Osmani. Consistent updating of geographical database as emergent property over influence system. *International Journal of Modelling Identification and Control, special Issue Modelling Complex Systems*, 2007.

[RI-73] Max Kanovich et Jacqueline Vauzeilles. Strong planning under uncertainty in domains with numerous but identical elements (a generic approach). *Theoretical Computer Science*, 379(1–2) :84–119, 2007.

[RI-72] C. Lakos et L. Petrucci. Modular state space exploration for timed Petri nets. *Journal of Software Tools for Technology Transfer*, 9(3–4) :393–411, juin 2007.

[RI-71] Christian Lavault, Jean-Franois Marckert et Vlady Ravelomanana. Quasi-optimal energy-efficient leader election algorithms in radio networks. *Information and Computation*, 205(5) :679–693, 2007.

[RI-70] Christian Lavault, Jean-Franois Marckert et Vlady Ravelomanana. Quasi-optimal leader election algorithms in radio networks. *Information and Computation*, 205(5) :679–693, 2007.

[RI-69*] Jérôme Monnot et Sophie Toulouse. Approximation results for the weighted p_4 partition problem. *Journal of Discrete Algorithms*, 2007.

[RI-68*] Jérôme Monnot et Sophie Toulouse. The complexity of the p_k partition problem and related problems in bipartite graphs. *Operations Research Letters*, 2007.

[RI-67*] Fabien Picarougne, Hanane Azzag, Gilles Venturini et Christiane Guinot. A new approach of data clustering using a flock of agents. *Evolutionary Computation*, 2007.

[RI-66*] Thierry Poibeau. Semantic annotation : Mapping text to ontologies. *International Journal of Metadata Semantics and Ontologies*, 2(2), 2007.

[RI-65] Vlady Ravelomanana. Another proof of Wright's inequalities. *Information Processing Letters*, 104(1) :36–39, 2007.

[RI-64] Vlady Ravelomanana. Optimal gossiping and initialization protocols for random wireless multihop network. *IEEE Transactions on Parallel and Distributed Systems*, 18(1) :246–257, 2007.

[RI-63] Arthur Tenenhaus, Alain Giron, Emmanuel Viennet, Michel Béra, Gilbert Saporta et Bernard Fertil. Kernel logistic PLS : a tool for supervised nonlinear dimensionality reduction and binary classification. *Computational Statistics & Data Analysis*, 51(9) :4083–4100, mai 2007.

2006 [RI-62] Rafael Castro De Andrade, Abdel Lisser, Nelson Maculan et Gérard Plateau. Enhancing a branch-and-bound algorithm for two-stages stochastic integer problems. *Management Science*, 52(9) :1450–1455, 2006.

[RI-61] Khalid Benabdeslem et Younès Bennani. Dendogram-based svm for multi-class classification. *International Journal of Computing and Information Technology*, 14(4) :288–291, 2006.

[RI-60] D. Berrabah, F. Boufarès et C.F. Ducateau. A new tool to check the coherence of constraints defined on UML class diagrams. *International Journal of Computer Science and Network Security*, 6 :162–171, 2006.

[RI-59] Christophe Cérin, Jean-Luc Gaudiot et Michel Koskas. A multithreaded SQL service. *Parallel Processing Letters*, 16(2) :245–259, 2006.

[RI-58] C. Choppy et G. Reggio. A formally grounded software specification method. *Journal of Logic and Algebraic Programming*, 67(1-2) :52–86, 2006.

[RI-57] U. Dal Lago et P. Baillot. On Light Logics, Uniform Encodings and Polynomial Time. *Mathematical Structures in Computer Science*, 16(4) :713–733, 2006. (une version préliminaire avait été présentée au *Workshop on Logics for Resources, Processes, and Programs* LRPP'04).

[RI-56] G.H.E. Duchamp, J-G. Luque, K-A. Penson et C. Tollu. Free quasi-symmetric functions, product actions and quantum field theory of partitions. *Séminaire Lotharingien de Combinatoire*, 54A :12 pages, 2006.

[RI-55] Mourad Hakem et Franck Butelle. Critical path scheduling parallel programs on an unbounded number of processors. *Int. J. of Foundations Computer Science*, 17(2) :287–301, 2006.

[RI-54*] L. Jaff, G.H.E. Duchamp et C. Bertelle. Shift operators and complex systems. *Int. Journal of Modeling, Identification and Control*, 2006.

[RI-53] R. C. King, C. Tollu et F. Toumazet. The hive model and the factorisation of kostka coefficients. *Séminaire Lotharingien de Combinatoire*, 54A :22 pages, 2006.

[RI-52] R. C. King, C. Tollu et F. Toumazet. The hive model and the polynomial nature of stretched Littlewood-Richardson coefficients. *Séminaire Lotharingien de Combinatoire*, 54A :19 pages, 2006.

[RI-51] Christian Lavault et Guy Louchard. Asymptotic analysis of a leader election algorithm. *Theoretical Computer Science*, 359(2) :239–254, 2006.

[RI-50] Christian Lavault, Gabriel Paillard et Felipe Franca. A distributed prime sieving algorithm based on smer. *Scalable Computing Practice and Experience, special issue of ISPDC'2005 (selected papers)*, 3(8) :61–69, 2006.

[RI-49*] J.-G. Luque, J.-Y. Thibon et F. Toumazet. Unitary invariants of qubit systems. *Mathematical Structures in Computer Science*, page 19 pages, 2006. À paraître.

[RI-48] Anass Nagih et François Soumis. Nodal aggregation of resource constraints in a shortest path problem. *European Journal of Operational Research*, 172 :500–514, 2006.

[RI-47] Vlady Ravelomanana. The average size of giant components between the double-jump. *Algorithmica*, 46(3) :529–555, 2006.

[RI-46] Babacar Thiongane, Anass Nagih et Gérard Plateau. Lagrangean heuristics combined with reoptimization for the 0-1 biknapsack problem. *Discrete Applied Mathematics*, 154 :2200–2211, 2006.

2005 [RI-45] Foto Afrati, Euripidis Bampis, Lucian Finta et Ioannis Milis. Scheduling trees with large communication delays on two identical processors. *Journal of Scheduling*, 8(2) :179–190, 2005.

[RI-44] Rafael Castro De Andrade, Abdel Lisser, Nelson Maculan et Gérard Plateau. B & b frameworks for the capacity expansion of high speed telecommunication network under uncertainty. *Annals of Operations Research*, 140(1) :49–65, 2005.

[RI-43] Cyril Banderier et Sylviane Schwer. Why Delannoy numbers? *J. Statist. Plann. Inference*, 135(1) :40–54, 2005.

[RI-42] Cyril Banderier et Sylviane R. Schwer. Why Delannoy numbers? *Journal of Statistical Planning and Inference*, 135(1) :40–54, 2005.

[RI-41] J. Billington, G. E. Gallasch et L. Petrucci. FAST verification of the class of stop-and-wait protocols modelled by coloured Petri nets. *Nordic Journal of Computing*, 12(3) :275–307, automne 2005.

[RI-40] P. Blasiak, A. Horzela, K. A. Penson, G. H. E. Duchamp et A.I. Solomon. Normal ordering via substitutions and Sheffer-type polynomials. *Phys. Lett. A*, 338 :108–, 2005.

[RI-39] P. Blasiak, K.A. Penson, A.I. Solomon, A. Horzela et G.H.E. Duchamp. Some useful combinatorial formulas for bosonic operators. *J. Math. Phys.*, 46 :052110–, 2005.

[RI-38] F. Boufarès, D. Berrabah, C.F. Ducateau et F. Gargouri. Les conflits entre les contraintes dans les schémas conceptuels de bases de données UML-EER. *Information Sciences for Decision Making Journal (ISDM Journal)*, 19 :234–244, 2005.

[RI-37] F. Boufarès et S. Hamdoun. Integration techniques to build a data warehouse using heterogeneous data sources. *Journal of Computer Sciences*, pages 48–55, 2005.

[RI-36] Christophe Cérin, Olivier Cozette, Gil Utard, Hazem Fkaier et Mohamed Jemni. Parallel out-of-core sorting and fast accesses to disks. *International Journal on High Performance Computing and Networking*, 3 :188–202, 2005.

[RI-35] C. Choppy, D. Hatebur et M. Heisel. Architectural patterns for problem frames. *IEE Proceedings - Software, Special Issue on Relating Software Requirements and Architectures*, 152(4) :198–208, 2005.

[RI-34] C. Choppy et G. Reggio. A UML-Based Approach for Problem Frame Oriented Software Development. *Journal of Information and Software Technology*, 47 :929–954, 2005.

[RI-33] M-C. Costa, L. Létocart et F. Roupin. Multicut and integral multiflow : a survey. *European Journal of Operational Research*, 162(1) :55–69, 2005.

[RI-32] Tadeusz Czachorski et Ferhan Pekergin. A novel diffusion process with jumps to study an electronic-optical edge router. *Archives of Theoretical and App. Informatics*, 17(2) :76–88, 2005.

[RI-31] David Delahaye et Micaela Mayero. Dealing with Algebraic Expressions over a Field in Coq using Maple. 39(5) :569–592, 2005.

[RI-30] G. Duchamp, Hatem Hadj Kacem et Eric Laugerotte. Algebraic elimination of ϵ-transitions. *Discrete Mathematics and Theoretical Computer Science*, 7(1) :51–70, 2005.

[RI-29] Alejandro Guerra-Hernández, Amal El Fallah-Seghrouchni et Henry Soldano. On learning intentionality. *Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial*, 9(25) :9–18, 2005. Première version publiée dans Proceedings of IberAgents 04 Workshop.

[RI-28] Daniel Kayser et Farid Nouioua. About norms and causes. *International Journal on Artificial Intelligence Tools*, 14 :7–23, 2005.

[RI-27] Romain Mellier, Jean-Frédéric Myoupo et Vlady Ravelomanana. A non-token-based-distributed mutual exclusion algorithm for single-hop mobile ad hoc networks. *IFIP International Federation for Information Processing*, 162 :287–298, 2005.

[RI-26] Adeline Nazarenko et Touria Aït El Mekki. Building back-of-the-book indexes. *Terminology*, 11(1) :193–218, 2005.

[RI-25] Babacar Thiongane, Anass Nagih et Gérard Plateau. Adapted step size in a 0-1 biknapsack lagrangean dual solving algorithm. *Annals of Operations Research*, 139 :353–373, 2005.

[RI-24] Farida Zehraoui et Younès Bennani. New self-organising maps for multivariate sequences processing. *International Journal of Computational Intelligence and Applications*, 5(4) :439–456, 2005.

2004 [RI-23] Rafael Castro De Andrade, Abdel Lisser, Nelson Maculan et Gérard Plateau. Telecommunication network capacity design for uncertain demand. *Computational Optimization and Applications*, 29(2) :127–146, 2004.

[RI-22] P. Baillot. Stratified coherence spaces : a denotational semantics for Light Linear Logic. *Theoretical Computer Science*, 318(1-2) :29–55, 2004.

[RI-21] P. Baillot. Type inference for Light Affine Logic via constraints on words. *Theoretical Computer Science*, 328(3) :289–323, 2004.

[RI-20] Hachemi Bennaceur. A comparison between CSP and SAT techniques. *Constraints : an International Journal*, 9(2) :123–138, 2004.

[RI-19] Lélia Blin et Franck Butelle. The first approximated distributed algorithm for the minimum degree spanning tree problem on general graphs. *Int. J. of Foundations Computer Science*, 15 :507–516, 2004.

[RI-18] F. Boufarès et H. Bennaceur. Consistency problems in er-schemas for database systems. *Information Sciences Journal*, 163(4) :263–274, 2004.

[RI-17] F. Boufarès et H. Bennaceur. Consistency problems in ER-schemas for database systems. *Information Sciences Journal*, 163(4) :263–274, 2004.

[RI-16] Marc Bui, Franck Butelle et Christian Lavault. A distributed algorithm for the minimum diameter spanning tree problem. *Journal of Parallel and Distributed Computing*, 64(5) :571–577, 2004.

[RI-15] Marie-Odile Cordier, Philippe Dague, François Lévy, Jacky Montmain, Marcel Staroswiecki et Louise Travé-Massuyès. Conflicts versus analytical redundancy relations : a comparative analysis of the model based diagnosis approach from the artificial intelligence and automatic control perspective. *IEEE Transactions on Systems, Man and Cybernetics*, 34 :2163–2177, 2004.

[RI-14] G. Duchamp et J.-M. Champarnaud. Derivatives of rational expressions and related theorems. *Theoretical Computer Science*, 313(1) :31–44, 2004.

[RI-13] Nahla El-Zant et Henry Soldano. Finding repeated flexible relational words in sequences. *Journal of Systemics, Cybernetics and Informatics*, 2(4), 2004.

[RI-12] R. C. King, F. Toumazet et B. G. Wybourne. The square of the Vandermonde determinant and its q-generalisation. *J. Phys. A : Math. Gen.*, 37 :735–767, 2004.

[RI-11] Francine Krief. Self-aware management of IP networks with QoS guarantees. *International Journal on Network Management*, 2004.

[RI-10] K.A. Penson, P. Blasiak, G. Duchamp, A. Horzela et A.I. Solomon. Hierarchical Dobiński-type relations via substitution and the moment problem. *J. Phys. A : Math. Gen.*, 37 :3457–, 2004.

[RI-9] Vlady Ravelomanana. Extremal properties of three dimensional sensor networks with applications. *IEEE Transactions on Mobile Computing*, 3 :17–28, 2004.

[RI-8] Vlady Ravelomanana et Loys Thimonier. Forbidden subgraphs in connected graphs. *Theoretical Computer Science*, 314(1–2) :121–171, 2004.

2003 [RI-7] Jean-Michel Autebert et Sylviane R. Schwer. On generalized Delannoy paths. *SIAM Journal on Discrete Mathematics*, 16(2) :208–223, 2003.

[RI-6] Anne-Elisabeth Baert, Vlady Ravelomanana et Loys Thimonier. On the growth of components with non fixed excesses. *Discrete Applied Mathematics*, 130(3) :487–493, 2003.

[RI-5] Bruno Chassagne et Vlady Ravelomanana. Simple technique to widen the passband of a free-space optics multiplexer and demultiplexer. *Journal of The Optical Society of America*, 20(11) :2270–2274, 2003.

[RI-4] M-C. Costa, L. Létocart et F. Roupin. A greedy algorithm for multicut and integral multiflow in rooted trees. *Operations Research Letters*, 31 :21–27, 2003.

[RI-3] Daniel Kayser. Abstraction and natural language semantics. *Philosophical Transactions of the Royal Society of London*, 358 :1261–1268, 2003.

[RI-2] Jean-Frédéric Myoupo, Vlady Ravelomanana et Loys Thimonier. Average case analysis based-protocols to initialize packet radio networks. *Wireless Communications and Mobile Computing*, 3(4) :539–548, 2003.

[RI-1] Louise Travé-Massuyès, L. Ironi et Philippe Dague. Mathematical foundations of qualitative reasoning. *AI Magazine*, 24(4) :91–106, 2003.

Articles dans des revues nationales ou internationales

2008 [RE-40*] Sylvie Després et Hélène Richy. Métadonnées, ontologies et documents numériques. *Revue des Techniques de l'Ingénieur*, 2008.

2007 [RE-39*] Sujeevan Aseervatham et Emmanuel Viennet. Méthodes à noyaux appliquées aux textes structurés. *Revue des Nouvelles Technologies de l'Information* [ED-7], pages 185–205. 230 pages, publication en cours.

[RE-38*] Sophie Aubin, Julien Derivière, Thierry Hamon, Adeline Nazarenko, Thierry Poibeau et Davy Weissenbacher. Une infrastructure pour l'annotation linguistique de documents issus du web : le projet ALVIS. *Revue des Nouvelles Technologies de l'Information*, 2007.

[RE-37*] Hanane Azzag, David Da Costa, Christiane Guinot et Gilles Venturini. Un aperçu de la fouille visuelle de données. *Revue des Nouvelles technologies de l'Information (RNTI)*, 2007. Journées AAFD-2006 (Apprentissage Artificiel et Fouille de données).

[RE-36*] Delphine Battistelli, Jean-Luc Minel et Sylviane R. Schwer. Représentation des expressions calendaires dans les textes : vers une application à la lecture assistée de biographies. *revue TAL (traitement automatique des langues)*, 47-3, 2007. À paraître.

[RE-35] Khalid Benabdeslem et Younès Bennani. Approche connexionniste pour l'analyse des données issues d'usage d'internet : Classification et visualisation. *Revue des Nouvelles Technologies de l'Information*, 2007.

[RE-34*] Younès Bennani, Sébastien Guérif et Emmanuel Viennet. Réduction des dimensions de l'espace de description. *Revue des Nouvelles Technologies de l'Information*, 2007.

[RE-33] Sylviane R. Schwer. Traitement de la temporalité des discours : une analysis situs. *Cahiers Chronos*, 18 :7–21, 2007.

[RE-32*] Nicolas Turenne et Sylviane R. Schwer. Temporal representation of gene interaction networks from text databases - drosophila megalogaster and bacillus subtilis cases. *International Journal of data Mining and Bioinformatics (IJDMB)*, 2007. À paraître.

[RE-31*] Emmanuel Viennet et Françoise Fogelman Soulié. Le traitement des refusés dans le risque crédit. *Revue des Nouvelles Technologies de l'Information*, 2007.

[RE-30] Davy Weissenbacher. Les réseaux bayésiens : un formalisme adapté au traitement automatique des langues ? *Intelligence Artificielle*, 21(3) :371–389, Juin 2007.

2006 [RE-29] Jean-Michel Autebert et Sylviane R. Schwer. Henri Auguste Delannoy, une biographie. *Revue de Mathématiques & Sciences Humaines*, 174 :25–67, 2006.

[RE-28] Jérôme Azé, Mathieu Roche, Erik Alphonse, Ahmed Amrani, Thomas Heitz et Amar-Djallil Mezaour. Défi Fouille de Textes : reconnaissance automatique des auteurs de discours - Campagne DEFT'05 (TALN'05). *Revue des Nouvelles technologies de l'Information (RNTI)*, 2006.

[RE-27] Rushed Kanawati. Refléxions sur l'apport de l'exploration des trace d'usage pour améliorer le tri de résultats des moteurs de recherche sur le web. pages 150–154, Janvier 2006.

[RE-26] H. Karoui, R. Kanawati et L. Petrucci. An intelligent peer-to-peer multi-agent system for collaborative management of bibliographic databases. *SGAI's expert update magazine*, 8(3) :22–31, printemps 2006.

[RE-25] F. Kordon et L. Petrucci. Toward formal methods oecumenism ? *IEEE Distributed Systems Online*, 7(7), juillet 2006.

[RE-24] Sylviane R. Schwer. Représentation mathématique du temps : après Reichenbach. *Tranel (Travaux neuchâtelois de linguistique)*, 45 :167–186, 2006.

[RE-23*] Sylviane R. Schwer. Traitement de la temporalité des discours : une analysis situs. *Revue Chronos*, 2006. À paraître.

[RE-22*] Davy Weissenbacher. Les réseaux bayésiens : un formalisme adapté au traitement automatique des langues ? *Revue d'Intelligence Artificielle, Numéro spécial : "Modèles Graphiques Probabilistes"*, 2006. À paraître.

[RE-21] Haifa Zargayouna, Touria Ait El Mekki, Laurent Audibert et Adeline Nazarenko. IndDoc : An aid for the back-of-the-book indexer. *The Indexer*, 25(2) :122–125, 2006.

2005 [RE-20] Sylvie Després et Sylvie Szulman. Construction d'une ontologie formelle à partir d'un texte de droit communautaire. *International journal of theoretical and applied issues in specialized communication*, 11(2) :261–281, 2005.

[RE-19] L. Hillah, F. Kordon, L. Petrucci et N. Trèves. Building an API for ISO/IEC 15909, based on model engineering techniques. *Petri Net Newsletter*, 69, octobre 2005.

[RE-18] L. Petrucci. Cover picture story : Experiments with modular state spaces. *Petri Net Newsletter*, 68 :Cover page and 5–10, avril 2005.

[RE-17] Catherine Recanati. Raisonner avec des diagrammes : perspectives cognitives et computationnelles. *Intellectica*, 40 :9–42, 2005.

[RE-16] Véronique Ventos et Henry Soldano. Treillis de Galois alpha. *Revue d'Intelligence Artificielle RSTI série RIA*, 19(4-5) :799–227, 2005.

2004 [RE-15] Touria Aït El Mekki et Adeline Nazarenko. L'index de fin de livre, une forme de résumé indicatif ? *Traitement Automatique des Langues*, 45(1) :121–150, 2004.

[RE-14] Philippe Dague et Louise Travé-Massuyès. Raisonnement causal en physique qualitative. *Numéro spécial Raisonnement Causal, Intellectica*, 1(38), novembre 2004.

[RE-13] Natalia Grabar et Thierry Hamon. Les relations dans les terminologies structurées : de la théorie à la pratique. *Revue d'Intelligence Artificielle*, 18/1 :57–85, 2004.

[RE-12] Gaston Gross et Adeline Nazarenko. Quand la langue cause : paramètres d'une étude linguistique de la causalité. *Intellectica*, 38(1) :15–41, 2004.

[RE-11] Daniel Kayser et François Lévy. Modélisations symboliques du raisonnement causal. *Intellectica*, 38 :291–323, 2004.

[RE-10] François Lévy. Statistiques et contrefactuels, quelques explications. *Intellectica*, 38 :187–191, 2004. Introduction à "Expérimentation, inférence statistique et analyse causale" de B. Lecoutre dans le même numéro.

[RE-9] Adeline Nazarenko. Primitives logiques et réseaux sémantiques : quel bilan pour la terminologie ? *Revue d'Intelligence Artificielle*, 18(1) :41–56, 2004.

[RE-8] Thierry Poibeau. Annotation d'informations textuelles : le cas du Web sémantique. *Revue d'Intelligence Artificielle*, 18(1) :139–157, 2004.

[RE-7] Olivier Sigaud et Pierre Gérard. Apprentissage par renforcement indirect dans les systèmes de classeurs. *JEDAI*, 2004.

2003 [RE-6] Christophe Fouqueré et Fabrice Issac. Corpus issus du web : constitution et analyse informationnelle. *Revue Québecoise de Linguistique*, 13(1) :111–134, 2003.

[RE-5] Daniel Kayser. Des modèles de nature différente peuvent approcher une même réalité ; réponse à Giuseppe Longo. *Intellectica*, 35 :175–177, 2003.

[RE-4] Daniel Kayser. Quelle place pour l'Intelligence Artificielle ? *lettre du département SHS du CNRS*, 67 :11–12, 2003.

[RE-3] Nelson Maculan, Gérard Plateau et Abdel Lisser. Integer linear models with a polynomial number of variables and constraints for some classical combinatorial problems. *Pesquisa Operacional*, 23(1) :161–168, 2003.

[RE-2] Chantal Morley, Bruno Defude, Franck Butelle et Daniel Lang. Les outils du commerce électronique. *Annales des télécommunications*, 58(1–2) :8–58, 2003.

[RE-1] Babacar Thiongane, Anass Nagih et Gérard Plateau. Analyse de sensibilité pour les problèmes linéaires en variables 0-1. *RAIRO-Operations Research*, 37(4) :291–309, 2003.

Conférences invitées (INV)

Communications invitées dans une conférence internationale

2008 [IN-8*] C. Choppy, L. Petrucci et G. Reggio. A modelling approach with coloured Petri nets. *Proc. 13th Int. Conf. on Reliable Software Technologies — Ada-Europe, Venice, Italy*, Lecture Notes in Computer Science. Springer, juin 2008.

[IN-7] Christian Lavault, Stephan Olariu et Vlady Ravelomanana. Random geometric graphs and wireless network. *Proc. of the 1st Advanced Research Workshop on Information Security in Wireless Networks (ARW 2006)*, volume 13 de *NATO Security through Science Series D*, pages 18–51, 2008.

2007 [IN-6] K. Klai et L. Petrucci. Modular construction of the symbolic observation graph. *Proc. 1st International Workshop on Verification and Evaluation of Computer and Communication Systems (VECOS'07), Algiers, Algeria*, eWiC, page 6. British Computer Society, mai 2007.

2006 [IN-5] Christian Lavault. Link reversal routing algorithms for manets. *Information Security in Wireless Networks - NATO ARW*, Suceava, Romania (tutoriel invité), 2006.

2005 [IN-4] L. Petrucci. Modularity and Petri nets. *Proc. 7th International Symposium on Programming and Systems (ISPS'2005), Algiers, Algeria*, pages 7–8, mai 2005.

2004 [IN-3] Christophe Fouqueré. Le web comme source d'informations : quelles méthodes? *Proc. 72ème congrès de l'ACFAS Montréal, Canada*, mai 2004.

[IN-2] Thierry Poibeau. Content annotation for the semantic web. Sergio Bolasco, éditeur, *The NEMIS 2004 Final Conference*, pages 1–15, Rome, Italie, 2004. Universita di Roma1 « La Sapienza ». Conférence invitée lors de la conférence de clôture Nemis. Reprise ensuite en volume chez Springer.

2003 [IN-1] Daniel Kayser. Une approche inférentielle de la sémantique du langage naturel. *6th International Symposium on Programming and Systems*, page 9, Alger, Algérie, 2003.

Autres communications sur invitation

2007 [AI-3] M. Elati et C. Rouveirol. Génération d'associations candidates sous contrainte de support adaptative. *BDI'2007 : atelier bases de données inductives, Plate-forme AFIA*, 2007.

[AI-2] F. Kordon et L. Petrucci. A formal approach to designing autonomous systems : from Intelligent Transport Systems to robots. *Proc. 2nd National Workshop on Control Architectures of Robots : from models to execution on distributed control architectures, Paris, France*, mai 2007.

[AI-1] L. Petrucci. ISO/IEC 15909-2, concepts of high-level nets and CPN. *Presentation at the Workshop on Petri Nets Standards (PNS'07, associated with PETRI NETS'07), Siedlce, Poland*, juin 2007.

Tutoriaux dans des conférences internationales

[TU-3] P. Baillot. From Proof-Nets to Linear Logic Type Systems for Polynomial Time Computing. *Proceedings of Typed Lambda Calculi and Applications (TLCA'07)*, volume 4583 de *LNCS*, pages 2–7. Springer, 2007. Exposé Invité.

2006 [TU-2] C. Choppy, S. Haddad, H. Klaudel, F. Kordon, L. Petrucci et Y. Thierry-Mieg. Tutorial on formal methods for distributed and cooperative systems. *Proc. 3rd Int. Coll. on Theoretical Aspects of Computing (1-day tutorial at ICTAC'06), Tunis, Tunisia*, volume 4281 de *Lecture Notes in Computer Science*. Springer, novembre 2006.

2004 [TU-1] Tadeusz Czachorski et Ferhan Pekergin. Diffusion approximation as a modelling tool in congestion control and performance evaluation. *Proceedings of First Workshop on New Trends in Modelling, Quantitative Methods and Measurements, EURONGI, Next Generation Internet*, Poland Zakopane, 2004. 44 pages, Tutoriel, version longue.

Communications avec actes (ACT)

Communications dans des conférences internationales avec comité de lecture

2008 [CI-176*] Hervé Daudé et Vlady Ravelomanana. Random 2-xorsat at the satisfiability threshold. *Proceedings of 8th Latin American Theoretical Informatics – LATIN 2008*, http ://www.latin08.org/papers.htm, 2008.

[CI-175*] Gabriel Paillard et Vlady Ravelomanana. Limit theorems and lifetime in large sensor networks. *Proceedings of The 27th Conference on Computer Communications – IEEE INFOCOM 2008*, 2008.

2007 [CI-174*] Erick Alphonse et Aomar Osmani. Phase transition and heuristic search in relational learning. *Sixth Int. Conference on Machine Learning and Applications (ICMLA 2007)*. IEEE press, 2007.

[CI-173] Sujeevan Aseervatham, Emmanuel Viennet et Younès Bennani. A semantic kernel for semi-structured documents. *ICDM 2007*, octobre 2007. International Conference on Data Mining (ICDM'07).

[CI-172] Jérôme Azé, Lucie Gentils, Claire Toffano-Nioche, Valentin Loux, Jean-François Gibrat, Philippe Bessières, Céline Rouveirol, Anne Poupon et Christine Froidevaux. Towards a semi-automatic functional annotation tool based on decision tree techniques. *International Workshop on Machine Learning in Systems Biology (MLSB 2007)*, 2007.

[CI-171] Hanane Azzag, Younès Bennani et Mustapha Lebbah. A stochastic algorithm for clustering inspired from cockroaches behavior. *V International Conference on Microelectronics and Computer Science*, 2007.

[CI-170] Hanane Azzag, Julien Lavergne, Christiane Guinot et Gilles Venturini. On building graphs of documents with artificial ants. *16th International World Wide Web Conference (WWW2007)*, pages 1299–1301, Banff, Alberta, Canada, May, 8-12th 2007.

[CI-169] P. Baillot, P. Coppola et U. Dal Lago. Light Logics and Optimal Reduction : Completeness and Complexity. *Proceedings of Symposium on Logic in Computer Science (LICS'07)*, pages 421–430. IEEE Computer Society, 2007.

[CI-168] Etienne Birmlé, Mohamed Elati, Céline Rouveirol et Christophe Ambroise. Identification of functional modules based on transcriptional regulation structure. *International Workshop on Machine Learning in Systems Biology (MLSB 2007)*, 2007.

[CI-167] Khellaf Bouandas et Aomar Osmani. Mining association rules in temporal sequences. *IEEE Symposium on Computational Intelligence and Data Mining, CIDM 2007, part of the IEEE Symposium Series on Computational Intelligence 2007*, pages 610–615, 2007.

[CI-166*] Amanda Bouffier. From Texts to Structured Documents : The Case of Health Practice Guidelines. *Proceedings of the 6th International Semantic Web Conference (ISWC 2007)*, South Korea, Nov. 2007.

[CI-165*] Amanda Bouffier et Thierry Poibeau. Automatically Restructuring Practice Guidelines using the GEM DTD. Kevin Bretonnel Cohen, Dina Demner Fushman, Carol Friedman, Lynette Hirschman et John Pestian, éditeurs, *Proceedings of Biomedical Natural Language Processing (BioNLP'07)*, Prague, République Tchèque, 2007. ACL.

[CI-164*] Amanda Bouffier et Thierry Poibeau. Restructuring documents : A discourse based approach. *Proc. Int. Conf. Recent Advances in Natural Language Processing (RANLP'07)*, Bulgarie, Oct. 2007.

[CI-163] Gauvain Bourgne, Henry Soldano et Amal ElFallah Seghrouchni. Smile : Sound multiagent incremental learning. *International Conference on Autonomous Agents and Multiagent Systems*, pages 164–171, Honolulu, Hawai, 5 2007. ACM.

[CI-162] Laure Bourgois, Alexandre Delteil et François Lévy. A Formalism for Web Services Properties. *Proceedinfgs of the International Conference on Internet and Web Applications and Services (ICIW'07)*, page 28, 05 2007.

[CI-161] Franck Butelle, Lucian Finta et Mourad Hakem. Real-time distributed scheduling of precedence graphs on arbitrary wide networks. *Proceedings of the 21st IEEE International Parallel and Distributed Processing Symposium International Parallel and Distributed Processing Symposium*, page 220 (article sur CDROM). IEEE Computer Society, 2007.

[CI-160] Guénaël Cabanes et Younès Bennani. A simultaneous two-level clustering algorithm for automatic model selection. *(ICMLA'07), International Conference on Machine Learning and Applications*, Cincinnati, Ohio, USA, décembre 2007.

[CI-159] Jean Cardinal, Vlady Ravelomanana et Mario Valencia-Pabon. Chromatic edge strength of some multigraphs. *Proceedings of The IV Latin-American Algorithms, Graphs and Optimization Symposium – LAGOS 2007*, 2007. to appear in Electronic Notes in Discrete Mathematics.

[CI-158] Pascal Coupey, Christophe Fouqueré et Jean-Vincent Loddo. Tree components programming : An application to xml. *4th International Colloquium on Theoretical Aspects of Computing (ICTAC07)*, volume 4711 de *Lecture Notes in Computer Science*, pages 139–153. Springer, septembre 2007.

[CI-157*] Sylvie Després et Valentina Ceausu. Semantic case-based reasoning framework for text categorization. *Proceedings of the 6th International Semantic Web Conference and the 2nd Asian Semantic Web Conference 2007 (ISWC 2007 + ASWC 2007)*, Corée, 11 2007.

[CI-156] Sylvie Després et Sonia Mhiri. Ontology Visualization Tool for Indexing DICOM Structured Reporting (SR) Documents. *Proceedings of the 10th International Protégé Conference*, pages 150–153, Hongrie, 07 2007.

[CI-155*] Sylvie Després et Sonia Mhiri. Usability via a Visualization Tool for the Semantic Indexing of Medical Reports (DICOM SR). *Proceedings of USAB 07 (Usability & HCI for Medicine and Health Care3)*, Autriche, 11 2007.

[CI-154] G.H.E.D. Duchamp, J.-G. Luque, J.-C. Novelli, C. Tollu et F. Toumazet. Hopf algebras of diagrams. *Proceedings of FPSAC'07*, Tianjin, China, 2007.

[CI-153] Christophe Fouqueré. A logical calculus for modelling interferences. *4th International Colloquium on Theoretical Aspects of Computing (ICTAC07)*, volume 4711 de *Lecture Notes in Computer Science*, pages 216–230. Springer, septembre 2007.

[CI-152] Nistor Grozavu, Younès Bennani et Sébastien Guérif. Towards a dimensionality reduction through unsupervised learning and local variable weighting. *ICMCS'07*, Chisinau, Moldavie., September 19-21, 2007 2007.

[CI-151] Sébastien Guérif et Younès Bennani. Dimensionality Reduction Through Unsupervised Features Selection. *Proceedings of the 10th International Conference on Engineering Applications of Neural Networks (EANN 2007)*, pages 98–106, Thessaloniki, Greece, aug 2007. Publishing Centre Alexander T.E.I. of Thessaloniki.

[CI-150*] Mourad Hakem et Franck Butelle. Reliability and scheduling on systems subject to failures. *Proc. of the 36th IEEE Intenational Conference on parallel Processing (ICPP)*, 2007.

[CI-149*] Thierry Hamon, Julien Derivière et Adeline Nazarenko. Ogmios : a scalable NLP platform for annotating large web document collections. *Proc. Corpus Linguistics 2007*, Birmingham, UK, Jul. 2007. Electronic publication. ALVIS STREP Project.

[CI-148] Thierry Hamon, Adeline Nazarenko, Thierry Poibeau, Sophie Aubin et Julien Derivière. A Robust Linguistic Platform for Efficient and Domain specific Web Content Analysis. *Proc. Recherche d'Information Assistée par Ordinateur (RIAO 2007) - Large-Scale Semantic Access to Content (Text, Image, Video and Sound)*, Pittsburgh, USA, May 2007. ALVIS STREP Project.

[CI-147] K. Klai, L. Petrucci et M. Reniers. An incremental and modular technique for checking LTL\X properties of Petri nets. *Proc. 27th International Conference on Formal Methods for Networked and Distributed Systems (FORTE'07), Tallinn, Estonia*, volume 4574 de *Lecture Notes in Computer Science*. Springer, juin 2007.

[CI-146] Julien Lavergne, Hanane Azzag, Christiane Guinot et Gilles Venturini. Construction incrémentale et visualisation de graphes de voisinage par des fourmis artificielles. Monique Noirhomme-Fraiture et Gilles Venturini, éditeurs, *Extraction et gestion des connaissances (EGC'2007)*, volume I de *Revue des Nouvelles Technologies de l'Information (RNTI-E-9)*, pages 135–146, Namur, Belgique, Janvier 2007. Cépadués Éditions.

[CI-145*] Julien Lavergne, Hanane Azzag, Christiane Guinot et Gilles Venturini. Incremental construction of neighborhood gra-vphs using the ants self-assembly behavior. *the annual IEEE International Conference on Tools with Artificial Intelligence (ICTAI'2007)*, Patras, Greece, October 29-31 2007. the Conference and Cultural Center of the University of Patras, accepted full paper.

[CI-144] Julien Lavergne, Hanane Azzag, Christiane Guinot et Gilles Venturini. Une approche incrémentale d'une méthode de classification non supervisée par nuages d'insectes volants. *XIVe Rencontre de la Société Francophone de Classification (SFC'2007)*, Paris, France, 5-7 septembre 2007. ENST.

[CI-143] Mustapha Lebbah, Nicoleta Rogovschi et Younés Bennani. BeSOM : Bernoulli on self organizing map. *Cap'2007 : conférence francophone sur l'apprentissage automatique, Plate-forme AFIA*, 2007.

[CI-142] Mustapha Lebbah, Nicoleta Rogovschi et Younés Bennani. BeSOM : Bernoulli on self organizing map. *International Joint Conferences on Neural Networks. IJCNN 2007, Orlando, Florida, August 12-17*, 2007.

[CI-141] V. Mogbil et V. Rahli. Uniform circuits, & boolean proof nets. Sergei Artemov et Anil Nerode, éditeurs, *Logical Foundations of Computer Science*, volume 4514 de *Lecture Notes in Computer Science*, pages 401–421. Springer-Verlag Heidelberg, 2007. International Symposium, LFCS 2007, New York, NY, USA, June 4-7.

[CI-140] Thibault Mondary, Amanda Bouffier et Adeline Nazarenko. Between browsing and search, a new model for navigating through large documents. Stella Vosniadou, Daniel Kayser et Athanassios Protopapas, éditeurs, *Proc. 2nd European Cognitive Science conf. (EuroCogSci'07)*, pages 634–639, Delphi, Grèce, May 2007. Lawrence Erlbaum Associates.

[CI-139] Catherine Recanati. Characteristics of diagrammatic reasoning. Stella Vosniadou, Daniel Kayser et Athanassios Protopapas, éditeurs, *Proc. 2nd European Cognitive Science conf. (EuroCogSci'07)*, pages 510–515, Delphi, Grèce, May 2007. Lawrence Erlbaum Associates.

[CI-138*] Catherine Recanati, Nicoleta Rogovschi et Younès Bennani. Sequencing of verbs - a study on tense and aspect using unsupervised learning. *Proc. Int. Conf. Recent Advances in Natural Language Processing (RANLP'07)*, Borovets, Bulgary, Sep. 2007. Institute for Parallel Processing, Bulgarian Academy of Sciences.

[CI-137*] Catherine Recanati, Nicoleta Rogovschi et Younès Bennani. The structure of verbal sequences analyzed with unsupervised learning techniques. *Proc. 3rd Language & Technology Conference : Human Language Technologies as a Challenge for Computer Science and Linguistics*, Poznań, Poland, Oct. 2007.

[CI-136] Nicoleta Rogovschi et Emmanuel Viennet. Finding community structure in social networks using statistical approaches. *V International Conference on Microelectronics and Computer Science*, septembre 2007.

[CI-135*] Davy Weissenbacher et Adeline Nazarenko. A Bayesian approach combining surface clues and linguistic knowledge : Application to the anaphora resolution problem. *Proc. Int. Conf. Recent Advances in Natural Language Processing (RANLP'07)*, Borovets, Bulgary, Sep. 2007. Institute for Parallel Processing, Bulgarian Academy of Sciences.

2006 [CI-134] Erick Alphonse et Aomar Osmani. On the connection between the phase transition of the covering test and the learning success rate. *16th International Conference on Inductive Logic Programming (ILP 2006)*, Lecture Notes in Artificial Intelligence, pages 27–29. Springer, 2006.

[CI-133] Erick Alphonse et Céline Rouveirol. Extension of the top-down data-driven strategy to ILP. *16th International Conference on Inductive Logic Programming (ILP 2006)*, /LNAI 4455/, pages /49–63/, 2006.

[CI-132] Sujeevan Aseervatham, Aomar Osmani et Emmanuel Viennet. BitSPADE : A lattice-based sequential pattern mining algorithm using bitmap representation. *Proceedings of the 2006 IEEE International Conference on Data Mining.*, pages 792–797, décembre 2006.

[CI-131] V. Atassi, P. Baillot et K. Terui. Verification of Ptime reducibility for system F terms via Dual Light Affine Logic. *Proceedings of Computer Science Logic (CSL'06)*, volume 4207 de *Lecture Notes in Computer Science*, pages 150–166. Springer, 2006.

[CI-130] Sophie Aubin et Thierry Hamon. Improving Term Extraction with terminological resources. Tapio Salakoski, Filip Ginter, Sampo Pyysalo et Tapio Pahikkala, éditeurs, *Advances in Natural Language Processing (5th International Conference on NLP, FinTAL 2006)*, LNAI 4139, pages 380–387. Springer, 2006.

[CI-129] Cyril Banderier et Bernhard Gittenberger. Analytic combinatorics of lattice paths : Enumeration and asymptotics for the average area. *Mathematics and computer science, IV (Nancy, 2006)*, Discrete Mathematics & Theoretical Computer Science Proceedings, pages 345–356. 2006.

[CI-128] Khalid Benabdeslem et Younès Bennani. Dendogram based SVM for multi-class classification. *28th International Conference Information Technology Interfaces*, pages 173–178, Cavtat / Dubrovnik, Croatia., juin 2006.

[CI-127] D. Berrabah, C.F. Ducateau et F. Boufarès. Active mechanisms for checking participation constraints in UML. *Proceedings of the 8th International Conference on Enterprise Information Systems*, pages 310–313, 2006.

[CI-126] Laure Bourgois. Finding alternative Web services to Parry Breakdowns. *Proceedings of the Third Starting AI Researchers' Symposium*, pages 232–243. IOS Press, 2006.

[CI-125] Laure Bourgois, Alexandre Delteil et François Lévy. Web Services subsumption with a specific PDL. *Advanced International Conference on Telecommunications and International Conference on Internet and Web Applications and Services (AICT/ ICIW)*, page 158. IEEE Computer Society, 2006.

[CI-124] Christophe Cérin, Dominique Bernardi, Michel Koskas, Mohamed Jemni et Hazem Fkaier. Geometrical interpretation for data partitioning on a grid architecture. *15th International Symposium on High Performance Distributed Computing (HPDC 2006)*, pages 355–356, France Paris, 2006. IEEE Computer Society.

[CI-123] Christophe Cérin, Jean-Christophe Dubacq et Jean-Louis Roch. Methods for partitioning data to improve parallel execution time for sorting on heterogeneous clusters. Yeh-Ching Chung et José E. Moreira, éditeurs, *Advances in Grid and Pervasive Computing*, volume 3947 de *Lecture Notes in Computer Science*, pages 175–186, Taichung, Taiwan, mai 2006. Springer.

[CI-122] C. Choppy, D. Hatebur et M. Heisel. Component composition through architectural patterns for problem frames. Pankaj Jalote, éditeur, *Proc. of the Asia Pacific Software Engineering Conference (APSEC-2006)*, pages 27–34. IEEE, 2006.

[CI-121] C. Choppy et G. Reggio. Requirements Capture and Specification for Enterprise Applications : a UML Based Attempt. Jun Han et Mark Staples, éditeurs, *Proc of the Australian Software Engineering Conference (ASWEC-2006)*, pages 19–28. IEEE, 2006.

[CI-120] Hakima Kadri Dahmani, Gérard Duchamp, Rawan Ghnemat, Hatem Hadj Kacem et Cyril Bertelle. Emerging decision support systems for geographical informations systems. *ICELM2*, 2006.

[CI-119] Paulin Jacobé de Naurois. A measure of space for computing over the reals. Arnold Beckmann, Ulrich Berger, Benedikt Löwe et John V. Tucker, éditeurs, *Logical Approaches to Computational Barriers, Second Conference on Computability in Europe, CiE 2006, Swansea, UK, June 30-July 5, 2006, Proceedings*, volume 3988 de *Lecture Notes in Computer Science*, pages 231–240. Springer, 2006.

[CI-118] David Delahaye et Micaela Mayero. Quantifier Elimination over Algebraically Closed Fields in a Proof Assistant using a Computer Algebra System. 151(1) :57–73, 2006.

[CI-117] Julien Derivière, Thierry Hamon et Adeline Nazarenko. A scalable and distributed NLP architecture for Web document annotation. Tapio Salakoski, Filip Ginter, Sampo Pyysalo et Tapio Pahikkala, éditeurs, *Advances in Natural Language Processing (5th International Conference on NLP, FinTAL 2006)*, LNAI 4139, pages 56–67. Springer, 2006.

[CI-116] Julien Derivière, Thierry Hamon, Adeline Nazarenko, Thierry Poibeau et Davy Weissenbacher. A robust linguistic infrastructure for efficient web content analysis : the ALVIS project. *Symposium on Digital Semantic Content across Cultures*, Paris, France, 2006. 15 pages.

[CI-115] Sylvie Després et Sylvie Szulman. Terminae Method and Integration Process for Legal Ontology Building. *Advances in Applied Artificial Intelligence : 19th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, IEA/AIE 2006, Annecy, France, June 27-30, 2006.*, pages 1014–1023, Annecy, France, June 2006. Springer Berlin/Heidelberg.

[CI-114] B.T. Doan, C. Lavault, S. Olariu et V. Ravelomanana. An energy-efficient initialization algorithm for random radio networks. *4th International Conference Research, Innovation, and Vision for the Future*, pages 121–139, 2006.

[CI-113] G. H. E. Duchamp, P. Blasiak, A. Horzela, K. A. Penson et A. I. Solomon. Feynman graphs and related Hopf algebras. *Journal of Physics : Conference Series, SSPCM'05*, volume 30, pages 107–118, Myczkowce, Poland, 2006. IoP.

[CI-112] R. Ghnemat, K. Khatatneh, S. Oqeili, C. Bertelle et G.H.E. Duchamp. Automata-based adaptive behavior for economic modeling using game theory. M.A. Aziz-Alaoui et C. Bertelle, éditeurs, *Emergent Properties in Natural and Artificial Dynamical Systems*, Understanding Complex Systems series, pages 173–185, Gammarth, Tunisia, 2006. Springer.

[CI-111] Natalia Grabar et Thierry Hamon. Terminology structuring through the derivational morphology. Tapio Salakoski, Filip Ginter, Sampo Pyysalo et Tapio Pahikkala, éditeurs, *Advances in Natural Language Processing (5th International Conference on NLP, FinTAL 2006)*, LNAI 4139, pages 652–663. Springer, 2006.

[CI-110] Sébastien Guérif et Younès Bennani. Selection of clusters number and features subset during a two-levels clustering task. *Proceeding of the 10th International Conference Artificial intelligence and Soft Computing 2006*, pages 28–33, Palma de Mallorca, Spain, août 2006.

[CI-109] Sébastien Guérif, Younès Bennani et Claude Baudoin. Connectionist and ethological approaches for discovering salient facial movement features in human gender recognition. *Proceeding of the 28th International Conference Information Technology Interfaces*, pages 189–194, Cavtat, Croatia, juin 2006.

[CI-108] Nejla Amara Hachmi. An ontology-based model for mobile agents adaptation in pervasive environments. *4th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA'06)*, pages 1106–1109, mars 2006.

[CI-107] L. Hillah, F. Kordon, L. Petrucci et N. Trèves. PN standardisation : a survey. *Proc. 26th International Conference on Formal Methods for Networked and Distributed Systems (FORTE'06), Paris, France*, volume 4229 de *Lecture Notes in Computer Science*, pages 307–322. Springer, septembre 2006.

[CI-106] Zeina Jrad, Francine Krief, Lahcène Dehni et Younès Bennani. Artificial Intelligence techniques in the dynamic negotiation of QoS : A user interface for the Internet New Generation. *International Federation for Information Processing (IFIP) TC6 Conference Smartnet in Autonomic Networking*, septembre 2006.

[CI-105] Zeina Jrad, Francine Krief et Rim Hammi. Video applications quality improvement in wireless systems : QoS negotiation and rate control. *International Conference on Digital Telecommunications ICDT'06*, France, septembre 2006.

[CI-104] Hager Karoui, Rushed Kanawati et Laure Petrucci. COBRAS : Cooperative CBR System for Bibliographical Reference Recommendation. *8th European Conference on Case-Based Reasonning (ECCBR'06)*, volume 4106 de *LNCS*, pages 76–90, Fethiye, Turkey, 2006. springer.

[CI-103] M. I. Khemmoudj et H. Bennaceur. Clique inference process for solving Max-CSP. *12th International Conference on Principles and Practice of Constraint Programming (Nantes, 2006)*, pages 246–250. Springer- Berlin Heidelberg, 2006.

[CI-102] M. I. Khemmoudj, M. Porcheron et H. Bennaceur. When constraint programming and local search solve the scheduling problem of électricité de France nuclear power plant outages. *12th International Conference on Principles and Practice of Constraint Programming (Nantes, 2006)*, pages 271–283. Springer- Berlin Heidelberg, 2006.

[CI-101] Ugo Dal Lago. Context semantics, linear logic and computational complexity. *21th IEEE Symposium on Logic in Computer Science (LICS 2006), 12-15 August 2006, Seattle, WA, USA*, pages 169–178, 2006.

[CI-100] Maria Malek et Rushed Kanawati. CBR-based sequence analysis : An overview. *Case-Based Reasoning in Knowledge Discovery and Data Mining*, juin 2006.

[CI-99] Adeline Nazarenko, Erick Alphonse, Julien Derivière, Thierry Hamon, Guillaume Vauvert et Davy Weissenbacher. The ALVIS format for linguistically annotated documents. *Proceedings of the 5th international conference on Language Resources and Evaluation, LREC 2006*, pages 1782–1786. ELDA, 2006.

[CI-98] J. Nzeutchap, F. Toumazet et F. Butelle. Kostka numbers and Littlewood-Richardson coefficients : distributed computation. R.C. King, M. Bylicki et J. Karwowski, éditeurs, *Symmetry, Spectroscopy and SCHUR, Proc. Prof. Brian G. Wybourne Commemorative Meeting*, Torun, Poland, 2006. Nicolaus Copernicus University Press.

[CI-97] K. A. Penson, P. Blasiak, G. Dattoli, G. H. E. Duchamp, A. Horzela et A. I. Solomon. Monomiality principle, Sheffer-type polynomials and the normal ordering principle. *Journal of Physics : Conference Series, SSPCM'05*, volume 30, pages 86–97, Myczkowce, Poland, 2006. IoP.

[CI-96] Thierry Poibeau. Dealing with metonymic readings of Named Entities. *The 28th Annual Conference of the Cognitive Science Society (CogSci 2006)*, pages 1962–1968, Vancouver, Canada, 2006. Cognitive Science Society.

[CI-95] Sampo Pyysalo, Tapio Salakoski, Sophie Aubin et Adeline Nazarenko. Lexical adaptation of Link Grammar to the biomedical sublanguage : a comparative evaluation of three approaches. *Proceedings of the Second International Symposium on Semantic Mining in Biomedicine (SMBM 2006)*, pages 60–67. Sophia Ananiadou, Juliane Fluck, 2006.

[CI-94] Vlady Ravelomanana et Alphonse Laza Rijamame. Creation and growth of components in a random hypergraph process. *Proceedings of 12th Annual International Computing and Combinatorics Conference 2006*, volume 4112 de *Lecture Notes in Computer Science*, pages 350–359. Springer, 2006.

[CI-93] A.I. Solomon, P. Blasiak, G. Duchamp, A. Horzela et K.A. Penson. Combinatorial physics, normal order and model Feynman graphs. B. Gruber, G. Marmo et N. Yoshinaga, éditeurs, *Proceedings of the Symposium Symmetries in Science XIII*, pages 150–166, Bregenz, Austria, 2006. Kluwer Academic Publishers.

[CI-92] F. Toumazet. Characterization of generic products in $Sp(2n, R)$ and $SO^*(2n)$. R.C. King, M. Bylicki et J. Karwowski, éditeurs, *Symmetry, Spectroscopy and SCHUR, Proc. Prof. Brian G. Wybourne Commemorative Meeting*, Torun, Poland, 2006. Nicolaus Copernicus University Press.

[CI-91] Davy Weissenbacher. Bayesian Network, a model for NLP ? *European Conference of the Association for Computational Linguistics (EACL 2006)*, pages 195–198, 2006.

2005 [CI-90] Touria Aït El Mekki et Adeline Nazarenko. Using NLP to build the hypertextuel network of a back-of-the-book index. *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP)*, pages 316–320. Institute for Parallel Processing, Bulgarian Academy of Sciences, 2005.

[CI-89] Sophie Aubin, Adeline Nazarenko et Claire Nédellec. Adapting a general parser to a sublanguage. G. Angelova, K. Bontcheva, R. Mitkov, N. Nicolov et N. Nikolov, éditeurs, *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP'05)*, pages 89–93. INCOMA Ltd., 2005.

[CI-88] P. Baillot et K. Terui. A Feasible Algorithm for Typing in Elementary Affine Logic. *Proceedings of 7th International Conference on Typed Lambda Calculi and Applications (TLCA 2005)*, volume 3461 de *Lecture Notes in Computer Science*, pages 55–70. Springer, 2005.

[CI-87] D. Berrabah, C.F. Ducateau et F. Boufarès. Analysing UML graphic constraints : How to cope with OCL. *Proceedings of the 3rd International Conference on Computer Science and its Applications*, pages 74–82, 2005.

[CI-86] Christophe Cérin et Michel Koskas. Mining traces of large scale systems. *6th International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP)*, volume 3719 de *Lecture Notes in Computer Science*, pages 132–138. Springer, 2005.

[CI-85] Tadeusz Czachorski et Ferhan Pekergin. A diffusion approximation model of an electronic-optical node. G. Zavattaro M. Bravetti, L. Kloul, éditeur, *Proceedings of Formal Techniques for Computer Systems and Business Processes : European Performance Engineering Workshop, EPEW 2005 and International Workshop on Web Services and Formal Methods, WS-FM 2005, Lecture Notes in Computer Science 3670/2005*, pages 187–199. Springer Verlag, 2005.

[CI-84] Hakima Kadri Dahmani, Cyril Bertelle, Gérard Duchamp et Aomar Osmani. Evolutive complex scheduling in interaction networks for quality improvement in geographical data base updating. *MCS 2005 - Modeling, Computation and Systems - in IEEE-ICECS*, 2005.

[CI-83] Lahcène Dehni, Francine Krief et Younès Bennani. Power control and clustering in wireless sensor networks. *IFIP Med-Hoc-Net 2005 : Mediterranean Ad Hoc Networking Conference*, pages 31–40. Springer, juin 2005.

[CI-82] Sylvie Després et Sylvie Szulman. Merging of Legal Micro-Ontologies from European Directives. *LOAIT - Legal Ontologies and Artificial Intelligences Techniques*, pages 149–155. IAAIL Workshop series, 2005.

[CI-81] G. H. E. Duchamp, H. Hadj-Kacem et E. Laugerotte. On the decomposition of boolean functions. J.-F. Michon, P. Valarcher et J.-B. Yunès, éditeurs, *Proceedings of BFCA'05 Conference, March 7–8, 2005 Rouen, France*, pages 115–128. Universités de Rouen et du Havre, 2005.

[CI-80] G. H.E. Duchamp, K. Khatatneh et C. Bertelle. Rewriting tables and memorized semirings. Noureddine Boudriga et Mohammad S. Obaidat, éditeurs, *Proc. of the 12th IEEE International Conference on Electronics, Circuits and Systems*, Gammarth, Tunisia, 2005. IoP.

[CI-79] G.H.E. Duchamp, J-G. Luque, K-A. Penson et C. Tollu. Free quasi-symmetric functions, product actions and quantum field theory of partitions. *Proceedings of FPSAC'05*, pages 649–660, Taormina, Italy, 2005.

[CI-78] Mourad Hakem et Franck Butelle. Efficient critical task scheduling parallel programs on a bounded number of processors. *Proc. of the 17th International Conference on Parallel and Distributed Computing and Systems*, ActaPress, pages 139–144. ActaPress, 2005.

[CI-77] S. Hamdoun et F. Boufarès. Building a data warehouse using heterogeneous data sources. *Proceedings of the International Conference on Sciences of Electronic, Technologies of Information and Telecommunications*, 2005.

[CI-76] S. Hamdoun et F. Boufarès. Relationships between attributes to integrate heterogeneous data sources : an SQL methodology. *Proceedings of the IASTED International Conference on Databases and Applications (IASTED : International Association of Science and Technology for Development)*, pages 23–28, 2005.

[CI-75] Zeina Jrad, Badr Benmammar, Joeseph Correa, Francine Krief et Nadek Mbarek. A user assistant for QoS negotiation in a dynamic environment using agent technology. *Second IFIP International Conference on Wireless and Optical Communications Networks WOCN*, mars 2005.

[CI-74] H. Kadri-Dahmani, C. Bertelle, G. H.E. Duchamp et A. Osmani. Evolutive complex scheduling in interaction networks for quality improvment in geographical data base updating. Noureddine Boudriga et Mohammad S. Obaidat, éditeurs, *Proc. of the 12th IEEE International Conference on Electronics, Circuits and Systems*, Gammarth, Tunisia, 2005. IoP.

[CI-73] Daniel Kayser. What logic can and cannot do in natural language semantics. *First Indian Conference on Logic and its relationship with other disciplines*, Actes en cours d'édition, Bombay, Inde, 2005.

[CI-72] M. I. Khemmoudj, H. Bennaceur et Anass Nagih. Combining arc-consistency and dual lagrangean relaxation for filtering CSPs. *2th International Conference on Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems (Praha, 2005)*, pages 258–272. Springer - Berlin Heidelberg, 2005.

[CI-71] R. C. King, C. Tollu et F. Toumazet. The hive model and the factorisation of Kostka coefficients. *Proceedings of FPSAC'05*, pages 721–732, Taormina, Italy, 2005.

[CI-70] Christian Lavault, Gabriel Paillard et Felipe Franca. A distributed prime sieving algorithm based on smer. *4th International Symposium on Parallel and Distributed Computing*, pages 139–146, ISPDC'2005, 2005.

[CI-69] Maria Malek et Rushed Kanawati. A data-driven CBR approach : A contineous maintenance strategy. *Proceedings of the first International conference on Advances in Intelligent systems : Theory and applications (AISTA'05)*, Luxembourg, décembre 2005.

[CI-68] Jérôme Monnot et Sophie Toulouse. Approximation results for the weighted p_4 partition problems. *Foundamentals of Computation Theory*, pages 388–396, 2005.

[CI-67] Jérôme Monnot et Sophie Toulouse. The complexity of the p_k partition problem and related problems in bipartite graphs. *SOFSEM 2007 : Theory and Practice of Computer Science*, volume 4362 de *Lecture Notes in Computer Science*, pages 422–433, 2005.

[CI-66] Hassine Moungla et Francine Krief. Conflict resolution in policy based management. *The 16th Annual IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, septembre 2005.

[CI-65] Alexandre Pauchet, Nathalie Chaignaud et Amal El Fallah-Seghrouchni. A cognitive model of interaction for software agents. *IEEE IAT - International Conference on Intelligent Agent Technology*, pages 281–284, septembre 2005.

[CI-64] Nadia Pisanti, Henry Soldano et Mathilde Carpentier. Incremental Inference of Relational Motifs with a Degenerate Alphabet. *16th Annual Symposium on Combinatorial Pattern Matching (CPM)*, volume 3537 de *Lecture Notes in Computer Science*, pages 229–240. Springer, 2005.

[CI-63] Vlady Ravelomanana. Distributed k-clustering algorithms for random wireless multihop networks. *Proceedings of the IEEE ICN 2005*, pages 109–116, 2005.

[CI-62] Véronique Ventos et Henry Soldano. Alpha Galois lattices : An overview. *Proceedings of ICFCA'05, Lecture Notes on Computer Science*, 3403 :298–313, 2005.

2004 [CI-61] Nejla Amara-Hachmi. A generic component-based methodology for extensible mobile agents design. *IEEE International Conference on Advances in Intelligent Systems : Theory and Applications AISTA'04*, novembre 2004.

[CI-60] Nejla Amara-Hachmi, Gilles Klein et Amal El Fallah-Seghrouchni. Towards MAS's design regarding communications and mobility. *Turkish Symposium on Artificial Intelligence and Neural Networks TAINN'04*, juin 2004.

[CI-59] P. Baillot et K. Terui. Light types for polynomial time computation in lambda-calculus. *Proceedings of 19th IEEE Symposium on Logic in Computer Science (LICS 2004)*, pages 266–275. IEEE Computer Society, 2004.

[CI-58] Patrick Baillot et Virgile Mogbil. Soft lambda-calculus : a language for polynomial time computation. *Proceedings of the 7th International Conference on Foundations of Software Science and Computation Structures (FOSSACS'04)*, volume 2987 de *Lecture Notes in Computer Science*, pages 27–41. Springer, 2004.

[CI-57] Cyril Banderier, Jean-Marie Le Bars et Vlady Ravelomanana. Generating functions for kernels of digraphs (enumeration & asymptotics for a constraint from game theory). *Proceedings FPSAC'04 (Vancouver, 2004)*, Formal Power Series and Algebraic Combinatorics, pages 91–105. (electronic), 2004.

[CI-56] Khalid Benabdeslem et Younès Bennani. An incremental SOM for web navigation patterns clustering. *26th International Conference Information Technology Interfaces ITI 2004*, juin 2004.

[CI-55] Badr Benmammar et Francine Krief. Agents for wireless environments. *International Conference on Telecommunication Systems, Modeling and Analysis (ICTSM 2004. IFIP WG 7.3)*, juillet 2004.

[CI-54] C. Choppy et G. Reggio. Improving Use Case Based Requirements Using Formally Grounded Specifications. *Fundamental Approaches to Software Engineering (FASE)*, volume 2984 de *Lecture Notes in Computer Science*, pages 244–260. Springer, 2004.

[CI-53] C. Choppy et G. Reggio. Using UML for Problem Frame Oriented Software Development. Walter Dosch et Narayan Debnath, éditeurs, *Proc of the ISCA 13th Int. Conf. on Intelligent and Adaptative Systems and Software Engineering (IASSE-2004)*, pages 239–244. The International Society for Computers and Their Applications (ISCA), 2004.

[CI-52] Tadeusz Czachorski et Ferhan Pekergin. Diffusion approximation as a modelling tool in congestion control and performance evaluation. Demetres Kouvatsos, éditeur, *Proceedings of HET-NETs '04 Performance Modelling and Evaluation of Heterogenous Networks*, United Kingdom Ilkley, 2004.

[CI-51] Hakima Kadri Dahmani et Aomar Osmani. How to update databases in a consistent way ? *UDMS 2004 : 24th Urban Data Management Symposium*, 2004.

[CI-50] Sylvie Després et Sylvie Szulman. Construction of a Legal Ontology from a European Community Legislative Text. Thomas Gordon, éditeur, *Jurix 2004- Berlin*, pages 79–88. IOS Press, 2004.

[CI-49] F. Djerourou, C. Lavault, G. Paillard et V. Ravelomanana. Assigning codes in a random wireless network. *International Conference on Telecommunications*, numéro 3124 Lecture Notes in Computer Science, pages 348–353, Fortaleza, Brasil, 2004.

[CI-48] G. Duchamp, K.A. Penson, A.I. Solomon, A. Horzela et P. Blasiak. One-parameter groups and combinatorial physics. J. Govaerts, M. N. Hounkonnou et A. Z. Msezane, éditeurs, *Proceedings of the Third International Workshop on Contemporary Problems in Mathematical Physics (COPROMAPH3)*, pages 436–, Porto-Novo, Benin, 2004. World Scientific Publishing.

[CI-47] Françoise Gayral, Daniel Kayser et François Lévy. Challenging the Principle of Compositionality in Interpreting Natural Language Texts. *Conference on Compositionality, Concepts and Cognition*, pages 111–127, Düsseldorf, Allemagne, 2004.

[CI-46] Alejandro Guerra-Hernández, Amal El Fallah-Seghrouchni et Henry Soldano. Distributed learning in intentional BDI multi-agent systems. *5th Mexican International Conference on Computer Science (ENC 2004)*, pages 225–232. IEEE Computer Society, 2004.

[CI-45] Mourad Hakem et Franck Butelle. A new on-line scheduling algorithm for distributed real-time system. *Proc. of the 3rd Int. Symp. and School on Advanced Distributed Systems (ISSADS)*, Lecture Notes in Computer Science, pages 241–251. Springer, 2004.

[CI-44] S. Hamdoun et F. Boufarès. Relationships between attributes to integrate heterogeneous data sources. *Proceedings of the 2004 International Arab Conference on Information Technology*, volume 2, pages 602–607, 2004.

[CI-43] A. Horzela, P. Blasiak, G. Duchamp, K. A. Penson et A.I. Solomon. A product formula and combinatorial field theory. J. Govaerts, M. N. Hounkonnou et A. Z. Msezane, éditeurs, *Proceedings of the XI International Conference on Symmetry Methods in Physics (SYMPHYS-11)*, Prague, Czech Republic, 2004. JINR Publishers.

[CI-42] Zeina Jrad et Francine Krief. An intelligent interface for the dynamic negotiation of QoS. *ARCADE. ACS/IEEE International Conference on Pervasive Services. ICPS'2004*, juillet 2004.

[CI-41] Daniel Kayser et Farid Nouioua. About norms and causes. Valerie Barr et Zadravko Markov, éditeurs, *The 17th Florida Artificial Intelligence Research Society Conference (FLAIRS'04)*, pages 502–507. AAAI Press, 2004.

[CI-40] Daniel Kayser et Farid Nouioua. Representing knowledge about norms. *Proc. 16th European Conference on Artificial Intelligence*, Valencia, Spain, pages 363–367. IOS Press, 2004.

[CI-39] R.C. King, C. Tollu et F. Toumazet. Stretched Littlewood-Richardson and Kostka coefficients. P. Winternitz, J. Harnad, C.S. Lam et J. Patera, éditeurs, *Symmetry in physics : in Memory of Robert T. Sharp*, volume 34 de *CRM Proceedings and Lecture Notes*, pages 99–112. Amer. Math. Soc., 2004.

[CI-38] Francine Krief et Dominique Bouthinon. A learning and intentional local policy decision point for dynamic QoS provisioning. *IFIP TC6, Third International Conference on Network Control and Engineering for QoS, Security and Mobility, NetCon 2004*, pages 277–288. Springer, 2004.

[CI-37] L. Kristensen et L. Petrucci. An approach to distributed state space exploration for coloured Petri nets. *Proc. 25th Int. Conf. Application and Theory of Petri Nets (ICATPN'2004), Bologna, Italy, June 2004*, volume 3099 de *Lecture Notes in Computer Science*, pages 474–483. Springer, juin 2004.

[CI-36] C. Lakos et L. Petrucci. Modular analysis of systems composed of semiautonomous subsystems. *Proc. 4th Int. Conf. on Application of Concurrency to System Design (ACSD'04), Hamilton, Canada, June 2004*, pages 185–194. IEEE Comp. Soc. Press, juin 2004.

[CI-35] Maria Malek et Rushed Kanawati. A Data Driven CBR Approach : A Contineous Maintenance Strategy. *International Conference on Advances in Intelligent Systems : Theory and Applications*, page 6. IEEE, Novembre 2004.

[CI-34] Hassine Moungla et Francine Krief. Service differentiation over GMPLS. *IEEE conference/ ICT'2004. Lecture Notes on Computer Science*. Springer, août 2004.

[CI-33] Thierry Poibeau. Automatic extraction of paraphrastic phrases from medium size corpora. Sergei Nirenburg, éditeur, *Actes de la conférence Computational Linguistics (COLING'2004)*, pages 638–644. Association for Computational Linguistics, 2004.

[CI-32] Thierry Poibeau et Bénédicte Goujon. Semi-automatic generation of command grammar. *Language and Resource Evaluation Conference (LREC 2004)*, pages 2147–2150, Lisbonne, Portugal, 2004.

[CI-31] Sidi Mohammed Sedjelmaci. The accelerated euclidean algorithm. *Proc. of the Meeting on Computer Algebra and Applications (EACA)*, pages 283–287, 2004.

[CI-30] A.I. Solomon, G. Duchamp, P. Blasiak, A. Horzela et K. A. Penson. Partition functions and graphs : A combinatorial approach. J. Govaerts, M. N. Hounkonnou et A. Z. Msezane, éditeurs, *Proceedings of the XI International Conference on Symmetry Methods in Physics (SYMPHYS-11)*, Prague, Czech Republic, 2004. JINR Publishers.

[CI-29] A.I. Solomon, G. Duchamp, P. Blasiak, A. Horzela et K.A. Penson. Normal order : Combinatorial graphs quantum theory and symmetries. P.C. Argyres, T.J. Hodges, F. Mansouri, J.J. Scanio, P. Suranyi et L.C.R. Wijewardhana, éditeurs, *Proceedings of the 3rd International Symposium*. World Scientific Publishing, 2004.

[CI-28] Véronique Ventos, Henry Soldano et Thibaut Lamadon. Alpha Galois lattices. *Int. Conf. on Data Mining (ICDM'04)*, pages 555–558. IEEE Press, 2004.

[CI-27] Farida Zehraoui et Younès Bennani. M-SOM-ART : Growing self organizing map for sequence clustering and classification. *ECAI'04, European Conference on Artificial Intelligence*, 2004.

[CI-26] Farida Zehraoui et Younès Bennani. M-SOM : Matricial self organizing map for sequences clustering and classification. *IJCNN'04, International Joint Conference on Neural Networks*, juillet 2004.

[CI-25] Farida Zehraoui, Rushed Kanawati et Sylvie Salotti. CASEP2 : Hybrid Case-based Reasoning System for Sequence Processing. *Advances in Case-Based Reasoning, 7th European Conference (ECCBR 2004)*, pages 449–463, 2004.

2003

[CI-24] Anne-Elisabeth Baert et Vlady Ravelomanana. Limit theorem on the creation of multicyclic components. *Proceedings FPSAC'03 (Linköping University – Sweden)*, Formal Power Series and Algebraic Combinatorics, pages 176–187. Linköping University, 2003.

[CI-23] Cyril Banderier, René Beier et Kurt Mehlhorn. Smoothed analysis of three combinatorial problems. *Mathematical foundations of computer science 2003*, volume 2747 de *Lecture Notes in Comput. Sci.*, pages 198–207. Springer, Berlin, 2003.

[CI-22] Hachemi Bennaceur et Aomar Osmani. Computing lower bound for MAX-CSP problems. *International Conference on Industrial, Engineering Applications of Artificial Intelligence and Expert Systems*, Lecture Notes in Artificial Intelligence, pages 614–624. Springer, 2003.

[CI-21] Khellaf Bouandas et Aomar Osmani. Optimal algorithm for temporal patterns discovery. *FLAIRS-2003*, pages 455–460. AAAI press, 2003.

[CI-20] F. Boufarès, H. Bennaceur et A. Osmani. On the consistency of cardinality constraints in UML modeling. *Proc. of the 10th ISPE International Conference on Concurrent Engineering : Research and Applications, (ISPE : International Society for Productivity Enhancement)*, pages 287–292, 2003.

[CI-19] Faouzi Boufares, Hacemi Bennaceur et Aomar Osmani. On the consistency of cardinality constraints in uml modelling. *10th ISPE International Conference on Concurrent Engineering*, volume 1, pages 287–292. A. A. Balkema Publishers, 2003.

[CI-18] C. Choppy, D. Hatebur et M. Heisel. Use of Patterns in Formal Development : Systematic Transition From Problems to Architectural Designs. *Recent Trends in Algebraic Development Techniques*, volume 2755 de *Lecture Notes in Computer Science*, pages 205–220. Springer, 2003.

[CI-17] Tadeusz Czachorski et Ferhan Pekergin. Diffusion approximation : modelling transient flows of multiclass queueing network. *Proceedings of Int. Network Optimization Conf.*, France Evry, 2003.

[CI-16] G. Duchamp, C. Tollu et F. Toumazet. Combinatorics of the q-shuffle product and Bargmann-Fock spaces of the $q$uon Weyl algebra. *Physical and Mathematical Aspects of Symmetry, Proc. 24th Int. Conf. on Group-theoretical Methods in Physics*, volume 173 de *Conference series*, pages 803–807. IoP, 2003.

[CI-15] S. Hamdoun et F. Boufarès. An XML approach of construction and self-maintenance of a web warehouse. *Proceedings of the 2003 International Arab Conference on Information Technology*, 2003.

[CI-14] Zeina Jrad, Francine Krief et Badr Benmammar. A intelligent user interface for the dynamic negotiation of QoS. *International Conference on Telecommunications, ICT'2003*, février 2003.

[CI-13] Rushed Kanawati. A Multi-agent System for Improving Result Ranking Service of Web Search Engines. *10th ISPE International conference on concurrent engineering (CE 2003)*, volume 1, pages 91–99, Funchal, Madeira, June 2003. Balkema.

[CI-12] Max I. Kanovich et Jacqueline Vauzeilles. Coping polynomially with numerous but identical elements within planning problems. *Computer Science Logic, 17th International Workshop, CSL 2003, 12th Annual Conference of the EACSL, and 8th Kurt Güdel Colloquium, KGC 2003, Vienna, Austria, August 25-30, 2003, Proceedings*, volume 2803 de *Lecture Notes in Computer Science*, pages 285–298. Springer, 2003.

[CI-11] Christian Lavault, Jean-Franois Marckert et Vlady Ravelomanana. Quasi-optimal leader election algorithms in radio networks with loglogarithmic awake time slots. *10th IEEE International Conference on Telecommunications*, pages 1113–1119, 2003.

[CI-10] Larbi Ait Mohamed et Aomar Osmani. A model based on domain theory for the supervision of the telecommunications networks. *The 10th International Conference on telecommunications ICT-2003*, pages 31–40. IEEE Springer, 2003.

[CI-9] Hassine Moungla et Francine Krief. For an intelligent policy-based telecommunication network management. *International Conference on Telecommunication Systems, Modelling and Analysis. ICTSM'2003. IFIP-IEE. WG 7.3*, octobre 2003.

[CI-8] Hassine Moungla et Francine Krief. Vers une gestion intelligente des réseaux d'opérateurs offrant des services IP. *International Congress On Networks And Services Mangement, GRES'2003*, février 2003.

[CI-7] Aomar Osmani. Qualitative point sequential patterns. *7th International Conference on Knowledge Based Intelligent Information KES2003*. Springer, 2003.

[CI-6] Aomar Osmani. STCSP : A representation model for sequential patterns. *Foundations and Applications of Spatio-Temporal Reasoning. AAAI Sumposium*, 2003.

[CI-5] Alexandre Pauchet, Nathalie Chaignaud et Amal El Fallah-Segrouchni. BDIGGY : a cognitive architecture for cooperative problem solving. *IEEE IAT - International Conference on Intelligent Agent Technology*, pages 523–526, octobre 2003.

[CI-4] Thierry Poibeau. The multilingual Named Entity recognition framework. Ann Copestake et Jan Hajic, éditeurs, *European Conference of the Association for Computational Linguistics (EACL 2003)*, pages 155–158, Budapest, Hongrie, 2003. Association for Computational Linguistics.

[CI-3] Arnaud Zeboulon, Younès Bennani et Khalid Benabdeslem. Hybrid connectionist approach for knowledge discovery from web navigation patterns. *ACS/IEEE International Conference on Computer Systems and Applications*, juillet 2003.

[CI-2] Farida Zehraoui, Rushed Kanawati et Sylvie Salotti. Case base maintenance for improving prediction quality. *The 6th International Conference on Case-Based Reasoning (ICCBR'03)*, volume 2689 de *Lecture Notes in Computer Science*, pages 261–275. springer, June 2003.

[CI-1] Rushed Kanawati et Maria malek. Computing Social Networks for Information Sharing : A Case-based Approach, volume = 4564, year = 2007. Schuler Douglas, éditeur, *Second international conference on Online Communities and Social Computing*, LNCS, pages 86–95, Beijing, China, July. Springer.

Workshops et autres conférences avec comité de lecture

2008 [CO-194*] Sylvie Després. Construction dï¿½ontologies à partir de textes dans le domaine juridique. *Actes du colloque annuel de la Société Française de Terminologie (SFT)*, France, 2008.

2007 [CO-193] Laurent Alfandari. A greedy approximation algorithm for the soft-capacitated facility location problem. *EURO XXII Congress*, Prague, République Tchèque, 2007.

[CO-192] Laurent Alfandari. Approximation du problème de localisation de centres à capacités variables. *ROADEF 2007*, pages 35–36, Grenoble, 2007.

[CO-191] Laurent Audibert. Désambiguïsation lexicale automatique : sélection automatique d'indices. Nabil Hathout et Philippe Muller, éditeurs, *Actes de la 14ème conférence sur le Traitement Automatique des Langues Naturelles (TALN'07)*, pages 13–22, Toulouse, France, Juin 2007. IRIT Press.

[CO-190] Delphine Battistelli, Jean-Luc Minel et Sylviane R. Schwer. Calendar expressions ordering in text corpora, juin 2007.

[CO-189*] A. BenHadjAli, F. Boufarès et S. BenAhmed. Handling integrity constraints via model transformation. *Proceedings of the 16th International conference on Information Systems Development (Galway, Ireland, August 29-31, 2007)*, 2007. À paraître.

[CO-188*] D. Berrabah et F. Boufarès. Constraints checkinh in UML class diagrams : SQL vs OCL. *Proceedings of the 18th International conference on Database and Expert Systems Applications*, 2007. À paraître.

[CO-187] D. Berrabah, F. Boufarès et M. Badri. Un métamodèle pour l'étude de la cohérence globale des contraintes dans les bases de données. *Actes du Colloque sur l'Optimisation et les Systèmes d'Information*, pages 485–497, 2007.

[CO-186] O. Bertrand, P. Carle et C. Choppy. Chronicle modelling using automata and colored Petri nets. *The 18th International Workshop on Principles of Diagnosis (DX-07)*, pages 229–234, 2007.

[CO-185] O. Bertrand, P. Carle et C. Choppy. Vers une exploitation des simulations distribuées par les chroniques. *8e Rencontres nationales des Jeunes Chercheurs en Intelligence Artificielle (RJCIA 07)*, pages 15–30, 2007.

[CO-184] Aurélien Bossard. Vers une ressource prédicative pour l'extraction d'information. Nabil Hathout et Philippe Muller, éditeurs, *Actes de la 14ème conférence sur le Traitement Automatique des Langues Naturelles (TALN'07)*, page 154, Toulouse, France, Juin 2007. IRIT Press.

[CO-183*] Amanda Bouffier et Thierry Poibeau. Exploiter la structure des documents : un enjeu pour la fouille de textes. *Actes du colloque Veille Stratégique Scientifique & Technologique (VSST'07)*, Maroc, Oct. 2007.

[CO-182] C. Boukala et L. Petrucci. Towards distributed verification of Petri nets properties. *Proc. 1st International Workshop on Verification and Evaluation of Computer and Communication Systems (VECOS'07), Algiers, Algeria*, eWiC, pages 15–26. British Computer Society, mai 2007.

[CO-181*] Valentina Ceausu et Sylvie Després. Learning Term to Concept Mapping Through Verbs : A Case Study. *Proceedings of Semantic Authoring, Annotation and Knowledge Markup Workshop of Conference on Knowledge Capture (KCap07)*, Canada, 10 2007.

[CO-180] Lucile Chalet, Hanane Azzag, Christiane Guinot et Gilles Venturini. Un algorithme préliminaire pour la structuration visuelle de données inspiré des modéles de construction de nids chez les insectes sociaux. *Atelier Fouille du Données et Algorithmes Biomimétiques aux 7ème EGC'07 Extraction et Gestion des Connaissances*, pages 13–24, jan 2007.

[CO-179] C. Choppy, L. Petrucci et G. Reggio. Designing coloured Petri net models : a method. *Proc. Workshop on Practical Use of Coloured Petri Nets, Aarhus, Denmark*, octobre 2007.

[CO-178] Sylvie Despres, Frédéric Fürst et Sylvie Szulman. Construction d'une ontologie du domaine HSE. *Ingénierie des Connaissances - 18èmes journées francophones*, pages 133–144, Grenoble, France, 2007. Cépaduès.

[CO-177] Sylvie Després, Sonia Mhiri et Ezzaddine Zagrouba. Unification d'ontologies hétérogènes pour l'indexation sémantique de comptes rendus standardisés d'examens du patient. *Actes de l'atelier Ontologies et gestion de l'hétérogénéité sémantique (OGHS). Plateforme AFIA 2007*, pages 60–62, Grenoble, France, 07 2007.

[CO-176] M. Elati, P. Neuvial, C. Rouveirol, F. Radvanyi et E. Barillot. A methodology for discovering cooperative regulatory pattern from ternary gene expression data. *Cap'2007 : conférence francophone sur l'apprentissage automatique, Plate-forme AFIA*, 2007.

[CO-175] Sylviane R. Schwer et Jean-Paul Cardinal. 1870 : une catastrophe, une aubaine pour l'université française et les scientifiques. *Actes des 14èmes Journées de Rochebrune, Rencontres interdisciplinaires sur les systèmes complexes naturels et artificiels. Catastrophes, discontinuités, ruptures, limites, frontières.*, pages 289–300, janvier 2007.

[CO-174] Sébastien Guérif et Younès Bennani. Sélection de Variables en Apprentissage Numérique Non Supervisé. *Conférence francophone sur l'Apprentissage automatique (CAp'2007)*, pages 221–236, Grenoble, France, July 2007.

[CO-173] Thierry Hamon, Julien Derivière et Adeline Nazarenko. Ogmios : une plate-forme d'annotation linguistique. Nabil Hathout et Philippe Muller, éditeurs, *Actes de la*

14ème conférence sur le Traitement Automatique des Langues Naturelles (TALN'07), pages 103–112, Toulouse, France, Juin 2007. IRIT Press. ALVIS STREP Project.

[CO-172] Rushed Kanawati. A community aware web search engine. *In proceedings HCI International 2007*, Beijing, china, july 2007. poster.

[CO-171] Hager Karoui. Systèmes de recommandation de références bibliographiques et de formation de comités. *Actes des 3èmes Rencontre Inter-Associations (RIAs 07) : Recherche et Extraction d'Information*, Toulouse - IRIT, 2007.

[CO-170] Mohand-Ou-Idir Khemmoudj et Hachemi Bennaceur. Valid inequality based lower boubnds for WCSP. *CP07*, Providence, USA, 2007.

[CO-169] C. Lakos et L. Petrucci. Modular state spaces and place fusion. *Workshop on Petri Nets and Software Engineering (PNSE'07, associated with PETRI NETS'07), Siedlce, Poland*, pages 175–190. University of Podlasie, juin 2007.

[CO-168] Luc Lamontagne, Rim Bentibebel, Erwan Miry et Sylvie Després. Finding Lexical Relations for the Reuse of Investigation Reports, Beyond Retrieval. *Proceedings of Fourth Workshop on Textual CBR, ICCBR'07*, pages 252–256, Irlande, 08 2007.

[CO-167] Julien Lavergne, Hanane Azzag, Christiane Guinot et Gilles Venturini. Construction et exploration interactive d'un graphe de documents. *3émes Rencontres Inter-Associations (RIA's 2007)*, Toulouse, France, 12-13 mars 2007. IRIT.

[CO-166] Lucas Létocart. Inégalités valides pour le problème de multiflot maximum. *2ème journée scientifique Polyèdres et Optimisation Combinatoire*, Paris, France, 2007.

[CO-165] Lucas Létocart, Anass Nagih et Gérard Plateau. Re-optimization in the resolution of the 0-1 quadratic knapsack problem. *EURO XXII*, Prague, République tchèque, 2007.

[CO-164] Lucas Létocart, Anass Nagih et Gérard Plateau. Réoptimisation dans les méthodes lagrangiennes pour le sac-à-dos quadratique. *1ère journée KnapSack et Optimisation*, Paris, France, 2007.

[CO-163] Lucas Létocart, Marie-Christine Plateau et Gérard Plateau. New bounds for the 0-1 exact k-item quadratic knapsack problem. *ECCO XX*, Limassol, Chypre, 2007.

[CO-162] Jean-Vincent Loddo et Thierry Hamon. Marionnet/Iutoppix : un environnement pédagogique pour la simulation de réseaux locaux. *Environnement Informatique pour l'Apprentissage Humain (EIAH'07)*, Juil. 2007. Démonstration.

[CO-161*] Jean-Vincent Loddo et Luca Saiu. Marionnet – How to Implement a Virtual Network Laboratory in Six Months and be Happy. *ACM SIGPLAN Workshop on ML*. ACM Press, 2007.

[CO-160] Donazzan Marta, Sylviane R. Schwer et Lucia Tovena. Apport du treillis associé à l'interprétation d'adverbes itératifs, septembre 2007.

[CO-159] Marie-Chirstine Plateau, Leo Liberti et Laurent Alfandari. Edge cover by bipartite subgraphs. *6th CTW Workshops*, Enschende, Pays-Bas, 2007.

[CO-158*] Thierry Poibeau. Up13 : Knowledge poor methods (sometimes) perform poorly. *Proceedings of the ACL SemEval workshop ACL SemEval (Semantic Evaluation) workshop*, Tchèque, République Prague, 2007. Association for Computational Linguistics.

[CO-157*] Thierry Poibeau et Aurélien Bossard. Adaptation d'une ressource prédicative pour l'extraction d'information. *Actes du Colloque Lexique et Grammaire 2007 Colloque Lexique et Grammaire 2007*, Bonifacio, France, 2007. LADL.

[CO-156] Catherine Recanati et Nicoleta Rogovschi. Enchaînements verbaux - étude sur le temps et l'aspect utilisant des techniques d'apprentissage non supervisé. Nabil Hathout et Philippe Muller, éditeurs, *Actes de la 14ème conférence sur le Traitement Automatique des Langues Naturelles (TALN'07)*, pages 379–388, Toulouse, France, Juin 2007. IRIT Press.

[CO-155] Nicoleta Rogovschi, Younès Bennani et Catherine Recanati. Apprentissage neuro-markovien pour la classification non supervisée de données structurées en séquences.

Actes de Extraction et Gestion des Connaissances (EGC'07), Atelier « Fouille de données et algorithmes biomimétiques », pages 49–60, Namur, Belgique, Jan. 2007. Cépaduès Editions.

[CO-154] Nicoleta Rogovschi et Emmanuel Viennet. Finding community structure in large social networks. Poster at NATO Advanced Study Institute on Mining Massive Data Sets for Security, septembre 2007.

[CO-153] Sylviane R. Schwer. Au sujet de la représentation et du raisonnement pour les informations temporelles : vers un modèle purement algébrique., mai 2007.

[CO-152] Sylviane R. Schwer. Topologie d'une cité fantôme [d'alain robbe-grillet] : une approche catastrophique ou chaotique. *Actes des 14èmes Journées de Rochebrune, Rencontres interdisciplinaires sur les systèmes complexes naturels et artificiels. Catastrophes, discontinuités, ruptures, limites, frontières.*, pages 273–287, janvier 2007.

[CO-151*] Sylvie Szulman et Sylvie Despres. Exploitation linguistique d'une ontologie générique. *TIA 2007 Terminologie et Intelligence artificielle*, Sophia-Antipolis, France, 10 2007.

[CO-150] Nora Touati, Lucas Létocart et Anass Nagih. Diversification des solutions et réoptimisation pour la résolution de problèmes de plus court chemin avec contraintes de ressources dans un schéma de génération de colonnes. *ROADEF 2007*, pages 419–420, Grenoble, France, 2007.

[CO-149] Sophie Toulouse. Approximation différentielle : histoire, motivations, enjeux. *ROADEF 2007*, pages 423–424, Grenoble, 2007.

[CO-148] Davy Weissenbacher et Adeline Nazarenko. A bayesian classifier for the recognition of the impersonal occurrences of the 'it' pronoun. *Proc. 6th Discourse Anaphora and Anaphor Resolution colloquium (DAARC'07)*, pages pp.145–150, Lagos, Portugal, May 2007.

[CO-147] Davy Weissenbacher et Adeline Nazarenko. Identifier les pronoms anaphoriques et trouver leurs antécédents : l'intérêt de la classification bayésienne. *Actes de la 14ème conférence sur le Traitement Automatique des Langues Naturelles (TALN'07)*, pages 145–155, Toulouse, France, Juin 2007. IRIT Press.

[CO-146*] Haifa Zargayouna, Olivier Hamon et Adeline Nazarenko. Evaluation des outils terminologiques : état des lieux et propositions. *Actes de la 7ème Conférence : Terminologie et Intelligence Artificielle (TIA'07)*, Sophia Antipolis, France, 10 2007.

[CO-145] Haifa Zargayouna, Victor Rosas et Sylvie Salotti. Shallow parsing of INEX Queries. *Comparative Evaluation of XML Information Retrieval Systems 5th International Workshop of the Initiative for the Evaluation of XML Retrieval, INEX 2006*, pages 284–293, Dagstuhl, Deutschland, 02 2007. LNCS, Volume Number 4518.

2006 [CO-144] Touria Aït El Mekki et Adeline Nazarenko. An application-oriented terminology evaluation : the case of back-of-the book indexes. Rute Costa, Fidélia Ibekwe-SanJuan, Susanne Lervad, Marie-Claude L'Homme, Adeline Nazarenko et Henrik Nilsson, éditeurs, *Workshop on Terminology design : quality criteria and evaluation methods (TermEval)*, pages 18–21, Genoa, Italie, 2006.

[CO-143] Laurent Alfandari, Jean-Luc Lemalade, Anass Nagih et Gérard Plateau. A mip flow model for crop-rotation planning in a sustainable development context. *EURO Congress*, Reykjavik, Iceland, 2006.

[CO-142] Nejla Amara-Hachmi et Amal El Fallah-Seghrouchni. Modélisation d'information contextuelle pour des agents mobiles sensibles au contexte. *Nouvelles technologies de la répartition (NOTERE'06)*, Toulouse, juin 2006.

[CO-141] Sophie Aubin, Julien Derivière, Thierry Hamon, Adeline Nazarenko, Thierry Poibeau et Davy Weissenbacher. ALVIS : A robust linguistic infrastructure for efficient semantic content analysis. *Symposium on Digital Semantic Content across Cultures*, Paris, France, 2006.

[CO-140] M. Badri, F. Boufarès, C.F. Ducateau et S. Hamdoun. Maintenance des entrepôts de données hétérogènes. *Proceedings of the 9th Maghrebian Conference on Software Engineering and Artificial Intelligence*, pages 152–158, 2006.

[CO-139] M. Badri, F. Boufarès, C.F. Ducateau et R. Nefoussi. Maintenance des vues matérialisées hétérogènes sous Oracle. *Actes de l'atelier Système d'Information Décisionnels, Informatique des Organisations et Systèmes d'Information et de Décision*, pages 8–14, 2006.

[CO-138] P. Baillot, U. Dal Lago et J.-Y. Moyen. On Quasi-interpretations, Blind Abstractions and Implicit Complexity. 8th International Workshop on Logic and Computational Complexity (LCC'06) (Satellite de LICS'06), Seattle, Prepublication HAL ccsd-00023668, août 2006.

[CO-137] P. Baillot et M. Pedicini. An Embedding of the BSS Model of Computation in Light Affine Lambda-Calculus. 8th International Workshop on Logic and Computational Complexity (LCC'06) (Satellite de LICS'06), Seattle, Prepublication HAL ccsd-00085547, août 2006.

[CO-136] Delphine Battistelli, Jean-Luc Minel et Sylviane R. Schwer. Quelques exemples d'utilisation des s-langages pour le traitement de la temporalité en linguistique, juin 2006.

[CO-135] S. Baudic et G. H. E. Duchamp. Modelling and reasoning on emotions : theories and clinical implications. *Complex Systems and Self-organization Modelling, CoSSOM'06 (Toulouse)*, 2006.

[CO-134] Didier El Baz, Moussa Elkihel, Laurent Gély et Gérard Plateau. Improved time and space complexity for kianfar's inequality rotation algorithm. *META'O6*, Hammamet, Tunisie, 2006.

[CO-133] Khalid Benabdeslem et Younès Bennani. Classification et visualisation des données d'usages d'Internet. *Atelier " Fouille du Web " 6ème EGC'06 (Extraction et Gestion des Connaissances*, pages 29–40, janvier 2006.

[CO-132] D. Berrabah, F. Boufarès, C.F. Ducateau et V. Heiwy. Sémantique des contraintes dans les diagrammes de classes UML : une expression en langage OCL. *Proceedings of the 9th Maghrebian Conference on Software Engineering and Artificial Intelligence*, pages 534–538, 2006.

[CO-131] Philippe Bessières, Robert Bossy, Alain-Pierre Manine, Erick Alphonse et Claire Nédellec. Getting the unknown from the known in bacteria, and the role of text mining. *Int. ECML'06 Workshop on Data and Text Mining for Integrative Biology*, 2006.

[CO-130] Amanda Bouffier et Thierry Poibeau. Segmentation et structuration de textes procéduraux pour l'aide à la modélisation de connaissances : le rôle de la structure visuelle. *Semaine de la Connaissance*, pages 195–205, Nantes, France, 2006.

[CO-129] Laure Bourgois, Alexandre Delteil et François Levy. Réparation de Web services composites. *Atelier Conception des systèmes d'information et services Web, INFORSID'06*, 2006. 12 pages, //sisw06.loria.fr/SISW$_$files/articles/BourgoisArticle.pdf.

[CO-128] Frederik Cailliau. Unifier la gestion de ressources linguistiques en contexte multilingue. *Actes de la 13ème conférence sur le Traitement Automatique des Langues Naturelles TALN'2006*, pages 455–461. Presses universitaires de Louvain, 2006.

[CO-127] Christophe Cérin, Heithem Abbes, Nabil Abdennadher et Mohamed Jemni. Architecture d'une grille de calcul tunisienne. *JTEA*, Tunisie Hammamet, 2006.

[CO-126] Julien Derivière et Thierry Hamon. Annotation linguistique de documents Web dans une architecture distribuée et adaptable. *The French Perl Workshop (FPW2006)*, Nov. 2006. Communication orale, ALVIS STREP project.

[CO-125] Fayçal Djerourou et Hachemi Bennaceur. Inférence d'inconsistances pour la résolution du Max-CSP. *Journées Francophones de Programmation par Contraintes*, pages 129–138, Alès, France, 2006.

[CO-124] Fayçal Djerourou et Hachemi Bennaceur. Infering inconsistencies to enhance branch and bound for Max-CSP. *Soft 2006 Workshop-Constraint Programming CP'06*, pages 47–62, Nantes, France, 2006.

[CO-123] Jérome Galy, Emmanuel Viennet et Maria Ruiz. Création des départements Télécommunications et Réseaux dans les IUT au Vénézuela. *Vème congrès France-Amérique latine, Coopération technologique et professionnelle*, novembre 2006.

[CO-122] R. Ghnemat, C. Bertelle et G.H.E. Duchamp. Self-organization simulation over geographical information system based on multi-agent platform. *Complex Systems and Self-organization Modelling, CoSSOM'06 (Toulouse)*, 2006.

[CO-121] Mourad Hakem et Franck Butelle. A bi-objective algorithm for scheduling parallel applications on heterogeneous systems subject to failures. *Proc. Rencontres Francophones du Parallélisme (RenPar)*, pages 25–35, 2006.

[CO-120] S. Hamdoun, F. Boufarès et M. Badri. Construction et maintenance des entrepôts de données hétérogènes. *Proceedings of the 9th Maghrebian Conference on Software Engineering and Artificial Intelligence*, pages 160–168, 2006.

[CO-119] H. Kadri-Dahmani, C. Bertelle, G.H.E. Duchamp et A. Osmani. The evolution process of geographical data base within self-organized topological propagation area. *Complex Systems and Self-organization Modelling, CoSSOM'06 (Toulouse)*, 2006.

[CO-118] Hager Karoui, Rushed Kanawati et Laure Petrucci. Cooperative CBR system for peer agent committee formation. *Proc. 5th Workshop on Agents and Peer-to-Peer Computing (AP2PC'06)*, pages 76–90, Hakodate, Japan, mai 2006.

[CO-117] Hager Karoui, Rushed Kanawati et Laure Petrucci. Une approche ràpc pour la formation de comité d'agents dans un système de recommandation égal à égal. *14ème Atelier de Raisonnement à Partir de Cas*, pages 63–74, Besançon, France, Mars 2006.

[CO-116] Daniel Kayser et Farid Nouioua. Une expérience de sémantique inférentielle. *Actes de la 13ème conférence sur le Traitement Automatique des Langues Naturelles TALN'2006*, pages 246–255. Presses Universitaires de Louvain, 2006.

[CO-115] M. I. Khemmoudj et H. Bennaceur. Inference de cliques pour la résolution de Max-CSP. *2e Journées Francophones de Programmation par Contraintes*, pages 229–238, Nîmes, France, 2006.

[CO-114] Jean Luc Lemalade, Laurent Alfandari, Anass Nagih et Gérard Plateau. Un modèle de planification des rotations culturales dans un contexte de lutte contre la déforestation à madagascar. *7e congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision*, Lille, 2006.

[CO-113] L. Létocart, A. Nagih, R. Ouaras et G. Plateau. Réoptimisation du sac-à-dos quadratique. *7e congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision*, Lille, France, 2006.

[CO-112] S. Mazouz et L. Petrucci. Modular discrete pseudo-state graphs for time Petri nets. *Workshop on Modelling of Objects, Components and Agents (MOCA'06, associated with PETRI NETS'06), Turku, Finland*, numéro 272 Universität Hamburg Bericht, pages 223–244, juin 2006.

[CO-111] Adeline Nazarenko, Claire Nédellec, Erick Alphonse, Sophie Aubin, Thierry Hamon et Alain-Pierre Manine. Semantic annotation in the Alvis project. *Int. Workshop on Intelligent Information Access (IIIA 2006)*, 2006.

[CO-110] Adeline Nazarenko, Claire Nédellec, Erik Alphonse, Sophie Aubin, Thierry Hamon et Alain-Pierre Manine. Semantic annotation in the ALVIS project. *Proceeding of IIIA-2006 : International Workshop on Intelligent Information Access*, Helsinki, Finlande, 2006.

[CO-109] Farid Nouioua et Pascal Nicolas. Using answer set programming in an inference-based approach to natural language semantics. Alexander Koller Johan Bos, éditeur, *Inference in Computational Semantics ICoS-5*, pages 77–86, Buxton, France, 2006.

[CO-108] Aomar Osmani et Khellaf Bouandas. DATTES : An algorithm for discovering patterns in temporal sequence. *Proceedings of ECAI temporal and Spatial reasoning Workshop*, 2006.

[CO-107] Thierry Poibeau. Linguistically grounded models of language change. Ron Sun, éditeur, *The 28th Annual Conference of the Cognitive Science Society (CogSci 2006)*, page 2587, Vancouver, Canada, 2006.

[CO-106] Sylviane R. Schwer. Le système des formes verbales de Marc Wilmet revisité dans une perspective néo-reichenbachienne, septembre 2006.

[CO-105] Emmanuel Viennet, Françoise Fogelman Soulié et Benoît Rognier. évaluation de techniques de traitement des refusés pour l'octroi de crédit. Société Française de Statistique, éditeur, *38èmes Journées de Statistique*, mai 2006.

[CO-104] Haifa Zargayouna et Nejla Amara. LiteMap : An ontology mapping approach for Mobile Agents' context-awareness. *OTM Workshops 2006 International Workshop on Context-Aware Mobile Systems*, pages 1934–1943, Montpellier, France, 2006. LNCS 4278.

2005 [CO-103] Touria Aït El Mekki et Adeline Nazarenko. Histoire et avenir des index de fin de livre. *Colloque international Indice, Index, Indexation*, Lille, France, 2005.

[CO-102] Touria Aït El Mekki et Adeline Nazarenko. ï¿½valuer le réseau terminologique d'un index de document. *Rencontres Terminologie et Intelligence Artificielle ; Atelier Evaluation des outils terminologiques*, Rouen, France, 2005.

[CO-101] Erick Alphonse, Sophie Aubin, Philippe Bessières, Gilles Bisson, Eric Gicquello, Thierry Hamon, Alain-Pierre Manine, Adeline Nazarenko et Claire Nédellec. Information extraction from MedLine on bacillus subtilis regulations. *BACELL 2005 European Workshop*, 2005.

[CO-100] Nejla Amara-Hachmi. A framework for building adaptive mobile agents. *Doctoral Mentoring Program at the 4th International Joint Conference on Autonomous agents and Multi-Agent Systems AAMAS 2005*, page 1369. ACM Press, juillet 2005.

[CO-99] Nejla Amara-Hachmi et Amal ElFallah-Seghrouchni. A framework for context-aware mobile agents. *4th Workshop on Ambient Intelligence - Agents for Ubiquitous Environments AmbiAgents05*, juillet 2005.

[CO-98] Nejla Amara-Hachmi et Amal ElFallah-Seghrouchni. GAMA : Architecture générique d'agents mobiles adaptables. *Plate-forme de l'Association Française pour l'Intelligence Artificielle - AFIA05, Conférence : Rencontres Jeunes Chercheurs en Intelligence Artificielle RJCIA'05*, juin 2005.

[CO-97] Nejla Amara-Hachmi et Amal ElFallah-Seghrouchni. Towards a generic architecture for self-adaptive mobile agents. *5th European Workshop on Adaptive Agents and MultiAgent Systems*, mars 2005.

[CO-96] Sujeevan Aseervatham et Aomar Osmani. Mining short sequential patterns for hepatitis type detection. *Discovery Challenge (ECML/PKDD)*, 2005.

[CO-95] M. Badri, F. Boufarès, C.F. Ducateau et F. Gargouri. Etat de l'art de la maintenance des entrepôts de données issus des systèmes d'information hétérogènes. *Actes des cinquièmes journées scientifiques des jeunes chercheurs en Génie Electrique et Informatique*, pages 13–18, 2005.

[CO-94] J. Billington, G. E. Gallasch et L. Petrucci. Transforming couloured Petri nets to counter systems for parametric verification : A stop-and-wait protocol case study. *Proc. 2nd workshop on MOdel-based Methodologies for Pervasive and Embedded Software (MOMPES'05, satellite of ACSD'05), Rennes, France*, volume 39, pages 37–55. TUCS general publication, juin 2005. ISBN 952-12-1556-9.

[CO-93] Tadeusz Czachorski, Ferhan Pekergin et Bartlomiej Szady. Forming optical packets, a diffusion approximation and simulation model. *Proceedings of EuroNGI (European New Generation Internet) 2nd Workshop IA.8.1*, Portugal Aveiro, 2005.

[CO-92] Tadeusz Czachorski, Ferhan Pekergin et Bartlomiej Szady. Forming optical packets, a diffusion approximation model. *Proceedings of 12th Polish Teletraffic Symposium PSRT*, Poland Poznan, 2005.

[CO-91] Lahcène Dehni, Younès Bennani et Francine Krief. Low energy adaptive connectionist clustering for wireless sensor networks. *IEEE/IFIP MATA 2005 : 2nd International Workshop on Mobility Aware Technologies and Applications*, volume 3744 de *Lecture Notes in Computer Science*, pages 405–415. Springer, octobre 2005.

[CO-90] Lahcène Dehni, Younès Bennani et Francine Krief. Une approche neuronale adaptative de routage minimisant la consommation d'énergie dans les réseaux de capteurs. *colloque GRES 2005 : Gestion de REseaux et de Services*, pages 115–127, février 2005.

[CO-89] Thierry Delbecque, Pierre Zweigenbaum, Jean-François Berroyer et Thierry Poibeau. Le système STIM/LIPN à EQueR 2004, tâche médicale. *Atelier EQueR (conférence TALN'2005)*, pages 89–94, Dourdan, France, 2005.

[CO-88] Julien Derivière, Thierry Hamon, Adeline Nazarenko et Guillaume Vauvert. Développement d'une plate-forme d'enrichissement des documents textuels : l'expérience du projet ALVIS. *Journée ATALA "Articuler les traitements sur corpus"*, Paris, France, 2005.

[CO-87] Sylvie Després et Sylvie Szulman. Construction d'une ontologie du droit documentaire. *IC'2005*, pages 85–96. Presses Universitaires de Grenoble, 2005.

[CO-86] Sébastien Guérif, Younès Bennani et Eric Janvier. μ-SOM : Weighting features during clustering. *Proceedings of the 5th Workshop On Self-Organizing Maps (WSOM'05)*, pages 397–404, Paris, France, septembre 2005.

[CO-85] Mourad Hakem et Franck Butelle. Dynamic critical path scheduling parallel programs onto multiprocessors. *Proc. of the 19th ACM/IEEE International Parallel & Distributed Processing Symposium – Advances in Parallel and Distributed Computational Models*, page 203b. IEEE Computer Society Press, 2005. (10 pages sur CDROM).

[CO-84] Thierry Hamon. Indexer les documents spécialisés : les ressources terminologiques contrôlées sont-elles suffisantes ? *Actes de Terminologie et Intelligence Artificielle 2005*, pages 71–82. PUR, 2005.

[CO-83] Zeina Jrad. Spécification d'une interface intelligente pour la prochaine génération de réseaux IP. *Journées Scientifiques DNAC 2005. Les réseaux IP : synthèse et perspective, controle, gestion, mobilité et sécurité.*, avril 2005.

[CO-82] Zeina Jrad, Francine Krief et Sébastien Guérif. Une interface utilisateur pour faciliter l'accès à l'internet nouvelle génération. *Journées doctorales Informatique et Réseau (JDIR 2005)*, Troyes, France, décembre 2005.

[CO-81] Rushed Kanawati. Un système coopératif pour le tri des résultats de moteurs de recherche. *13ième atelier sur le raisonnement à partir de cas - Plate-forme AFIA 2005*, mai 2005.

[CO-80] Hager Karoui. Agents RàPC pour la gestion coopérative de bases bibliographiques personelles. *Atelier raisonnement à partir de cas, plate-forme AFIA*, pages 86–97, Nice, mai 2005.

[CO-79] M. I. Khemmoudj et H. Bennaceur. Using constraint programming for scheduling of EDF nuclear power plant outages. *1st International Workshop on Optimization Frameworks for Industrial Applications*. Paris, France, 2005. (Poster).

[CO-78] M. I. Khemmoudj, H. Bennaceur et Marc Porcheron. Application de la programmation logique par contraintes au problème du placement des arrêts des réacteurs nucléaires d'EDF. *7th International Symposium on Programming an Systems*, pages 177–186, Alger, 2005.

[CO-77] M. I. Khemmoudj, F. Djerourou et H. Bennaceur. Résolution de Max-CSP par recherche locale guidée par la relaxation lagrangienne. *1th Journées Francophones de Programmation par Contraintes*, pages 305–314, Lens, France, 2005.

[CO-76] M. I. Khemmoudj et M. Porcheron. Using constraint programming and local search for scheduling of edf nuclear power plant outages. 1st International Workshop on Combination of metaheuristic and local search with Constraint Programming, 2005. (Poster).

[CO-75] C. Lakos et L. Petrucci. Distributed and modular state space exploration for timed Petri nets. *Proc. Workshop on Practical Use of Coloured Petri Nets, Aarhus, Denmark*, pages 191–210, octobre 2005. Proceedings published as Report DAIMI-PB 576, Aarhus, DK.

[CO-74] Dorra Maghri et Nejla Amara-Hachmi. Intégration d'un agent mobile dans un système multi-agents hierarchique. *Manifestation des Jeunes Chercheurs francophones dans les domaines des STIC MajecSTIC*, Rennes, novembre 2005.

[CO-73] Nicolas Marcos, David De Almeida, Daniel Gauyacq, Alexandre Quivet, Anass Nagih, Laurent Alfandari et Gérard Plateau. Un outil générique d'aide à la décision pour la gestion des locomotives fret à la SNCF. *6e congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision*, pages 295–312, Tours, 2005.

[CO-72] Hassine Moungla et Francine Krief. A probabilistic heuristic for conflict detection in policy based management of Diffserv networks. *Proceedings of IEEE/IFIP MATA 2005 : 2nd International Workshop on Mobility Aware Technologies and Applications*, octobre 2005.

[CO-71] Goritsa Ninova, Adeline Nazarenko, Thierry Hamon et Sylvie Szulman. Comment mesurer la couverture d'une ressource terminologique pour un corpus ? *Actes de la 12ème conférence sur le Traitement Automatique des Langues Naturelles TALN'2005*, pages 293–302. ATALA, 2005.

[CO-70] Farid Nouioua. Norm based causal reasoning in textual corpus. *Proceedings of the Sixth International Workshop on Computational Semantics IWCS-6*, pages 396–400, Tilburg, France, 2005.

[CO-69] Farid Nouioua. Raisonnement stratifié à base de normes pour inférer les causes dans un corpus textuel. *The Seventh International Symposium On Programming and Systems*, pages 81–92. USTHB d'Alger, 2005.

[CO-68] Farid Nouioua. Why did the accident happen ? A norm-based reasoning approach. *Logical Aspects of Computational Linguistics, student session*, pages 31–34. Université de Bordeaux, 2005.

[CO-67] Alexandre Pauchet, Nathalie Chaignaud et Amal El Fallah-Segrouchni. Un modèle cognitif de l'interaction pour agents logiciels. *3èmes Journées Francophones des Modèles Formels de l'Interaction (MFI'05)*, mai 2005.

[CO-66] Thierry Poibeau. Architectures articulant des représentations hétérogènes : l'exemple de Gate. B. Habert, A. Salem et S. Heiden, éditeurs, *Journée ATALA (Association pour le Traitement Automatique des Langues)*, Paris, France, 2005. Atala. www.limsi.fr/\-Individu/\-habert/\-Projets/\-JourneeATALA05ArticulerLesTraitements/\-Poibeau.ppt.

[CO-65] Thierry Poibeau. Parcours interprétatifs et terminologie. Yannick Toussaint, éditeur, *Actes du colloque Terminologie et Intelligence Artificielle*, pages 12–25. Université de Rouen, 2005.

[CO-64] Thierry Poibeau. Sur le statut référentiel des entités nommées. Michèle Jardino, éditeur, *Conférence Traitement Automatique des Langues 2005*, pages 173–183, Dourdan, France, 2005.

[CO-63] Davy Weissenbacher. A Bayesian Network for the resolution of non-anaphoric pronoun It. *Bayesian Methods for Natural Language Processing Workshop at NIPS*, Canada, 2005.

2004 [CO-62] Laurent Alfandari, Marc Demange et Jérôme Monnot. Boolean inverse optimization problems. *EURO XX Congress*, Rhodos, Grèce, 2004.

[CO-61] Erick Alphonse, Sophie Aubin, Philippe Bessières, Gilles Bisson, Thierry Hamon, Sandrine Lagarrigue, Alain-Pierre Manine, Adeline Nazarenko, Claire Nédellec, Mohamed Ould Abdel Vetah, Thierry Poibeau et Davy Weissenbacher. Extraction d'information appliquée au domaine biomédical : apprentissage et traitement automatique de la langue. *Actes de la Conférence Internationale sur la Fouille de textes (CIFT'04)*, pages 7–19, La Rochelle, France, 2004.

[CO-60] Erick Alphonse, Sophie Aubin, Philippe Bessières, Gilles Bisson, Thierry Hamon, Sandrine Lagarrigue, Adeline Nazarenko, Alain-Pierre Manine, Claire Nédellec, Mohamed Ould Abdel Vetah, Thierry Poibeau et Davy Weissenbacher. Event-based information extraction for the biomedical domain : the Caderige project. Nigel Collier, Patrick Ruch et Adeline Nazarenko, éditeurs, *Proceedings of the International Joint Workshop on Natural Language Processing in Biomedicine and Its Applications (COLING'04)*, pages 43–49, Geneva, Suisse, 2004.

[CO-59] S. Bardin et L. Petrucci. COAST : des réseaux de Petri à la planification assistée. *Proc. 6ème Conférence sur les Approches Formelles dans l'Assistance au Développement de Logiciels (AFADL'2004), Besançon, France*, pages 285–298, juin 2004.

[CO-58] S. Bardin et L. Petrucci. From PNML to counter systems for accelerating Petri nets with FAST. *Proc. of the Workshop on Interchange Formats for Petri Nets (at ICATPN 2004)*, pages 26–40, juin 2004.

[CO-57] D. Berrabah, F. Boufarès, C.F. Ducateau, F. Gargouri et V. Heiwy. Les conflits entre les contraintes dans les schémas conceptuels de bases de données UML-EER. *Proceedings of the 8th Maghrebian Conference on Software Engineering and Artificial Intelligence*, pages 437–448, 2004.

[CO-56] Jean Charlet, Touria Aït El Mekki, Didier Bourigault, Adeline Nazarenko, Régine Teulier et Baruk Toledano. CEDERILIC : constitution d'un livre et d'un index numériques. *Actes de 7ème conférence Internationale sur le Document numérique (CIDE'04)*, pages 187–204. Semaine du document numérique, 2004.

[CO-55] C. Choppy et M. Heisel. Une approche à base de "patrons" pour la spécification et le développement de systèmes d'information. *Proceedings Approches Formelles dans l'Assistance au Développement de Logiciels - AFADL'2004*, pages 61–76, 2004.

[CO-54] C. Choppy et L. Petrucci. Towards a methodology for modelling with Petri nets. *Proc. Workshop on Practical Use of Coloured Petri Nets, Aarhus, Denmark*, pages 39–56, octobre 2004. Proceedings published as Report DAIMI-PB 570, Aarhus, DK.

[CO-53] C. Choppy et G. Reggio. A UML-Based Method for the Commanded Behaviour Frame. K. Cox, J.G. Hall et L. Rapanotti, éditeurs, *Proc. of the 1st International Workshop on Advances and Applications of Problem Frames (IWAAPF 2004)*, pages 27–34. An ICSE 2004 workshop, IEEE, 2004.

[CO-52] Louisette Emirkanian, Christophe Fouqueré et Fabrice Issac. Corpus issus du web : analyse des pertinences thématiques et informationnelles. *Actes du Congrès JADT2004*, mars 2004.

[CO-51] Françoise Gayral, Daniel Kayser et François Lévy. The role of time in considering collections. *Journées de Sémantique et Modélisation*, pages 30–34, Lyon, France, 2004.

[CO-50] Natalia Grabar et Thierry Hamon. Repérage de relations sémantiques entre termes : sur la piste de la morphologie. *Actes de Journées francophones d'Ingénierie des Connaissances (IC'2004)*, pages 237–248. Presses Universitaires de Grenoble, 2004.

[CO-49] Alejandro Guerra-Hernandez, Amal El Fallah-Seghrouchni et Henry Soldano. Learning in BDI multi-agent systems. *Fourth International Workshop on Computational Logic in Multi-Agent Systems (CLIMA IV)*, volume 3259 de *Lecture Notes in Computer Science*, pages 218–233. Springer, 2004.

[CO-48] Thierry Hamon. Plateforme d'annotation terminologique de corpus des résumés d'articles de biologie. *Actes des 15ièmes journées francophones d'Ingénierie des Connaissances (IC'2004)*, Lyon, France, 2004. Presses universitaires de Grenoble. Poster.

[CO-47] H. Karoui, R. Kanawati et L. Petrucci. An intelligent peer-to-peer multi-agent system for collaborative management of bibliographic databases. *Proc. Workshop on Case Based Reasoning, Cambridge, England*, décembre 2004.

[CO-46] Hassine Moungla. La différentiation de service dans les réseaux MPLS/GMPLS. *DNAC 2004*, février 2004.

[CO-45] Adeline Nazarenko et Touria Aït El Mekki. Une mesure de pertinence pour le tri de l'information dans un index de fin de livre. *Actes de la 11ème conférence sur le Traitement Automatique des Langues Naturelles TALN'2004*, pages 1–10. ATALA, 2004.

[CO-44] Aomar Osmani. Cyclic and linear intervals for reasoning about repetitions in time. *Proceedings of AAAI Workshop*, pages 1–6, 2004.

[CO-43] Thierry Poibeau. Pré-analyse de corpus. G. Purnelle, C. Fairon et A. Dister, éditeurs, *Journées d'Analyse des Données Textuelles (JADT 2004)*, pages 897–904, Louvain, Belgique, 2004. Presses Universitaires de Louvain.

[CO-42] Thierry Poibeau et Bénédicte Goujon. Acquisition semi-automatique de grammaires de commande. I. Marlien et B. Bel, éditeurs, *Journées d'etude de la Parole (JEP'2004)*, pages 405–409, Fès, Maroc, 2004. Association Francophone de la Communication Parlée. Collaboration avec Thales Recherche et Technologie.

[CO-41] Sylviane R. Schwer. Relations temporelles qualitatives et langages formels. *Actes des 11èmes Journées de Rochebrune, Rencontres interdisciplinaires sur les systèmes complexes naturels et artificiels, Le temps dans les systèmes complexes naturels et artificiels*, pages 329–340, janvier 2004.

[CO-40] A. Suna, A El Fallah-Seghrouchni, C. Fouqueré et P. Baillot. Mobile Multi-Agent Systems : A Programming Language and its Semantics. *3rd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2004)*, pages 1386–1387. IEEE Computer Society, 2004. Poster.

[CO-39] Sylvie Szulman et Brigitte Biébow. OWL et Terminae. *IC'2004 - 15ème journées francophones d'Ingénierie des Connaissances*, pages 41–52. Presses Universitaires de Grenoble, 2004.

[CO-38] Véronique Ventos, Henry Soldano et Thibaut Lamadon. Treillis de Galois alpha. *Conférence d'Apprentissage (CAP'04)*, pages 175–190, Montpellier, 2004. Presses Universitaires de Grenoble.

[CO-37] Davy Weissenbacher. La relation de synonymie en génomique. *Actes des Rencontres des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RECITAL)*, pages 298–303. ATALA, 2004.

[CO-36] Haifa Zargayouna et Sylvie Salotti. Mesure de similarité dans une ontologie pour l'indexation sémantique de documents XML. *15ème journées francophones d'Ingénierie des Connaissances (IC'2004)*, pages 249–260. Presses Universitaires de Grenoble, 2004.

[CO-35] Haifa Zargayouna et Sylvie Salotti. Mesure de similarité sémantique pour l'indexation de documents semi-structurés. *12ème Atelier de Raisonnement à Partir de Cas*, pages 19–29, Villetaneuse, France, 2004. LIPN.

[CO-34] Haifa Zargayouna et Sylvie Salotti. Semindex : a model of semantic indexing on XML documents. *26th European Conference on Information Retrieval (ECIR'2004)*, 2004. Poster.

[CO-33] Farida Zehraoui et Younès Bennani. SOM-ART : Incorporation des propriétés de plasticité et de stabilité dans une carte auto-organisatrice. *Atelier FDC : Fouille de Données Complexes dans un processus d'extraction de connaissances (EGC 2004)*, pages 169–180, janvier 2004.

[CO-32] Farida Zehraoui, Rushed Kanawati et Sylvie Salotti. Case Representation, Retrieval and Reuse in the Hybrid System for Sequence Processing CASEP2. Rushed Kanawati et Sylvie Salotti, éditeurs, *Second International workshop on applying case-base reasoning on time series prediction*, pages 339–350, September 2004.

2003 [CO-31] Touria Aït El Mekki et Adeline Nazarenko. Le réseau terminologique, un élément central pour la construction d'index de document. *Actes des 5èmes rencontres Terminologie et Intelligence Artificielle (TIA'03)*, pages 1–10, Strasbourg, France, 2003.

[CO-30] Touria Aït El Mekki et Adeline Nazarenko. Semi-automatic method for the construction of a document index. *Workshop :"Applied Natural Language Processing, possible applications for the Semantic Web" associated with the EUROLAN-2003 summer school*, Bucarest, Roumanie, 2003.

[CO-29] Laurent Alfandari. Modèles de localisation pour l'aide multicritère à la décision. *5e Congrès de la Société Francaise de Recherche Opérationnelle et d'Aide à la Décision*, Avignon, 2003.

[CO-28] Rafael Castro De Andrade, Abdel Lisser, Nelson Maculan et Gérard Plateau. Synthèse de réseaux multiservice à demande incertaine. *5e congrès de la Société Francaise de Recherche Opérationnelle et d'Aide à la Décision*, Avignon, 2003.

[CO-27] Nathalie Aussenac-Gilles, Brigitte Biébow et Sylvie Szulman. D'une méthode à un guide pratique de modélisation de connaissances à partir de textes. *Actes des 5èmes rencontres Terminologie et Intelligence Artificielle (TIA'03)*, pages 41–53. LIIA - ENSAIS, 2003.

[CO-26] Badr Benmammar et Francine Krief. La technologie agent et les réseaux sans fil. *Congrès Des Nouvelles Architectures pour les Communications. DNAC 2003*, octobre 2003.

[CO-25] Lélia Blin et Franck Butelle. The first approximated distributed algorithm for the minimum degree spanning tree problem on general graph. *proc. IEEE International Parallel & Distributed Processing Symposium – Advances in Parallel and Distributed Computational Models*, page 161. IEEE Computer Society Press, 2003. PDF de 10 pages sur CDROM.

[CO-24] F. Boufarès, D. Berrabah, C.F. Ducateau et V. Heiwy. Etude de la cohérence des contraintes de multiplicité dans un schéma conceptuel de BD avec UML. *Actes du Colloque National de la Recherche en IUT*, pages 375–382, 2003.

[CO-23] F. Boufarès et S. Hamdoun. Construction et auto-maintenance d'un web warehouse. *Actes du Colloque National de la Recherche en IUT*, pages 366–374, 2003.

[CO-22] M-C. Costa, L. Létocart et F. Roupin. Multicoupe minimale et multiflot entier maximal dans un anneau. *Les Journées Graphes, Réseaux et Modélisation*, Paris, France, 2003.

[CO-21] M-C. Costa, L. Létocart et F. Roupin. Multicoupes minimales et multiflots maximaux en nombres entiers dans les anneaux. *5e congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision*, Avignon, France, 2003.

[CO-20] Fayçal Djerourou et Hachemi Bennaceur. Diversification de la recherche locale dans la résolution du VCSP avec pénalisation de couples. *5e congrès de la Société Francaise de Recherche Opérationnelle et d'Aide à la décision*, pages 399–400, Avignon, France, 2003.

[CO-19] Fayçal Djerourou et Hachemi Bennaceur. Une méthode exacte de résolution du Max-CSP. *5e congrès de la Société Francaise de Recherche Opérationnelle et d'Aide à la décision*, pages 332–333, Avignon, France, 2003.

[CO-18] Christophe Fouqueré. Process calculus and mobile code, octobre 2003. 2ème école franco-mexicaine sur les systèmes répartis coopératifs.

[CO-17] Françoise Gayral, Daniel Kayser et François Lévy. Collections évolutives et partitions. *Colloque "La partition en langue et en discours"*, France, 2003.

[CO-16] S. Jungjariyanonn et Sylviane R. Schwer. Extended trivalent computations. *Proceedings of KRR1 (Spatial and Temporal Reasoning) of the 18th International Joint Conference on Artificial Intelligence*, août 2003.

[CO-15] Gilles Klein et Francine Krief. Mobile agents for dynamic sla negotiation. *International Workshop on Mobile Agents for Telecommunication Applications. MATA 2003*, Lecture Notes in Computer Science. Springer, octobre 2003.

[CO-14] Francine Krief et Zeina Jrad. An intelligent environment for dynamic negotiation, provisioning and control of QoS in IP networks. *Colloque Gestion de Réseau et de Service. GRES 2003*, février 2003.

[CO-13] K. Lellahi et A. Zamulin. Algebraic semantics of statements of sequential java. *proc. ECOOP Workshop on Formal Techniques for Java Programs (FTFJP-2003), Darmstadt Germany*, juin 2003.

[CO-12] Yassir Miladi, Anass Nagih et François Soumis. Problème de plus court chemin avec contraintes de ressources à l'intérieur d'un algorithme de génération de colonnes. *Colloque des sciences mathématiques du Québec*, Moncton, Canada, 2003.

[CO-11] Anass Nagih et François Soumis. Nodal aggregation of resource constraints in a shortest path problem. *5e Congrès de la Société Francaise de Recherche Opérationnelle et d'Aide à la Décision*, Avignon, 2003.

[CO-10] Adeline Nazarenko et Touria Aït El Mekki. Une mesure de pertinence pour le tri de l'information dans un document. *Atelier Acquisition, apprentissage et exploitation de connaissances sémantiques pour l'accès au contenu textuel, Plateforme AFIA*, Laval, France, 2003.

[CO-9] Alexandre Pauchet, Nathalie Chaignaud et Amal El Fallah-Seghrouchni. Résolution coopérative de problèmes intégrant planification et interaction. *Journées Francophones sur les Systèmes Multi-Agents*, pages 309–313, 2003.

[CO-8] Alexandre Pauchet, Nathalie Chaignaud et Amal El Fallah-Segrouchni. Modelling planning and interaction in the framework of human cooperative problem solving. *AAMAS Workshop on Human and Multi-Agent Systems*, pages 31–38, juillet 2003.

[CO-7] Sylviane R. Schwer. Temporal reasoning without transitive table. *Proceedings of KRR1 (Spatial and Temporal Reasoning) of the 18th International Joint Conference on Artificial Intelligence*, août 2003.

[CO-6] Olivier Sigaud et Pierre Gérard. Apprentissage par renforcement indirect dans les sytèmes de classeurs. *Actes des journèes PDMIA*, 2003.

[CO-5] Babacar Thiongane, Anass Nagih et Gérard Plateau. Techniques de réoptimisation dans une séquence d'instances du problème du sac à dos unidimensionnel 0-1. *5e congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision*, Avignon, 2003.

[CO-4] Haifa Zargayouna et Sylvie Salotti. Indexation sémantique de documents XML. *Extraction et gestion des connaissances (EGC'2003). Atelier : Fouille de données et recherche d'informations dans des bases de données multimédia semi-structurées*, 2003. Poster.

2002 [CO-3] Christophe Fouqueré et Fabrice Issac. Constitution de corpus du web et analyse informationnelle : premiers résultats. Colloque TALN, Corpus et Web, MSH Paris Nord, 26 et 27 novembre 2002.

[CO-2] K. Lellahi et A. Zamulin. Towards an algebraic formalization of sequential java-like programs with implicit states. *Proc. ECOOP Workshop on Formal Techniques for Java Programs (FTFJP-2002), Malaga, Spain*, pages 131–132. Computing Science Department, University of Nijmegen, technical report No. NIII-R0204, juin 2002.

[CO-1*] L. Jaff, Hatem Hadj-Kacem, G. H.E. Duchamp et C. Bertelle. Moderate growth time series for dynamic combinatorics modelisation. submitted to the *International Journal of Electronic Commerce*.

Direction d'ouvrages (DO)

Livres

2008 [LI-8*] F. Kordon, L. Pautet et L. Petrucci. *Systèmes répartis en action : de l'embarqué au large échelle*. Hermès, 2008.

2006 [LI-7] Younès Bennani, éditeur. *Apprentissage Connexionniste*. Hermès, 2006, ISBN 2-7462-1337-0.

[LI-6] S. Haddad, F. Kordon et L. Petrucci, éditeurs. *Méthodes formelles pour les systèmes répartis et coopératifs*. Hermès, octobre 2006, ISBN 2-7462-1447-4.

2005 [LI-5] Benoît Charroux, Aomar Osmani et Yann Thierry-Mieg. *UML 2*. Pearson, juillet 2005, ISBN 2-7440-7124-2.

2004 [LI-4] Daniel Kayser, Adina Magda Florea et Stefan Gheorghe Pentiuc. *Intelligence artificielle et agents intelligents*. Printech, 2004, ISBN 973-652-976.

2003 [LI-3] Cyril Banderier et Christian Krattenthaler. *Discrete Random Walks : Theory and applications in combinatorics, computational biology, computer science, probability theory, and statistical physics*. Discrete Mathematics & Theoretical Computer Science, 2003.

[LI-2] Thierry Poibeau. *Extraction automatique d'information : Du texte brut au Web sémantique*. Lavoisier, 2003, ISBN 2-7462-0610-2.

[LI-1] Louise Travé-Massuyès et Philippe Dague, éditeurs. *Modèles et raisonnements qualitatifs*. IC2. Hermès, octobre 2003, ISBN 2-7462-0744-3.

Chapitres de livre

2008 [CL-36*] E. Badouel, P. Darondeau et L. Petrucci. *Synthesis of Distributed Petri Nets from Coordinated Automata*, M. Mukund, éditeur, *Perspectives in Concurrency Theory*. décembre 2008.

[CL-35*] Sylvie Szulman, Nathalie Aussenac-Gilles et Sylvie Despres. The Terminae Method and Platform for Ontology Engineering from Texts. *Bridging the Gap between Text and Knowledge - Selected Contributions to Ontology Learning and Population from Text*. IOS press, 2008.

2007 [CL-34] Rushed Kanawati. *Agents RàPC coopératifs pour la recherche d'information*, Jean Renaud, Brigitte Chebel Morello, Béatrice Fuchs et Jean Lieber, éditeurs, *Raisonnement à partir de cas 2, surveillance, diagnostic et maintenance*, volume 2 de *Informatique et système d'information*, Chapitre 6, pages 177–195. Hermès, 2007.

[CL-33] Maria Malek et Rushed Kanawati. *Case-based Reasoning in Knowledge Discovery and Data Mining*, Sankar K. Pal, David W. Aha et Kalyan M. Gupta, éditeurs, *Case-Based Reasoning in Knowledge Discovery and Data Mining*, Chapitre CBR-based sequence analysis : An overview. Wiely, 2007.

[CL-32] Jérôme Monnot et Sophie Toulouse. *Complexity and approximation results for bounded-size paths packing problems*, *Combinatorial Optimization and Theoretical Computer Science : Interfaces and Perspectives*, Chapitre 18. ISTE (V. Paschos ed.), 2007.

[CL-31] Jérôme Monnot et Sophie Toulouse. *Le voyageur de commerce et ses variations : un tour d'horizon de ses résolutions*, *Optimisation Combinatoire 5 : problèmes paradigmatiques et problématiques nouvelles*, Chapitre 2. Hermès (V. Paschos ed.), 2007.

[CL-30] Adeline Nazarenko et Touria Aït El Mekki. Building back-of-the-book indexes. Fidelia Ibekwe SanJuan, Anne Condamines et Maria Teresa Cabré, éditeurs, *Application-Driven Terminology Engineering*, Current Topics 2, pages 179–201. John Benjamins, 2007.

[CL-29] Gérard Plateau et Anass Nagih. *Le problème du sac à dos en variables 0-1*, *Optimisation combinatoire 4 : problèmes paradigmatiques*, Chapitre 6, pages 163–190. Hermès (V. Paschos ed.), Paris, 2007.

[CL-28] Antoine Rozenknop et Jean-Cédric Chappelier. Syntactic Processing. Martin Rajman et Vincenzo Pallota, éditeurs, *Speech and Language Engineering*, pages 177–204. Presses Polytechniques et Universitaires Romandes (CH)/ Marcel Dekker (US), 2007.

[CL-27] Farida Zehraoui, Rushed Kanawati et Sylvie Salotti. *CASEP2 : Système Hybride pour le traitement de séquences de navigation sur le Web*, Jean Renaud, Brigitte Cheble-Morello, Bétarice Fuchs et Jean Lieber, éditeurs, *Raisonnement à Partir de Cas 2 : surveillance, diagnostic et maintenance*, volume 2 de *Informatique et système d'information*, pages 151–176. Hermès, 2007.

2006 [CL-26] Laurent Alfandari et Anass Nagih. *Optimisation de la construction de rotations équipages en transport aérien*, *Optimisation combinatoire 3 : applications*, Chapitre 1. Hermès (V. Paschos ed.), Paris, 2006.

[CL-25] Younès Bennani. *Outils de simulation des systèmes connexionnistes*, Chapitre 11, pages 343–358. Volume 1 de Bennani [LI-7], 2006, ISBN 2-7462-1337-0.

[CL-24] Younès Bennani. *Réseaux à dictionnaires : Learning Vector Quantization*, Chapitre 5, pages 123–140. Volume 1 de Bennani [LI-7], 2006, ISBN 2-7462-1337-0.

[CL-23] Younès Bennani. Réseaux de neurones artificiels. J. Akoka et I. Wattiau, éditeurs, *Encyclopédie de l'informatique et systèmes d'information*, volume Section Intelligence Artificielle, pages 1217–1231. Editions Vuibert, 2006.

[CL-22] Younès Bennani. *Séparateurs connexionnistes linéaires*, Chapitre 1, pages 21–46. Volume 1 de Bennani [LI-7], 2006, ISBN 2-7462-1337-0.

[CL-21] C. Choppy et L. Petrucci. *Démarches de spécification*, Chapitre 2. Haddad et al. [LI-6], octobre 2006, ISBN 2-7462-1447-4.

[CL-20] H. Kadri-Dahmani, C. Bertelle, G.H.E. Duchamp et A. Osmani. *Consistent Updating of Geographical DataBase as Emergent Property over Influence System*, M. A. Aziz-Alaoui et C. Bertelle, éditeurs, *Emergent Properties in Natural and Artificial Dynamical Systems*. Understanding Complex Systems. Springer, 2006, ISBN 978-3-540-34822-1.

[CL-19] Rushed Kanawati. Raisonnement à partir de cas : principes, méthodes et applications industrielles. Jean Renaud, Béatrice Fuchs et Jean Lieber, éditeurs, *Raisonnement à Partir de Cas : principes, méthodes et applications industrielles*, Chapitre Agents RàPC coopératifs pour la recherche d'information sur le Web. Hermès, 2006.

[CL-18] Hassine Moungla. *L'Auto-configuration dans la gestion autonome, l'Autonomic communication.* Hermès, 2006.

[CL-17] Adeline Nazarenko. *Le point sur l'état actuel des connaissances en Traitement Automatique des Langues*, Gérard Sabah, éditeur, *Compréhension des langues et interaction*, pages 31–70. Hermès, 2006.

[CL-16] L. Petrucci. *Panorama des modèles et langages de spécification*, Chapitre 1. Haddad et al. [LI-6], octobre 2006, ISBN 2-7462-1447-4.

[CL-15] Nadia Pisanti, Henry Soldano, Mathilde Carpentier et Joël Pothier. Implicit and explicit representation of approximated motifs. C. S. Iliopoulos, K. Park et K. Steinhoefel, éditeurs, *Algorithms in Bioinformatics*, pages 1–15. College Press, 2006.

[CL-14] Emmanuel Viennet. *Réseaux à Fonctions de Base Radiales*, Chapitre 4, pages 105–122. Volume 1 de Bennani [LI-7], 2006, ISBN 2-7462-1337-0.

[CL-13] Farida Zehraoui, Rushed Kanawati et Sylvie Salotti. Raisonnement à partir de cas : principes, méthodes et applications industrielles. Jean Renaud, Béatrice Fuchs et Jean Lieber, éditeurs, *Raisonnement à Partir de Cas : principes, méthodes et applications industrielles*, Chapitre CASEP2 : Système hybride pour le traitement des séquences. Hermès, 2006.

2005 [CL-12] Brigitte Biébow, Sylvie Szulman et Nathalie Aussenac-Gilles. *Modélisation du domaine par une méthode fondée sur l'analyse de corpus*, R. Teulier ; J. Charlet ; P. Tchounikine, éditeur, *Ingénierie des connaissances*, pages 49–71. L'Harmattan, 2005.

[CL-11] Françoise Gayral, Daniel Kayser et François Lévy. *Challenging the Principle of Compositionality in Interpreting Natural Language Texts*, E. Machery, M. Werning et G. Schurz, éditeurs, *The Compositionality of Meaning and Content*, volume 2, pages 83–105. Ontos Verlag, 2005.

[CL-10] Zeina Jrad. *Les Plates-formes Multi-agents*, *Intelligence artificielle et réseaux*. Hermès, 2005.

[CL-9] Hassine Moungla. *L'Intelligence artificielle et la supervision des réseaux de télécommunication*, *Intelligence artificielle et réseaux*. Hermès, 2005.

[CL-8] Adeline Nazarenko. *Sur quelle sémantique reposent les méthodes automatiques d'accès au contenu textuel ?*, Anne Condamines, éditeur, *Sémantique et corpus*, pages 211–244. Traité IC2, série Cognition et traitement de l'information. Lavoisier, 2005.

[CL-7] Claire Nédellec et Adeline Nazarenko. *Ontology and Information Extraction : a necessary symbiosis*, Bernardo Magnini Paul Buitelaar, Philipp Cimiano, éditeur, *Ontology Design and Population*, pages 155–170. IOS Press, 2005.

[CL-6] Thierry Poibeau. *Content annotation for the semantic Web*, Spiros Sirmakessis, éditeur, *Knowledge Mining*, pages 250–262. Springer, 2005. ISBN 3-540-25070-0.

[CL-5] Thierry Poibeau. *Una metodologica per l'annotazione e l'estrazione d'informazione*, S. Bolasco, A. Canzonetti et F.M. Capo, éditeurs, *Knowledge Mining : Uno strumento strategico per imprese e institutioni*, pages 37–45. ASU, 2005. ISBN 88-7975-341-X.

2004 [CL-4] Francine Krief. *Application des Agents dans la Gestion par Politique*, *Intelligence artificielle et réseaux*. Hermès, 2004.

[CL-3] Adeline Nazarenko et Thierry Poibeau. *L'évaluation des systèmes d'analyse et de compréhension de textes*, Stéphane Chaudiron, éditeur, *L'évaluation des systèmes de traitement de l'information*, pages 127–148. Hermès, 2004. ISBN 2-7462-0862-8.

2003 [CL-2] Laurent Alfandari, Agnès Plateau et Pierre Tolla. *A path-relinking algorithm for the Generalized Assignment Problem*, *Metaheuristics : Computer Decision Making*, pages 1–18. Combinatorial Optimization Book Series (D-.Z. Du and P.M. Pardalos series ed.). Kluwer Academic Publishers (M.G.C Resende, J.P de Sousa ed.), Boston, 2003.

[CL-1] Hakima Kadri-Dahmani et Aomar Osmani. *Updating Data in GIS : How to maintain Database Consistency*, *Enterprise Information Systems IV book*, pages 163–169. Kluwer, 2003, ISBN 1-4020-1086-9.

Ouvrages scientifiques (OS)

Édition d'ouvrages collectifs

2008 [ED-9*] R. Janicki et L. Petrucci (eds.). Fifth Special Issue on Application of Concurrency to System Design. *Fundamenta Informaticae*, 2008. À paraître.

2007 [ED-8] Hanane Azzag et Gilles Venturini. Fouille du Données et algorithmes biomimétiques. *Atelier Fouille du Données et Algorithmes Biomimétiques aux 7ème EGC'07 Extraction et Gestion des Connaissances*, page 60, jan 2007.

[ED-7] Y. Bennani et E. Viennet, éditeurs. *Apprentissage artificiel et fouille de données, numéro spécial de la Revue des Nouvelles Technologies de l'Information*. Cepadues, 2007. 230 pages, publication en cours.

2006 [ED-6] K. Goossens et L. Petrucci, éditeurs. *Proceedings of the 6th International Conference on Application of Concurrency to System Design (ACSD'06), Turku, Finland*. IEEE Comp. Soc. Press, juin 2006, ISBN 0-7695-2556-3.

2004 [ED-5] Rushed Kanawati et Sylvie Salotti, éditeurs. *Proceedings of the second international workshop on applying case-based reasoning for sequence processing*, Madrid, September 2004.

[ED-4] Rushed Kanawati et Sylvie Salotti, éditeurs. *Second International Workshop on Applying CBR to Time Series Prediction*, Madrid, Spain, septembre 2004. In conjunction with ECCBR'04.

[ED-3] Rushed Kanawati, Sylvie Salotti et Farida Zehraoui, éditeurs. *12ème Atelier de Raisonnement à partir de cas (RàPC 2004)*, Villetaneuse, mars 2004. LIPN.

2003 [ED-2] Rushed Kanawati, Maria Malek et Sylvie Salotti, éditeurs. *First International Workshop on Applying CBR to Time Series Prediction*, Trondheim, Norway, juin 2003. NTNU, departement of information Science. In conjunction with ICCBR'03.

[ED-1] Rushed Kanawati, Maria Malek et Sylvie Salotti, éditeurs. *Proceedings of the first international workshop on applying Case-based reasoning for sequence processing*, Trondheim, june 2003.

Autres publications (AP)

Thèses et Habilitations

2007 [TH-28] Laure Bourgois. Exploitation d'ontologies pour la combinaison automatique de composants logiciels, 2007. Doctorat.

[TH-27] Fayçal Djerourou. Etude des méthodes de recherche locale pour la satisfaction de contraintes, 2007. Doctorat.

[TH-26] Mohand Ou Idir Khemoudj. Contribution de la programmation par contraintes à la résolution du problème de placement des arrêts des réacteurs nucléaires d'edf, 2007. Doctorat.

[TH-25] Farid Nouioua. Extraction et utilisation de normes pour le raisonnement causal sur un corpus textuel, 2007. Doctorat.

2006 [TH-24] Nejla Amara-Hachmi. Gama : une architecture générique pour le développement d'agents mobiles sensibles au contexte et auto-adaptatifs, 2006. Doctorat.

[TH-23] Sébastien Guérif. Approches connexionnistes pour l'exploration et l'extraction de connaissances à partir de données spatio-temporelles, 2006. Doctorat.

[TH-22] Mourad Hakem. Algorithmes d'ordonnancement distribué temps réel, 2006. Doctorat.

[TH-21] Sana Hamdoun. Modélisation, maintenance et integration des données hétérogènes, 2006. Doctorat.

[TH-20] Zeina El Ferkh Jrad. Apport des techniques de l'intelligence artificielle dans la négociation dynamique des niveaux de service : proposition d'une interface utilisateur pour l'internet de nouvelle génération, 2006. Doctorat.

[TH-19] Nicolas Marcos. Modélisation et intégration de la gestion tactique des ressources traction, 2006. Doctorat.

[TH-18] Hassine Moungla. Supervision distribuée de la qos dans les réseaux IP contrôlés par politique des opérateurs de telecommunications, 2006. Doctorat.

[TH-17] Alexandre Pauchet. Modélisation cognitive d'interactions humaines dans un cadre de planification multi-agents, 2006. Doctorat.

[TH-16] Frédéric Roupin, 2006. Thèse d'habilitation à diriger des recherches.

2005 [TH-15] Nahla El Zant Kadhi. Apprentissage de motifs relationnels dans des objets partiellement ordonnés et application à la caractérisation de motifs dans l'adn, 2005. Doctorat.

[TH-14] Hakima Kadri-Dahmani. Méthodes de modélisation pour la gestion de la notion de cycle de vie dans les systèmes d'information géographique, 2005. Doctorat.

[TH-13] Irène Loiseau. Sur la génération de colonnes en nombres entiers, 2005. Doctorat.

[TH-12] Gabriel Paillard. Quelques contributions à l'algorithmique distribuée. réseaux de capteurs et génération de nombres premiers, 2005. Doctorat.

2004 [TH-11] Brigitte Biébow. De Dasert à Terminae ; voyage au pays de la modélisation des connaissances à partir de textes, 2004. Thèse d'habilitation à diriger des recherches.

[TH-10] Alain Faye. Contributions à la programmation quadratique en 0-1, 2004. Thèse d'habilitation à diriger des recherches.

[TH-9] Touria Ait El Mekki. Construction semi-automatique d'index de fin de livre, 2004. Doctorat.

[TH-8] Anass Nagih. Contributions à la résolution des problèmes en nombres entiers : réoptimisation, problèmes de grande taille et planification en transport, 2004. Thèse d'habilitation à diriger des recherches.

[TH-7] Adeline Perrin-Nazarenko. Donner accès au contenu des documents textuels. acquisition de connaissances et analyse de corpus spécialisés, 2004. Thèse d'habilitation à diriger des recherches.

[TH-6] Guillaume Vauvert. Atteinte de consensus basé sur l'échange de préférences entre agents rationnels et autonomes : application à la formation d'alliances, 2004. Doctorat.

[TH-5] Farida Zehraoui. Systèmes d'apprentissage connexionniste et raisonnement à partir de cas pour la classification et le classement des séquences, 2004. Doctorat.

2003 [TH-4] Khalid Benabdeslem. Approches connexionnistes pour la visualisation et la classification des données issues d'usages d'internet, 2003. Doctorat.

[TH-3] Alejandro Guerra-Hernandez. Apprentissage d'agents rationnels bdi dans un univers multi-agents, 2003. Doctorat.

[TH-2] Didier Mallat-Desmortiers. Raisonnement automatique sur les croyances et les incertitudes d'un agent formalisé au sein de la théorie de l'interaction rationnelle, 2003. Doctorat.

[TH-1] Babacar Thiongane. Réoptimisation dans le dual lagrangien du biknapsack en variables 0-1, 2003. Doctorat.

Rapports de contrats

2007 [RC-34] L. Alfandari, L. Létocart, A. Nagih, A. Plateau et S. Toulouse. Méthodes de décomposition pour l'optimisation discrète. Rapport de contrat, LIPN, 2007.

[RC-33] Younès Bennani et Emmanuel Viennet. Réduction des dimensions de l'espace de description (Rapport d'avancement 1, ST3.21, Projet Infom@gic, Pôle Cap Digital). Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, juillet 2007.

2006 [RC-32] L. Alfandari, L. Létocart, A. Nagih, A. Plateau et S. Toulouse. Méthodes de génération de colonnes pour la planification optimale des locomotives. Rapport de contrat, LIPN, 2006.

[RC-31] L. Alfandari, L. Létocart, A. Nagih, A. Plateau et S. Toulouse. Techniques de résolution approchée : heuristiques et métaheuristiques. Rapport de contrat, LIPN, 2006.

[RC-30] Erick Alphonse, Sophie Aubin, Julien Derivière, Thierry Hamon, Dunja Mladenic, Adeline Nazarenko, Claire Nédellec, Thierry Poibeau, Davy Weissenbacher et Qiang Zhou. Deliverable D5.1 : Report on method and language for the production of the augmented document representations. ALVIS, 2006.

[RC-29] Sujeevan Aseervatham et Emmanuel Viennet. Méthodes à Noyau appliquées aux textes structurés (Rapport d'avancement 1, ST3.12, Projet Infom@gic, Pôle Cap Digital). Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, juillet 2006.

[RC-28] Gérard Bécher, Patrice Enjalbert, Estelle Fievé, Laurent Gosselin, François Lévy et Gérard Ligozat. Rapport technique du projet OGRE. 92 pages PI TCAN du CNRS, 2006.

[RC-27] Younès Bennani et Lahcène Dehni. Apprentissage non-supervisé pour la visualisation et la classification de données spatio-temporelles (Rapport d'avancement 1, ST3.21, Projet ANR SILLAGES Suivi Individuel par Label éLectronique Au sein de Groupes complexES). Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, juillet 2006.

[RC-26] Dominique Bouthinon et Henri Soldano. Concept learning from pieces of information. Rapport Technique 2006-03, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, 2006.

[RC-25] Julien Derivière et Adeline Nazarenko. Deliverable D5.2 : Report on theory and software of normalization options for IR (platform conception). Technical report, ALVIS IST-1-002068 STREP project, 2006.

[RC-24] Thierry Poibeau et Adeline Nazarenko. Rapport InfoM@gic "État de l'art des méthodes linguistiques et statistiques pour l'extraction d'informations". 25 pages, Projet Infom@gic, 2006.

2005 [RC-23] Armen Aghasaryan, Yves Cochet, Francine Krief, Eric Moreau, Hassine Moungla et Sophie Piekarec. State of the art on service contract expression languages, policy and network resource models. Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, janvier 2005. Rapport 3.1.1- Self Aware management project (SWAN- RNRT project).

[RC-22] L. Alfandari, L. Létocart, A. Nagih, A. Plateau et S. Toulouse. Modélisation et approches de résolutions. Rapport de contrat, LIPN, 2005.

[RC-21] L. Alfandari, L. Létocart, A. Nagih, A. Plateau et S. Toulouse. Modélisation et résolution de problèmes de couverture de tâches : logiciels du domaine libre, logiciel du commerce et heuristiques. Rapport de contrat, LIPN, 2005.

[RC-20] Sébastien Guérif et Younès Bennani. Apprentissage numérique non supervisé partir de données évolutives. Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, Villetaneuse, France, juin 2005. Deuxième rapport de recherche ANRT (CIFRE 654-2002).

[RC-19] Zeina Jrad, Badr Benmammar, Francine Krief et Nader Mbarek. Dynamique de l'environnement : Scénarios, simulations et maquette. Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, juillet 2005. Livrable de projet RNRT : IP-SIG (IP-Signaling).

[RC-18] Hassine Moungla. Specification of conflict resolution module. Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, mars 2005. Rapport 3.4.2- Self Aware management project (SWAN- RNRT project).

[RC-17] Hassine Moungla et Francine Krief. Policy based management specification. Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, mars 2005. Rapport 3.4.1- Self Aware management project (SWAN- RNRT project).

2004 [RC-16] Erick Alphonse, Philippe Bessières, Gaël Deral, Gaëtan Lehman, Alain-Pierre Manine, Claire Nédellec, Sophie Aubin, Julien Derivière, Thierry Hamon, Adeline Nazarenko, Thierry Poibeau et Davy Weissenbacher. Deliverable D6.2 : Requirements for integration of WP6 results into WP5 normalization and representation tasks and into WP9. ALVIS deliverable report ALVIS, 2004.

[RC-15] Sophie Aubin, Philippe Bessières, Thierry Hamon, Gaëtan Lehman, Alain-Pierre Manine, Adeline Nazarenko, Claire Nédellec, Mohamed Ould Abdel Vetah et Thierry Poibeau. Deliverable D6.1 : Requirement and training data for biological domain. ALVIS Deliverable Report ALVIS, 2004.

[RC-14] Sophie Aubin, Thierry Hamon et Adeline Nazarenko. Analyse terminologique (application et évaluation sur des textes de biologie). Contrat ExtraPloDocs (laboratoire LIPN, laboratoire INRA/MIG, Isoft), 2004.

[RC-13] Sophie Aubin et Adeline Nazarenko. Choix, évaluation et adaptation d'un analyseur syntaxique pour l'extraction d'information. Rapport ExtraPloDocs 3.3a. Contrat ExtraPloDocs (laboratoire LIPN, laboratoire INRA/MIG, Isoft), 2004.

[RC-12] Yamina Elmrabet, Sujeevan Aseervatham, Philippe Dague et Aomar Osmani. Réalisation du système de diagnostic pour le système EMB du véhicule. Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, 2004.

[RC-11] Thierry Gaudré, Vania Joloboff, Yamina Elmrabet, Sujeevan Aseervatham, Philippe Robin, Philippe Dague et Aomar Osmani. Réalisation d'un environnement système évolutif pour le diagnostic automobile,. Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, 2004.

[RC-10] Sébastien Guérif et Younès Bennani. Approches connexionnistes pour l'exploration et l'extraction de connaissances partir de données spatio-temporelles : application aux données géographiques. Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, Villetaneuse, France, avril 2004. Premier rapport de recherche ANRT (CIFRE 654-2002).

[RC-9] Zeina Jrad, Badr Benmammar, Francine Krief et Nader Mbarek. Dynamique de l'environnement. Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, juillet 2004. Rapport de projet RNRT : IP-SIG (IP-Signaling).

[RC-8] Hassine Moungla et Francine Krief. Spécification d'un modèle GMPLS dérivé du modèle générique. Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, février 2004. Modélisation et Apprentissage pour une Gestion Distribuée des Alarmes De bout En boUt (projet RNRT-MAGDA2).

[RC-7] Thierry Poibeau et Jean-François Berroyer. Analyse et repérage des entités nommées. 60 pages, rapport du Projet RNTL ExtraPloDocs 3.3b, 2004.

2003 [RC-6] Younès Bennani et Khalid Benabdeslem. Codage et compression de données issues d'internet. Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, janvier 2003. RT 1, Contrat de recherche Maxeem-NumSight.

[RC-5] Younès Bennani et Khalid Benabdeslem. Visualisation des traces de navigation par apprentissage non-supervisé. Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, mars 2003. RT 2, Contrat de recherche Maxeem-NumSight.

[RC-4] Zeina Jrad, Khaldoun Al Agha, Walid Dabbous, Ghamri Y. Doudane, Francine Krief, Frédéric Le Garrec, Sandrine Masson, Pascale Minet, Thi-Mai-Trang Nguyen, Guy Pujolle, Rafael Rizo et Rares Serban. ARCADE (ARchitecture de Contrôle Adaptative Des Environnements IP). Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, janvier 2003. Rapport final projet RNRT.

[RC-3] Larbi Ait Mohamed et Aomar Osmani. Sélection des MIB standards et des MIB CISCO. Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, janvier 2003.

[RC-2] Emmanuel Viennet. Estimation de la QoS de bout en bout par apprentissage numérique. Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, décembre 2003. Modélisation et Apprentissage pour une Gestion Distribuée des Alarmes De bout En boUt (projet RNRT-MAGDA2).

[RC-1] Arnaud Zeboulon, Younès Bennani et Khalid Benabdeslem. Classification et reconnaissance de séquences de navigations par SOM et modèle de mélanges. Rapport Technique, Laboratoire d'Informatique de Paris Nord - UMR CNRS 7030 - Institut Galilée - Université Paris 13, juin 2003. RT 3, Contrat de recherche Maxeem-NumSight.

Rapports

2007 [Ra-31] Heithem Abbes, Christophe Cérin, Jean-Christophe Dubacq et Mohamed Jemni. Performance analysis of publish/subscribe systems. RR 2007-XX, LIPN - CNRS UMR 7030 - Université Paris Nord, 2007.

[Ra-30] S. Bardin, A. Finkel, J. Leroux et L. Petrucci. Fast : Acceleration from theory to practice. Research Report LSV-07-16, Laboratoire Spécification et Vérification, ENS Cachan, France, avril 2007. 21 pages.

[Ra-29] Jean Cardinal, Vlady Ravelomanana et Mario Valencia-Pabon. Chromatic edge strength of some multigraphs. RR 2007-XX, Computer Science Dept.- Université Libre de Bruxelles & LIPN - CNRS UMR 7030 - Université Paris Nord, 2007.

[Ra-28] C. Choppy, L. Petrucci et G. Reggio. A method for modelling with coloured nets. Rapport Technique, Universite Paris 13, 2007.

[Ra-27] Philippe Darondeau et Laure Petrucci. Modular automata 2 distributed Petri nets 4 synthesis. Rapport Technique 6192, INRIA, mai 2007.

[Ra-26] L. Petrucci. ISO/IEC 15909 — Part 3 : Extensions, novembre 2007. Working document of ISO/IEC JTC1-SC7-WG19, ref. PA2-018.

[Ra-25] Nora Touati, Lucas Létocart et Anass Nagih. Méthodes de décomposition pour l'optimisation discrète. Rr, LIPN - CNRS UMR 7030 - Université Paris Nord, 2007.

2006 [Ra-24] Didier El Baz, Moussa Elkihel, Laurent Gély et Gérard Plateau. Improved time and space complexity for kianfar's inequality rotation algorithm. RR 06834, LAAS, Toulouse, 2006.

[Ra-23] Franck Butelle, Lucian Finta et Mourad Hakem. Real-time distributed scheduling of precedence graphs on arbitrary wide networks. RR 2006.01, LIPN - CNRS UMR 7030 - Université Paris Nord, 2006.

[Ra-22] C. Choppy, L. Petrucci et G. Reggio. A method for modelling with place/transitions nets. Rapport Technique, Universite Paris 13, 2006.

[Ra-21] Fayçal Djerourou et Hachemi Bennaceur. For more pruning values when solving Max-CSP. Rr, LIPN, 2006.

[Ra-20] Max Kanovich et Jacqueline Vauzeilles. Linear logic in temporal planning for actions with delayed effects. Rapport Technique, LIPN, 2006.

[Ra-19] M. I. Khemmoudj et H. Bennaceur. Clique based lower bounds for Max-CSP. RR 2006-02, Laboratoire d'informatique de Paris-Nord, 2006.

[Ra-18] Olivier Laval, Sophie Toulouse et Anass Nagih. Rapport de recherche sur le problème du plus court chemin contraint. RR 2006-05, LIPN - CNRS UMR 7030 - Université Paris Nord, 2006.

[Ra-17] J.-Y. Marion et J.-Y. Moyen. Heap-size analysis for assembly programs. Rapport Technique, LIPN, 2006.

[Ra-16] Catherine Recanati, Younès Bennani et Nicoleta Rogovschi. Analyse de séquences verbales - une segmentation numérique des temps et catégories aspectuelles dans des récits d'accidents. 19 pages, 2006.

[Ra-15] Nora Touati, Lucas Létocart et Anass Nagih. Sur l'accélération de la convergence de la génération de colonnes. RR 2006-4, LIPN - CNRS UMR 7030 - Université Paris Nord, 2006.

2005 [Ra-14] Laurent Alfandari. Improved approximation of the soft-capacitated facility location problem. RR 5003, ESSEC Research Center, 2005.

[Ra-13] David Delahaye et Micaela Mayero. Diophantus' 20th Problem and Fermat's Last Theorem for $n = 4$: Formalization of Fermat's Proofs in the Coq Proof Assistant. Rapport Technique, 2005.

[Ra-12] Christophe Fouqueré et Virgile Mogbil. Rewritings in polarized (partial) proof structures. Francois Lamarche Paola Bruscoli et James Stewart, éditeurs, *1st Workshop on Structures and Deductions*, volume ISSN 1430-211X de *Technical Report*, pages 95–109. Technische Universität Dresden, juillet 2005.

[Ra-11] Micaela Mayero, B. Monsuez et F. Védrine. How an "Incoherent Behavior" inside generic hardware component characterizes functional errors. Rapport Technique, 2005.

[Ra-10] Claire Nédellec et Adeline Nazarenko. Ontologies and Information Extraction. LIPN Internal Report, 2005.

2004 [Ra-9] Laurent Alfandari. Choice rule with size constraints for multiple-criteria decision making. RR 4002, ESSEC Research Center, 2004.

[Ra-8] C. Bentz, M-C. Costa, L. Létocart et F. Roupin. A bibliography on multicut and integer multiflow problems. RR 654, Centre d'Etude et de Recherche en Informatique du CNAM, Conservatoire National des Arts et Métiers, 2004.

[Ra-7] Christophe Fouqueré et Virgile Mogbil. Modules and logic programming. Rapport Technique, LIPN, Université Paris 13, 2004. cs.LO/0411029.

[Ra-6] F. Kordon et L. Petrucci. Proposal for an addendum to ISO/IEC 15909-1, novembre 2004. NWI submitted by AFNOR to ISO/IEC JTC1-SC7-WG19, ref. MAL-012.

[Ra-5] Catherine Recanati. Diagrammes pour résoudre le problème d'Einstein et celui d'un joueur de Mastermind. 29 pages, 2004.

2003 [Ra-4] K. Lellahi et A. Zamulin. A state-based algebraic semantics of statements of sequential java. Rapport Technique 3, LIPN, 2003.

[Ra-3] G. Reggio, E. Astesiano et C. Choppy. CASL-LTL : A CASL extension for dynamic reactive systems version 1.0 – summary. Rapport Technique DISI-TR-03-36, Univ. of Genova, 2003.

[Ra-2] Babacar Thiongane, Anass Nagih et Gérard Plateau. Algorithms and complexity for parametric 0-1 linear programs relative to the objective function. RR 2003-01, Laboratoire d'Informatique de Paris-Nord (LIPN), université Paris 13, 2003.

2002 [Ra-1] K. Lellahi et A. Zamulin. Implicit state approach for formalisation of sequential java-like programs. Rapport Technique 10, LIPN, 2002.

Outils

2006 [Lo-14] Sophie Aubin et Thierry Hamon. YaTeA - version 2006. YaTeA (Yet Another Term ExtrActor) aims at identifying and extracting noun phrases which are potential terms (*i.e.* term candidates). Projet Européen STREP IST-1-002068-STP - ALVIS, 2006.

[Lo-13] Franck Butelle, Ronald King et Frédéric Toumazet. Schur, an interactive program for calculating properties of lie groups and symmetric functions. Logiciel de calcul scientifique sous licence GPL, 2006.

[Lo-12] Julien Derivière et Thierry Hamon. Alvis NLP platform. A scalable architecture using existing NLP tools to annotate large collections of web documents, 2006.

[Lo-11] Thierry Hamon. SynoTerm - version 2006. Logiciel de détection de relations de synonymie entre termes, 2006.

[Lo-10] J.-V. Loddo. Iutoppix, septembre 2006. Cdrom amorçable, auto-configurable, dérivé de Knoppix 5.0.1, francisé et spécialisé pour les enseignements aux dép. R&T et INFO de l'IUT Villetaneuse. Licence GPL.

[Lo-9] J.-V. Loddo. Synopsys, juin 2006. Utilitaire permettant d'automatiser l'analyse lexicale, syntaxique et en partie sémantique des paramètres d'appel d'un script bash. Code source OCaml, licence GPL.

[Lo-8] Sylvie Szulman. Logiciel TERMINAE - version 2006. Plate-forme d'aide à la construction de terminologies et d'ontologies à partir de l'étude de textes. Intégre un concordancier (Linguae) et permet la visualisation des résultats de Lexter et/ou Syntex développés par D. Bourigault, 2006.

2005 [Lo-7] Laurent Audibert. Logiciel Concordanta. Logiciel de type concordancier, 2005.

[Lo-6] Laurent Audibert. Logiciel xexplora et explora. Outil de recherche et d'observation de phénomènes linguistiques dans les corpus écrits, 2005.

[Lo-5] S. Bardin, A. Finkel, J. Leroux et L. Petrucci. FAST : Fast Acceleration of Symbolic Transition systems, juin 2005. Tool presentation at ACSD'05.

[Lo-4] J.-V. Loddo. Netxkiss, mai 2005. Application (en mode texte) de simulation d'un réseaux d'ordinateurs qui exploite la technologie UML (User Mode Linux) et repose sur un DVD amorçable, auto-configurable, dérivé de Knoppix 3.8. Code source OCaml, licence GPL.

2004 [Lo-3] J.-V. Loddo. Knoppix4gtr, septembre 2004. Cdrom amorçable, auto-configurable, dérivé de Knoppix 3.8, francisé et spécialisé pour les enseignements au dép. GTR de l'IUT Villetaneuse. Licence GPL.

[Lo-2] Thierry Poibeau et Jean-François Berroyer. Tagen, un système de reconnaissance d'entités nommées. 60 pages, 2004.

[Lo-1] Jean-Vincent Loddo et Hamon Thierry. Marionnet/iutoppix : un Environnement Pédagogique pour la Simulation de Réseaux Locaux. Accepted for demonstration at the conference Environnement Informatique pour l'Apprentissage Humain (EIAH), Lausanne, 2007.

Travaux en cours de soumission

2007 [So-6] Sujeevan Aseervatham, Emmanuel Viennet et Younès Bennani. Semi-structured document categorization with a semantic kernel. *International Journal of Computational Intelligence and Applications*, 2007. Soumis en août 2007.

[So-5] Hanane Azzag, Fabien Picarougne, Pascale Kuntz, Christiane Guinot et Gilles Venturini. A survey of biomimetic algorithms for data clustering. *Journal of Soft Computing, Special Issue on Evolutionary and Metaheuristic-based Data Mining*, 2007.

[So-4] Mustapha Lebbah et Younès Bennani. A probabilistic self-organizing map for binary data topographic clustering. *International Journal of Computational Intelligence and Applications (IJCIA)*, 2007.

2006 [So-3] C. Bentz, M-C. Costa, L. Létocart et F. Roupin. Multicut and integral multiflow in rings. RR 1050, Centre d'Etude et de Recherche en Informatique du CNAM, Conservatoire National des Arts et Métiers, 2006. Soumis à EJOR.

[So-2] Christophe Fouqueré et Virgile Mogbil. Rewritings in polarized proof structures. *Mathematical Structures in Computer Science*.

[So-1] Max Kanovich et Jacqueline Vauzeilles. Linear logic foundations for artificial intelligence. *Annals of Mathematics and Artificial Intelligence*.

Miscellanea

2005 [Mi-1] Daniel Kayser. Compte-rendu de lecture du livre "Interpréter en contexte" sous la dir. de Francis Corblin et de Claire Gardent. *T.A.L.*, 46(1) :197–200, 2005.

Partie V
Acronymes

A³ Apprentissage Artificiel et Applications
ACI Action Concertée Incitative
AERES Agence d'évaluation de la recherche et de l'enseignement supérieur
AM allocataire moniteur
ANR Agence National de la Recherche
ARC Australian Research Council
AS Action Spécifique (CNRS)
ATER attaché temporaire d'enseignement et de recherche
AUF Agence Universitaire pour la Francophonie
B boursier
BQR Bonus Qualité-Recherche (université)
CDD Contrat à Durée Déterminée
CDI Contrat à Durée Indéterminée
CEDRIC Centre d'Etudes et de Recherche en Informatique du CNAM - EA 1395
CIFRE Conventions Industrielles de Formation par la Recherche
CNAM Centre National des Arts et Métiers
CNR Consiglio Nazionale delle Ricerche
CNRS Centre National de la Recherche Scientifique
COST european Cooperation in the field Of Scientific and Technical Research
CR chargé de recherche
DR directeur de recherche
DEA Diplôme d'Etudes Approfondies
DESS Diplôme d'Etudes Supérieures Spécialisées
DGE Direction Générale de l'Equipement (ministère)
EA Equipe d'Accueil
ECOS Evaluation-orientation de la COopération Scientifique
EDF Electricité de France
ENS Ecole Normale Supérieure
ESSEC Ecole Supérieure des Sciences Economiques et Commerciales
FRE Formation de Recherche en Evolution
GDR Groupement de Recherche
GDRE Groupement de Recherche Européen
HDR Habilitation à Diriger des Recherches
I3M Institut de Mathématiques et Modélisation de Montpellier - UMR 5030
IA Intelligence Artificielle
IG Institut Galilée
IML Institut de Mathématiques de Luminy - UMR 6206
INPL Institut National Polytechnique de Lorraine
Inria Institut National de la Recherche en Informatique et Automatique
Irisa Institut de Recherche en Informatique et en Systèmes Aléatoires
IRMA Institut de Recherche Mathématique Avancée
IUT Institut Universitaire Technologique
IUTB IUT de Bobigny
IUTV IUT de Villetaneuse
LAGA Laboratoire Analyse, Géométrie et Applications, université Paris 13 - UMR 7539

LAMSADE Laboratoire d'Analyse et Modélisation de Systèmes pour l'Aide à la Décision - UMR 7024

LARIA Laboratoire de Recherche en Informatique d'Amiens - FRE 2733

LCR Logique, Calcul et Raisonnement

LEEC Laboratoire d'Ethologie Expérimentale et Comparée, université Paris 13 - UMR 7153

LIENS Laboratoire d'Informatique de l'Ecole Normale Supérieure - UMR 8548

LIF Laboratoire d'Informatique Fondamentale - UMR 6166

LIFAR Laboratoire d'informatique fondamentale et appliquée de Rouen - EA 2655

LIFM Laboratoire d'Informatique Fondamentale de Marseille - UMR 6166

LIG Laboratoire d'Informatique de Grenoble

Lim&Bio Laboratoire d'Informatique Médicale et Bio-Informatique, université Paris 13 - EA

LIP Laboratoire de l'Informatique du Parallélisme

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

LLI Laboratoire de Linguistique-Informatique, université Paris 13 - FRE 2882

LRI Laboratoire de Recherches en Informatique, université Paris 11 - UMR 8623

LSV Laboratoire Spécification et Vérification

MAE Ministère des Affaires Etrangères

MCF maître de conférences

MC-TOK action Marie Curie - Transfer of Knowledge

MENESR Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche

MSTP Mission Scientifique Technique et Pédagogique, ministère délégué à l'Enseignement Supérieur et à la Recherche

MPRI Master Parisien de Recherche en Informatique

NII National Institute of Informatics

OC Optimisation Combinatoire

OCAD Optimisation Combinatoire et Algorithmique Distribuée

ONERA Office National d'Etudes et de Recherches Aérospatiales

PA professeur assistant

PAI Programme Actions Intégrées

PAN Polskiej Akademii Nauki – Académie des sciences de Pologne

PCRDT Programme Cadre de Recherche et Développement Technologique

PPF Plan Pluri-Formation

PPS Preuves Programmes et Systèmes - UMR 7126

PU professeur des universités

RCLN Représentation des Connaissances et Langage Naturel

RNRT Réseau National de Recherche en Télécommunication

RNTL Réseau National de recherche et d'innovation en Technologies Logicielles

SAIC Service d'Activité Industrielle et Commerciale

SMF Société Mathématique de France

SNCF Société Nationale des Chemins de Fer

SECyT Secretaría de Ciencia y Tecnológia – Université de Buenos Aires, Argentine

STIC Sciences et Technologies de l'Information et de la Communication

STREP Specific Targeted Research Projects (6ème PCRDT)

UE Union européenne

UMR Unité Mixte de Recherche

UniSA University of South Australia

USAR Unité Support de l'Agence Nationale de la Recherche