

BULLETIN

DE

L'AFIA



JUILLET 2014

N° 85

Association Française pour l'Intelligence Artificielle

Présentation du bulletin

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

Pour contacter l'AFIA

Président

Yves DEMAZEAU
L.I.G./C.N.R.S., Maison Jean
Kuntzmann, 110, avenue de la Chimie,
B.P. 53, 38041 Grenoble cedex 9
Tél : +33 (0)4 76 51 46 43
Fax : +33 (0)4 76 51 49 85
Yves.Demazeau@imag.fr
[http://membres-lig.imag.fr/
demazeau](http://membres-lig.imag.fr/demazeau)

Serveur WEB

<http://www.afia.asso.fr>

Adhésions, liens avec les adhérents

Thomas GUYET
Laboratoire Informatique d'Agrocampus-Ouest
65, rue de Saint-Brieuc
35042 Rennes cedex
Mél. : tresorier@afia.asso.fr

Personnes morales adhérentes à l'AFIA

ENSMSE, Université Paris Dauphine, LORIA, LIRIS, LIMSI, IRIT/SMAC, EDF/STEP, LIPADE, IFFSTAR, LIRMM, TAO, LIFL, GREYC, LIG, ONERA, IRSTEA-TETIS, INRA, LITIS

Conseil d'Administration de l'AFIA

Yves DEMAZEAU, président
Pierre ZWEIGENBAUM, vice-président
Amélie CORDIER, vice-présidente
Olivier BOISSIER, secrétaire
Catherine FARON-ZUCKER, secrétaire adjoint
Catherine TESSIER, secrétaire adjoint
Thomas GUYET, trésorier
Patrick REIGNIER, webmestre

Membres :

Carole ADAM, Patrick ALBERT, Christine BOURJOT, Serge GARLATI, Sébastien KONIECZNY, Vincent LEMAIRE, Nicolas MAUDET, Philippe MORIGNOT, Bruno PATIN, Laurent VERCOUTER, .

Comité de Rédaction

Charles Gouin-Vallerand
Rubrique « I.A. au Québec »

TÉLUC

5800, rue Saint-Denis, Montréal, Canada

charles.gouin-vallerand@teluq.ca

Nicolas Maudet

Rédacteur adjoint

LIP6, Université Pierre et Marie Curie

4, place Jussieu, 75005 Paris
maudetn@lip6.fr

Philippe Morignot

Rédacteur en chef

LIFEWARE, INRIA Rocquencourt

Domaine de Voluceau, B.P.105, 78150 Le Chesnay

pmorignot@yahoo.fr

Patrick Reignier

Rubrique « Résumés de thèse et HDR »

PRIMA, INRIA Rhône-Alpes 655, avenue de l'Europe, 38334 Saint-Ismier cedex

Patrick.Reignier@inrialpes.fr

Laurent Vercouter

Rédacteur adjoint

LITIS, INSA de Rouen

avenue de l'université, BP8 76801 St-Étienne-du-Rouvray

laurent.vercouter@insa-rouen.fr



3 lauréats du prix de thèse 2014 en I.A. !

Le jury du prix de thèse 2014 en Intelligence Artificielle, présidé par Amedeo Napoli, a rendu son verdict. Pour cette année 2014, ce sont deux thèses qui ont été désignées pour le premier prix *ex aequo* de la meilleure thèse 2014 en Intelligence Artificielle, et une autre thèse distinguée pour le 1er accessit.

Abdallah Saffidine pour une thèse intitulée "*Solving Games and All That*" soutenue à l'université Paris Dauphine et Michaël Thomazo pour une thèse intitulée "*Conjunctive Query Answering under Existential Rules. Decidability, Complexity and Algorithms*" soutenue à l'université Montpellier 2, se partagent le premier prix, et Émilie Morvant pour une thèse intitulée "*Apprentissage de vote de majorité pour la classification supervisée et l'adaptation de domaine : approches PAC-Bayésiennes et combinaison de similarités*" soutenue à l'université d'Aix-Marseille obtient le premier accessit. L'AFIA félicite une nouvelle fois les trois lauréats pour l'excellence de leurs travaux et remercie le jury du prix de thèse I.A. 2014 pour son travail d'évaluation. Les deux récipiendaires du premier prix ont été invités à présenter leurs travaux au prochain congrès RFIA qui se tiendra à Rouen. Vous trouverez dans ce bulletin un résumé de ces thèses (ainsi que de l'accessit).

Comme de tradition pour les bulletins estivaux de l'AFIA, le dossier spécial est consacré à la présentation d'équipes francophones de recherche de Intelligence Artificielle. Cinq descriptifs d'équipe de recherche composent ce dossier par un résumé de leurs axes principaux de recherche et contributions marquantes. Vous trouverez enfin dans ce bulletin un compte-rendu de la journée bilatérale consacrée aux liens entre l'Intelligence Artificielle et le domaine de la Santé, avec un focus plus spécifique sur l'Ingénierie des Connaissances.

Bonne lecture.

Nicolas Maudet, Philippe Morignot & Laurent Vercoeur
Rédacteurs en chef

Dossier “Présentation d’équipes de recherche en I.A.”

AMA - Grenoble

*(Analyse de données, Modélisation et Apprentissage
automatique)*

Laboratoire LIG - Université Joseph Fourier / Grenoble
1 / CNRS

Bâtiment CE4, Allée de la Palestine, 38610 Gières -
France

<http://ama.liglab.fr>

Responsable : Eric Gaussier

mel : Eric.Gaussier@imag.fr

tel : +33 (0)456 52 03 09

Chercheurs :

- AMBLARD, Cécile, MdC Université de Grenoble
- AMINI, Massih-Reza, Professeur Université de Grenoble
- AYCARD, Olivier, MdC Université de Grenoble, HdR
- BISSON, Gilles, CR CNRS
- BROUARD, Christophe, MdC Université de Grenoble
- DOUZAL, Ahlame, MdC Université de Grenoble
- GARBAY, Catherine, DR CNRS
- GAUSSIÉ, Eric, Professeur Université de Grenoble
- GORDON, Mirta, DR CNRS, Emérite

Mots-clés :

- Analyse de données structurées, non structurées, temporelles
- Théorie de l’apprentissage
- Classification large échelle, en grande dimension
- Interaction homme-machine en apprentissage
- Localisation et tracking d’objets en mouvements
- Suivi du comportement humain

- Modélisation des systèmes sociaux, diffusion de l’information et prédiction d’activité

Thématique de recherche

Les recherches de l’équipe s’inscrivent dans le champ de la "Science des données", avec un focus important sur l’analyse des données, l’apprentissage automatique et la modélisation de l’information. Nous développons de nouveaux modèles théoriques, des algorithmes et des systèmes pour l’analyse et la décision à partir de données complexes. Les recherches de l’équipe s’organisent selon 3 axes complémentaires :

Analyse de données et théorie de l’apprentissage

Nous nous intéressons dans cet axe à la conception et l’apprentissage de métrique dans des contextes supervisés et non supervisés pour différents types de données relationnelles, structurelles ou temporelles (comme les textes ou les séries chronologiques). Nous nous intéressons également à la classification large échelle (classification en grandes taxonomies comprenant typiquement des dizaines de milliers de catégories et des centaines de milliers de documents et propriétés).

Apprentissage et systèmes de perception

Deux perspectives complémentaires sont considérées dans cet axe : (i) celle de systèmes capables de naviguer de manière autonome au sein d’environnements dynamiques inconnus et (ii) celle de systèmes capables de suivre la physiologie et l’activité humaine. Ces deux perspectives soulèvent des problèmes difficiles en termes de fusion de données temporelles multi-capteurs, de localisation et de tracking, et de conception de processus de perception/interprétation de données sensibles au contexte.

DOSSIER “PRÉSENTATION D’ÉQUIPES DE RECHERCHE EN I.A.”

Modélisation des systèmes sociaux

Plusieurs problématiques sont abordées dans ce axe : (i) celle de la propagation de l’information au sein des réseaux de diffusion de contenus, dans une perspective de recommandation, (ii) celle de la prédiction des activités relatives à certains mots-clés dans les réseaux sociaux, sans connaissance a priori sur ces réseaux et (iii) celle de la modélisation des interactions entre agents humains et artificiels dans des contextes économiques ou de gestion de crise.

Production de logiciels

- Deux modèles d’accès à l’information que nous avons développés sont intégrés au sein des plateformes de recherche Lucene et Terrier.
- Les outils de détection et tracking d’objets que nous avons développés sont intégrés au sein d’un démonstrateur développé par FIAT.
- Les outils de prédiction d’activité dans les réseaux sociaux ont été développés en collaboration avec la société BestOfMedia/MediaTechNet. Ils sont disponibles sous licence GPL. Les bases de données correspondantes sont disponibles sur la base UCI.

Relations internationales

L’équipe AMA collabore et publie avec les personnes et équipes suivantes :

- Allemagne : A. Ngonga (Leipzig University),
- Australie : Longbing Cao, Guandong Xu, AAI (Advanced Analytics Institute) and UTS (Univ. Technology Sydney)
- Belgique : F. Maes (University KU Leuven),
- Canada : C. Goutte (National Research Council), F. Laviolette (Univ. Laval, Québec),
- Corée du Sud : Young-Min Kim, KISTI (Korean Institute of Science and Technology Information),
- États-Unis : L. Billard (Georgia University) and M. Taylor (Washington State University),
- Grèce : G. Paliouras (NCSR “Demokritos”), A. Kosmopoulos and I. Androutopoulos (AUEB) and A. Fachantidis (Aristotle University of Thessaloniki),
- Maroc : S. Ouatik (Fes University),
- Roumanie : Cluj University,
- Tunisie : C. Latiri, Y. Slimani (Univ. de la Manouba)
- Royaume-Uni : M. Lalmas (Yahoo Labs, London)

Projets marquants

Projet Européen

- Specific Support Action BioAsq (2012–14),
- IP Interactive (2010–13),
- Network of Excellence PASCAL2 (2008–12).

Projets Nationaux

- ANR CONTINT : IMAGIT (2010–13), FRAGRANCES (2009–12), METRICC (2009–12),
- ANR Blanc : CLASS-Y (2011–15), DYNOTEP (2012–16),
- MASTODONS de la Mission Interdisciplinarité du CNRS : PROSPECTOM (2012–), ARESOS (2012–), GARGANTUA (2013–),
- REI DGA : SUPERCO (2009–12).

Projets Locaux

L’équipe a obtenu des financements et bourses de thèses de la Région Rhône-Alpes et du Labex PERSYVAL.

Projets Industriels

- L’équipe a établi des collaborations avec plusieurs entreprises et PME (XEROX, Orange-Labs, Schneider Electric, TechMediaNetwork, Wizbii, ShortEdition) et dispose de financements directs avec ces entreprises, le plus souvent sous la forme de contrats CIFRE.
- FUI : Smart Support Center (2015–17). Partenaires : Coservit, Slaesforce, HP, Com6, LIG.

Publications significatives

Journaux internationaux

- A. Kosmopoulos, I. Partalas, E. Gaussier, G. Paliouras, I. Androutopoulos. Evaluation measures for hierarchical classification : a unified view and novel approaches. *Data Mining and Knowledge Discovery Journal*, Vol. 28, N. 5-6, 2014.
- C. Metzger and M. B. Gordon. A Model for Scaling in Firms’ Size and Growth Rate Distribution. *Physica A : Statistical Mechanics and its Applications*, pages 1–28, 2014.

DOSSIER “PRÉSENTATION D’ÉQUIPES DE RECHERCHE EN I.A.”

- B. Vettier and C. Garbay. Abductive Agents for Human Activity Monitoring. *International Journal on Artificial Intelligence Tools*, pages 1–34, 2014.
- A. Douzal-Chouakria and C. Amblard. Classification trees for time series. *Pattern Recognition*, 45(3) :1076–1091, Mar. 2012.
- B. Piwowarski, M.-R. Amini, and M. Lalmas. On using a quantum physics formalism for multidocument summarization. *Journal of the American Society for Information Science and Technology*, 63(5) :865–888, 2012.
- T-D Vu, J. Bulet, O. Aycard : Grid-based localization and local mapping with moving object detection and tracking. *Information Fusion* 12(1) : 58-69 (2011).

Conférences internationales

- R. Babbar, I. Partalas, E. Gaussier, and M.-R. Amini. On Flat versus Hierarchical Classification in Large-Scale Taxonomies. In *Advances in Neural Information Processing Systems 26*, pages 1824–1832, Etats-Unis, Dec. 2013.
- A. Fakeri-Tabrizi, M.-R. Amini, and G. Patrick. Multiview Semi-Supervised Ranking for Automatic Image Annotation. In *ACM Multimedia Conference, MM '13, Barcelona, Spain, October 21-25, 2013*, pages 513–516, Barcelone, Espagne, Oct. 2013.
- C. Frambourg, A. Douzal-Chouakria, and E. Gaussier. Learning Multiple Temporal Matching for Time Series Classification. In *Advances in Intelligent Data Analysis*, Springer, pages 198–209, Royaume-Uni, Oct. 2013.
- G. Bisson and C. Grimal. Co-clustering of Multi-View Datasets : a Parallelizable Approach. In *Proceedings of the IEEE International Conference on Data Mining (ICDM)*, pages 828–833, Bruxelles, Belgique, Dec. 2012.
- C. Brouard. Document Classification by Computing an Echo in a Very Simple Neural Network. In *Proceedings of the 24th IEEE International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 735–741, Athènes, Grèce, Nov. 2012. Session : Machine Learning and Data Mining III.

Paper awards

- “Best paper award” à la conférence ICTIR 2013, pour le papier intitulé “A Theoretical Analysis of Pseudo-Relevance Feedback Models” par Stéphane Clinchant et Eric Gaussier.

DUKe - Nantes

(Data User Knowledge)

*Laboratoire d'Informatique de Nantes Atlantique -
Université de Nantes / Ecole des Mines de Nantes /
CNRS*

<http://www.lina.univ-nantes.fr/?-DUKe-.html>

Responsable : Philippe Leray
mel : philippe.leray@univ-nantes.fr
tel : +33 (0)240 68 30 38

Chercheurs :

- Julien Blanchard, MCF
- Marc Gelgon, PR
- Régis Gras, PR émérite
- Fabrice Guillet, PR
- Mounira Harzallah, MCF
- Pascale Kuntz-Cosperec, PR
- Eric Languenou, MCF
- Hoel Le Capitaine, MCF
- Rémi Lehn, MCF
- Philippe Leray, PR
- José Martinez, PR
- Marie-Pierre Machouki, MCF
- Philippe Peter, MCF
- Fabien Picarougne, MCF
- Antoine Pigeau, MCF
- Yannick Prié, PR
- Gérard Ramstein, MCF HDR
- Guillaume Raschia, MCF
- Christine Sinoquet, MCF HDR
- Franky Trichet, MCF

Mots-clés :

- Apprentissage numérique et symbolique, découverte, classification
- Fouille de données, bases de données avancées, web sémantique
- Ingénierie et partage des connaissances, argumentation
- Intelligence collective, intelligence sociale, réseaux sociaux
- Interfaces intelligentes, interaction homme-machine, intelligence ambiante

DOSSIER "PRÉSENTATION D'ÉQUIPES DE RECHERCHE EN I.A."

- Raisonnement probabiliste et incertain, logique floue
- Représentation des connaissances, extraction et gestion des connaissances
- Web intelligence, internet du futur, protection de la vie privée

Thématique de recherche

La "science des données" recouvre le processus d'exploitation de jeux de données : acquisition, stockage, interrogation, découverte de connaissances. Elle s'appuie à la fois sur des questions de manipulation de données (indexation, interrogation) et de fouille de données (apprentissage, modélisation de connaissances, interactions, visualisation).

L'équipe DUKe (Data User Knowledge), résultat de la fusion des équipes COD et GRIM du LINA, est forte de ses compétences en manipulation de données (interrogation, anonymisation), en fouille de données (extraction de motifs, découverte de partitions, apprentissage de modèles graphiques probabilistes) et en interaction (visualisation de données/connaissances, immersion, annotations).

Notre ambition est de proposer des algorithmes de fouille ou d'apprentissage :

- prenant en compte des objectifs intégrant manipulation et fouille de données ;
- prenant en compte des connaissances expertes et/ou une interaction avec les utilisateurs grâce à des supports visuels adaptés.

Nos compétences générales dans le domaine des sciences des données illustrées par le triptyque manipulation-fouille-interaction s'appuient aussi par un engagement fort dans la recherche interdisciplinaire, avec un intérêt dans l'étude des traces numériques, avec des collaborations marquées dans le domaine des Arts, Lettres, Sciences humaines et sociales d'un côté, et de la Biologie de l'autre.

Projets marquants

Le projet ANR GRAPH-COMP (2005-08) nous a donné l'occasion de développer des méthodes d'analyse structurelle de grands réseaux. Ce projet, faisant interagir chercheurs en Histoire, Archéologie, Mathématiques et Informatique, portait sur la comparaison de grands graphes dans le cadre applicatif de la recherche de réseaux de sociabilités dans la société paysanne du Moyen-Age.

Le projet ANR Musimorphose (2013-...) est consacré à l'étude de l'impact du virage numérique sur la formation des goûts et les usages de la musique au quotidien. Nous y utilisons des algorithmes de fouille et de visualisation de données pour développer des outils à destination de sociologues de la culture ou d'ethnomusicologues, avec pour objectif d'aborder des questions sociales où réseaux sociaux et usages de l'internet sont primordiaux.

Le projet KIFANLO (2014-...) financé par la Fondation de France a pour but d'étudier avec des géographes les pratiques de pêche des flottilles ligériennes. Nous y contribuons par des travaux en ingénierie des ontologies dont la finalité est la valorisation du savoir des pêcheurs pour l'analyse des dynamiques spatiales et leur évolution dans le temps.

Dans le même objectif, nous contribuons cette année à 4 nouveaux projets ANR :

- ANTIMOINE - ANthropologie des TerrItaires - lecture du patriMOINE. Développement d'un environnement informatique intelligent pour faciliter l'accès aux objets du patrimoine culturel (disciplines partenaires SHS : linguistique)
- CIRESEFI - Contrainte et Intégration : pour une Réévaluation des Spectacles Forains et Italiens sous l'Ancien Régime (disciplines partenaires SHS : littérature, linguistique, histoire, économie du spectacle vivant)
- Episteme - Épistémologie transdisciplinaire des technologies numériques pour la conception d'un nouvel environnement de catégorisation contributive et d'instruments contributifs de recherche, d'enseignement et de diffusion associés (disciplines partenaires SHS : épistémologie, histoire, philosophie)
- HUBBLE - HUman oBServatory Based on anaLysis of e-LEarning traces. Création d'un observatoire pour la construction et le partage de processus d'analyse des traces e-learning massives (Disciplines partenaires SHS : didactique, sciences de l'éducation)

Nos compétences en fouille de graphes, visualisation ou apprentissage de modèles graphiques probabilistes sont aussi mises en œuvre via des projets dans le domaine des Sciences du Vivant. Citons le projet ANR MESOMORPH (2006-09) qui portait sur l'analyse quantitative des structures tri-dimensionnelles de nids d'insectes sociaux et l'analyse fonctionnelle (e.g. optimisation du transport et des communications à l'intérieur du nid). Nous avons ainsi contribué à l'analyse des réseaux 3D modélisant les intersections entre galeries issues des tomographies de nids réels.

DOSSIER “PRÉSENTATION D’ÉQUIPES DE RECHERCHE EN I.A.”

Le projet ANR SAMOGWAS (2013-...) porte sur le développement de modèles graphiques probabilistes pour les études d’association à l’échelle du génome, en collaboration avec généticiens et bio-informaticiens. Notre contribution concerne l’élaboration d’algorithmes avancés dédiés à la découverte d’associations génotypes-maladie complexe. Dans une approche intégrative, des données transcriptomiques et des connaissances tirées d’ontologies permettront d’enrichir les modèles et de guider cette découverte d’associations.

Nos travaux concernant l’analyse de réseaux sociaux et de manière plus générale l’analyse de traces numériques ont pu aussi être validés dans le cadre de plusieurs autres projets.

Le projet SOCIOPRISE (2009-12) financé dans le cadre de l’appel à projets "Web Innovant" du Plan de Relance Numérique, portait sur l’analyse statique, dynamique et sémantique de réseaux sociaux d’entreprises dédiée à l’organisation du travail et à la gestion des emplois et des compétences.

Le projet E-MAP (2010-12) concernait les réseaux sociaux informatiques. Nos contributions ont porté sur le développement de techniques d’analyse statistique des contenus postés sur ces réseaux, pour recommander des personnes ou des contenus, avec des problématiques de type détection de communauté, apprentissage de profil d’utilisateur, résumé d’information.

Le projet ANR Safimage (2007-10), en collaboration avec ALCATEL CIT, s’intéressait à l’analyse de traces numériques par des méthodes de classification dans le contexte de la sécurité informatique. Dans le même domaine, le projet ANR PLACID (2007-10) proposait une approche hybride à base de modèles graphiques probabilistes et de logique des préférences pour la corrélation d’alarmes dans les réseaux.

Ces mêmes compétences ont donné lieu à plusieurs collaborations industrielles dans le cadre de thèses CIFRE dans le domaine de l’informatique décisionnelle (Knowesia, PerformaSE SA, VM Matériaux, DataForPeople, Orange, EDF) ou l’ingénierie des connaissances (Tennaxia, SP2 Solutions).

10 publications significatives

- P. BRUNEAU, F. PICAROUGNE et M. GELGON. Interactive unsupervised classification and visualization for browsing an image collection. *Pattern Recognition*, 2010, 43(2) : 485-493. Elsevier
- P. BRUNEAU, M. GELGON et F. PICAROUGNE. A low-cost variational-Bayes technique for merging mixtures of probabilistic principal component analyzers. *Information Fusion*, 2013, (14) :268-280.
- P.-A. CHAMPIN, A. MILLE et Y. PRIE . Les traces numériques comme objets de premier niveau : une approche par les traces modélisées. *Intellectica*, 2013, (59) : 171–204.
- M. HARZALLAH, G. BERIO et A. OPDAHL. *New Perspectives in Ontological Analysis : Guidelines and Rules for Incorporating Modelling Languages into UEML*. *Information Systems*, 2012, V37, I5,484-507.
- H. LE CAPITAINE. A relevance-based learning model of fuzzy similarity measures. *IEEE Transactions on Fuzzy Systems*, 2012, 20(1) : 57-68.
- H. LE CAPITAINE. A unified view of class-selection with probabilistic classifiers. *Pattern Recognition*, 2014.
- C. MARINICA et F. GUILLET. Knowledge-Based Interactive Postmining of Association Rules Using Ontologies. *IEEE Transactions on Knowledge and Data Engineering*, 2010, 22(6) : 784-797.
- R. MOURAD, C. SINOQUET et P. LERAY. Probabilistic graphical models for genetic association studies. *Briefings in Bioinformatics*, 2012, 13(1) : 20–33.
- R. MOURAD, C. SINOQUET, N. ZHANG, T. LIU et P. LERAY. A survey on latent tree models and applications. *Journal of Artificial Intelligence Research*, 2013, 47 : 157-203.
- D. C. PORUMBEL, J.-K. HAO et P. KUNTZ. An efficient algorithm for computing the distance between close partitions. *Discrete Applied Mathematics*, 2011, 159(1) : 53-59.

AIDR - Paris

(Équipe Agents, Interaction, Decision, Raisonnement)
Laboratoire LAMSADE - CNRS UMR 7243 - Université
Paris Dauphine
Place du Maréchal de Lattre de Tassigny 75775 PARIS
Cedex 16

<http://www.lamsade.dauphine.fr/spip.php?article223>

Responsable : Tristan Cazenave, Suzanne Pinson
mel : cazenave@lamsade.dauphine.fr, pinson@lamsade.dauphine.fr
tel : +33 1 44 05 44 10, +33 1 44 05 45 51

DOSSIER “PRÉSENTATION D’ÉQUIPES DE RECHERCHE EN I.A.”

Chercheurs :

- Airiau Stéphane
- Balbo Flavien (jusqu’en novembre 2013)
- Cazenave Tristan
- Lang Jérôme
- Pigozzi Gabriella
- Pinson Suzanne
- Tsoukiàs Alexis

Étudiants Doctorants :

- Aknine-Arib Souhila
- Barrot Nathanael
- Cornu Marek
- Louati Amine
- Menif Alexandre
- Paget Nicolas

Mots-clés :

- Agents autonomes, systèmes multi-agents, systèmes autonomiques
- Applications de l’Intelligence Artificielle, méthodologie, évaluation
- Intelligence collective, intelligence sociale, réseaux sociaux
- Systèmes à base de règles, planification, aide à la décision

Thématique de recherche de l’équipe

Ce projet a pour objectif l’élaboration de modèles et d’algorithmes pour les processus de raisonnement et de prise de décision, avec en ligne de mire le développement de systèmes complexes de résolution de problèmes pour l’aide à la décision ou la décision automatisée. Plus précisément, nos travaux de recherche portent sur l’étude et la conception de **systèmes intelligents destinés à automatiser ou à aider la prise de décision**. Les types de problèmes de décision considérés vont de la décision individuelle à la décision de groupe, tandis que les outils utilisés sont tantôt centralisés, tantôt distribués. Le concept-clé de notre projet est celui **d’agent intelligent**, et se positionne assez largement dans le domaine de l’intelligence artificielle. Les agents intelligents peuvent présenter une certaine autonomie, ont des croyances et

des préférences qui leur sont propres, ainsi que des capacités de raisonnement, de communication, et de prise de décision.

Les recherches menées se regroupent autour de plusieurs axes :

- Modèles de coordination multi-agent
- Modèles d’agent
- Les jeux coopératifs,
- L’agrégation de jugements,
- Choix social computationnel,
- Raisonnement sur la connaissance et planification basée sur la connaissance

L’équipe AIDR du LAMSADE est impliquée en tant que partenaire dans divers projets ANR :

- ANR Project AMANDE - Advanced Multilateral Argumentation for DELiberation with CRIL, LIP6 and IRIT (Project coordinator : Nicolas Maudet). Start : December 2013. Duration : 4 years.
- ANR Project LARDONS - Learning And Reasoning for Deciding Optimally using Numerical and Symbolic information with GREYC, LIP6, LAMSADE and UBIA. Start : December 2010. Duration : 4 years.
- GDRI Algodex (Algorithmic Decision Theory) International Research Group on Algorithmic Decision Theory.

Résultats majeurs

Décision automatisée dans de grands espaces d’états

Concernant la recherche heuristique dans les grands espaces d’états, le résultat principal est la publication d’un nouvel algorithme de recherche arborescente Monte-Carlo : la recherche Monte-Carlo imbriquée. Le principe est de faire une recherche Monte-Carlo à plusieurs niveaux, chaque niveau utilisant le niveau en dessous pour faire des explorations aléatoires. Cet algorithme s’applique aux problèmes d’optimisation et aux jeux à un joueur. Il a été appliqué avec succès au voyageur de commerce avec fenêtres temporelles, au Morpion Solitaire et à SameGame (entre autres). Une seconde thématique concerne les algorithmes pour les jeux à deux joueurs.

Thèse en cours (Marek Cornu) : Recherche Monte-Carlo appliquée aux problèmes d’optimisation combinatoire multi-objectifs.

Thèse en cours (Alexandre Menif) : Planification pour les serious games.

DOSSIER "PRÉSENTATION D'ÉQUIPES DE RECHERCHE EN I.A."

Modèles d'interaction

Concernant les systèmes multi-agent (SMA) les points forts de nos recherches s'articulent autour de la coordination multi-agent et de l'interaction (communication, négociation) entre agents. La théorie de la décision (théorie multicritère, théorie de l'utilité, analyse des préférences) propose un cadre conceptuel original et intéressant dans la modélisation de la coopération, de la prise de décision et du contrôle dans les SMA. De nouveaux modèles d'interaction ont été développés et appliqués à différents domaines : systèmes d'aide à la décision dédiés au transport dans le cadre des travaux de Flavien Balbo et de la coopération avec l'IFSTTAR (SATIR : aide à la régulation de réseaux de bus, SAIV : aide à la décision pour le voyageur, Lacios-TAD : transport à la demande, SGC : aide à la décision en situation de crises, allocation en ligne de ressources locales : gestion temps réel et distribuée du stationnement urbain, gestion dynamique des feux de croisement et du trafic bi-modal), modèle d'enchères automatiques dans le commerce électronique, (stratégies de fixation de l'incrément sur le déroulement des enchères), coordination de services Web (modification dynamique de la coordination de services web). Dans ce domaine, l'importante augmentation du nombre de services disponibles dans les applications sociales (par exemple les réseaux sociaux) a engendré un nouveau défi à la communauté "services web" : la sélection de services à partir des réseaux sociaux plutôt qu'à partir de registres (par exemple, UDDI, ebXML). Nous proposons une approche multiagent basée sur la confiance pour une découverte distribuée de services dans les réseaux sociaux. Pour ce faire, nous définissons une mesure de confiance comme un concept bidimensionnel qui comprend (i) une confiance en la sociabilité indiquant si le fournisseur de services est socialement digne de confiance et (ii) une confiance en l'expertise évaluant la qualité des services offerts.

Thèse en cours (Amine Louati) : Une approche multi-agents basée sur la confiance pour la découverte et la coordination de services Web dans des réseaux sociaux.

Thèse en cours (Souhila Aknine-Arib) : Formations de coalitions pour le covoiturage.

Décision collective

Le choix social computationnel (ou l'étude des questions algorithmiques en décision collective) est un champ de recherche à l'intersection de l'économie mathématique (théorie du choix social) et de l'intelligence artificielle. Problème de votes. On s'est intéressé à des questions liées au vote séquentiel sur des domaines combinatoires (référendums multiples ou élections de comité) en étu-

diant les propriétés des méthodes de vote séquentielles. On a également étudié des modèles autorisant une représentation compacte de préférences sur des domaines combinatoires, par exemple les CP-nets. On a enfin étudié la question de l'agrégation de préférences incomplètes avec, par exemple, la détermination de « vainqueurs possibles » et de « vainqueurs nécessaires » Partage équitable de ressources. Ce thème est étudié, d'une part, du point de vue de la représentation compacte des préférences et de l'algorithmique et, d'autre part, sans exiger une élicitation préalable des préférences. Agrégation de jugements. Les aspects algorithmiques de l'agrégation de jugements exprimés en logique propositionnelle est un champ que nous avons commencé à aborder récemment.

Thèse en cours (Nathanael Barrot) : Aspects algorithmiques des élections à vainqueur multiples.

Thèse en cours (Nicolas Paget) : Le partage d'information pour la gestion des ressources naturelles.

Principales publications

- Saunier J., Balbo F. and Pinson S., A Formal Model of Communication and Context Awareness in Multi-agent Systems, *Journal of Logic, Language and Information*, 23(), 2014. (in press)
- Airiau S., Sen S. and Villatoro D., Emergence of conventions through social learning, *Autonomous Agents and Multi-Agent Systems 0()*, 2014. (in press)
- Airiau S., Endriss U., Multiagent resource allocation with sharable items, *Autonomous Agents and Multi-Agent Systems 0()*, 2014. (in press)
- Boutilier C., Lang J., Oren J. and Palacios H., Robust Winners and Winner Determination Policies under Candidate Uncertainty, *AAAI-2014*
- Louati A., EL Haddad J. and Pinson S., A Multilevel Agent-based Approach for Trustworthy Service Selection in Social Networks., *The 2014 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT 2014)*.
- Louati A., EL Haddad J. and Pinson S., Towards Agent-based and Trust-oriented Service Discovery Approach in Social Networks, *AAMAS Workshop TRUST2014*. , In *Proceedings CEUR "Trust in Agent Societies"*, pages 1-12, 2014.
- Boella B. and Pigozzi G., Normative Multiagent Systems : Guest Editors' Introduction, *Journal of Logic and Computation* 23(2), 2013.
- Cazenave C., Monte-Carlo Expression Discovery. *International Journal on Artificial Intelligence Tools* 22(1) (2013).

DOSSIER “PRÉSENTATION D’ÉQUIPES DE RECHERCHE EN I.A.”

- Wooldridge M., Endriss U., Kraus S. and Lang J., Incentive engineering for Boolean games., *Artificial Intelligence Journal* 195(418-439), 2013.
- Airiau S., Cooperative Games and Multiagent Systems, *The Knowledge Engineering Review* 28(4), pages 381–424, 2013.
- Saffidine A. and Cazenave C., Developments on Product Propagation, In *Proceedings Computers and Games 2013*, 2013.
- Bhouri N., Balbo F. and Pinson S., An Agent-Based Computational Approach for Urban Traffic Regulation, *Progress in Artificial Intelligence* 1(2), pages 139-147, 2012.
- Grossi D. and Pigozzi G., Introduction to Judgment Aggregation, In *Proceedings ESSLLI’11 Lecture Notes on Logic and Computation*, pages 160-209, 2012.
- Badeig F., Balbo F. and Pinson S., A contextual environment approach for multi-agent-based simulation, In *Proceedings 2nd International Conference on Agents and Artificial Intelligence (ICAART 2010)*, pages 212-217, 2010.
- Balbo F. and Pinson S., Using intelligent agents for Transportation Regulation Support System Design, *Transportation Research Part C : Emerging Technologies* 18(1), pages 140-156, 2010. "TOP25 Hot-test Articles in Decision Science" cité par Science Direct.
- Cazenave C., Balbo F. and Pinson S., Monte-Carlo Bus Regulation, In *Proceedings 12th Intern. IEEE Conf. on Intelligent Transportation Systems, ITSC’09*, pages 340-345, 2009.

Principaux livres et chapitres de livres

- Grossi D. and Pigozzi G., *Judgment Aggregation : A Primer*, Morgan & Claypool, 2014.
- Tsoukiàs A., *Aiding to decide : concepts and issues*, Book chapter, pages to appear, to appear. *Evaluation and Decision Models : real case studies*, Springer Verlag. Raymond Bisdor , Luis Dias, Vincent Mousseau, Marc Pirlot, editor(s)
- Bourgne G., Maudet N. and Pinson S., *Hypothesis Refinement : Building Hypotheses in an Intelligent System (ch.6)*, *Handbook on Reasoning-Based Intelligent Systems*, World Scientific Publishing Co, 2011
- Cazenave T., *Intelligence Artificielle une Approche Ludique*, Ellipses, 2011

MAGMA - Grenoble

(*Equipe Systèmes Multi-Agents*)

LIG (Laboratoire d’Informatique de Grenoble) UMR
5217

Maison Jean Kuntzmann

110 rue de la Chimie

Domaine universitaire

38400 Saint-Martin-d’Hères

<http://magma.imag.fr/>

Responsable : Julie Dugdale

mel : Julie.Dugdale@imag.fr

tel : 04 76 51 46 44

Chercheurs :

- Carole Adam
- Yves Demazeau
- Julie Dugdale
- Humbert Fiorino
- Damien Pellier
- Sylvie Pesty

Mots-clés :

- Comportements sociaux,
- modélisation du comportement humain,
- compagnons artificiels,
- algorithmes de planification sociaux,
- passage à l’échelle,
- robotique sociale,
- logiques BDI,
- simulation sociale,
- privacité,
- intelligence collective,
- formalisation des émotions.

Thématique de recherche

Thèmes de Recherche

Le projet de recherche de l’équipe MAGMA concerne les systèmes multi-agents socialement pertinents. Le terme système multi-agents (SMA) a été créé pour désigner une collection d’agents logiciels qui interagissent dans un environnement informatique. Très rapidement, le point de vue

DOSSIER “PRÉSENTATION D’ÉQUIPES DE RECHERCHE EN I.A.”

de la communauté des SMA a été étendu à d’autres types d’entités telles que les personnages virtuels, les robots et les humains, toutes en interaction dans une interprétation plus large de ce que peut être un système. A présent, notre vue est celle d’un système socio-technique composé d’une multitude de différents types d’agents, interagissant mutuellement pour le bénéfice de chacun et de tous. Dans ce nouveau système social, le terme d’utilité prend un sens nouveau, se référant à quelque chose qui a de la valeur pour la communauté mixte des agents dans son ensemble, ou pour un sous-ensemble de celle-ci.

Dans cette perspective, MAGMA se concentre sur l’étude, le développement et l’utilisation d’agents autonomes (AA) et de SMA pour résoudre les problèmes qui touchent à un problème ou à une question d’intérêt social. Son objectif est de produire des AA et des SMA utiles à la population, à la résolution des problèmes du monde réel et de la société. En se concentrant sur les AA et les SMA qui ont une utilité individuelle et sociale, MAGMA détermine de nouveaux rôles pour les AA et SMA dans la société. Pour étudier et développer des AA et des SMA qui soient socialement pertinents, nous devons répondre à un certain nombre de défis.

Challenges scientifiques

- Comment analyser, modéliser et reproduire le raisonnement humain, les comportements sociaux et les facteurs émotionnels,
- Quelles sont les formalisations et représentations appropriées pour les agents artificiels,
- Comment personnaliser des agents artificiels et des groupes d’agents,
- Quel est le socle des relations à long terme entre les hommes et leurs compagnons artificiels (robots sociaux ou personnages virtuels),
- Comment créer et valider nos modèles à base d’agents afin qu’ils soient bien fondés dans la réalité ou dans la théorie,
- Comment comprendre la société comme un système complexe, comprendre l’émergence de phénomènes collectifs et de l’intelligence collective, et concevoir des systèmes socio-techniques robustes et résistants,
- Comment développer des algorithmes de planification adaptatifs et conscients de la présence d’être humains,
- Comment concevoir des SMA de grande échelle ou pouvant passer à l’échelle,
- Comment garantir la vie privée et la confidentialité des données des agents,

- Comment observer microscopiquement et macroscopiquement des systèmes naturels et artificiels.

Approche

MAGMA s’intéresse au développement d’AA et de SMA qui soient ancrés dans le monde social. Cela nous conduit à suivre une approche multidisciplinaire, particulièrement avec les chercheurs d’autres disciplines, et à tenir compte activement des utilisateurs finaux et des parties prenantes. Les approches que nous utilisons pour relever ces challenges ont trait à :

- Simulation sociale orientée agent,
- Intelligence Artificielle et simulation du comportement,
- Planification distribuée / temps réel / continue / interactive,
- Simulation participative,
- Informatique affective, interaction intelligente,
- Agents virtuels intelligents, robotique sociale,
- Modélisation cognitive et architectures BDI (Beliefs, Desires, Intentions),
- Formalisation logique de théories psychologiques et sociologiques,
- Analyse spatio-temporelle de gros flux de données orientées agent,
- Analyse multi-échelle et visualisation d’information.

Domaines d’application

MAGMA s’intéresse au développement d’applications pratiques d’AA et de SMA conçues pour contribuer au bien-être de la société dans les domaines de :

- Gestion de l’énergie (énergie intelligente et éco-comportements),
- Gestion de crises et de situations d’urgence,
- Agents compagnons pour la vie quotidienne,
- Couplage robot et chien au service de l’aide à la mobilité humaine,
- Web intelligence, habitat numérique, villes intelligentes,
- Analyse de media de communication (News, blogs, media sociaux, etc.)

Projet de recherche

Nous entrons dans une nouvelle ère de l’informatique. Les agents logiciels sont de plus en plus répandus dans notre environnement ; ils sont intégrés dans le tissu de notre société à travers les technologies ambiantes, ils se

DOSSIER “PRÉSENTATION D’ÉQUIPES DE RECHERCHE EN I.A.”

retrouvent dans les supports technologiques que nous utilisons, et se manifestent à travers des personnages virtuels conversationnels ou des robots sociaux. Ceci conduit à une augmentation inévitable des interactions entre agents logiciels et humains, et à l’émergence de sociétés mixtes humains-agents.

Actuellement on sait très peu sur les interactions qui auront lieu dans ces nouvelles sociétés, quels types de comportements coopératifs s’y développeront, ou encore comment les comportements des humains et des agents intelligents seront modifiés par ces interactions. Nous savons que les humains se comportent inconsciemment face aux médias numériques comme s’ils étaient en interaction avec d’autres êtres humains. Pourrait-il en être de même lorsque les humains interagissent avec des agents intelligents? Les humains pourraient-ils présenter des comportements pro-sociaux à l’égard des agents logiciels comparables avec ceux qu’ils présentent avec les humains? Quels genres de comportements émergents seront observés dans les sociétés mixtes? Les agents artificiels seront-ils en mesure d’identifier et de réagir à ces phénomènes comme le font les gens dans les sociétés humaines? Les humains font déjà confiance aux agents en tant que soutien cognitif dans leur environnement : les agents artificiels se développeront-ils pour répondre à cette demande?

L’objectif de notre projet de recherche est donc double : d’une part, découvrir comment caractériser le comportement direct et indirect des humains et des agents, en modélisant et simulant ces sociétés mixtes; d’autre part, du point de vue des agents artificiels, explorer la façon dont les agents se comportent dans ces sociétés mixtes. Caractériser et comprendre l’interaction que les humains ont envers les agents logiciels embarqués dans les nouvelles technologies, nous permettra de construire des systèmes multi-agents qui agissent en synergie et en collaboration avec les humains.

Dans la pratique, nous nous efforçons de maintenir et d’accroître la cohérence scientifique de l’équipe. Cela implique par exemple de regrouper les travaux sur la simulation sociale et les robots de compagnie. Bien que les deux aspects ont comme tronc commun la modélisation du comportement humain, nous nous efforçons d’aller plus loin et de voir comment les interactions hommes-robots peuvent être simulées en utilisant la simulation orientée agent. Cela nous permettra de modéliser et de simuler les interactions au sein de notre société mixte dans le but d’améliorer la conception de tels systèmes. En effet, la simulation a un rôle à jouer comme un outil important pour aider à la conception de nouvelles technologies (habitats intelligents, agents compagnons, gestion de crise, etc).

Depuis plusieurs années, nous travaillons sur la modélisation et la simulation du comportement humain dans le domaine de la gestion de crises, telles que les catastrophes à grande échelle comme les tsunamis, les tremblements de terre et les incendies de forêt. Ce domaine en lui-même apporte de nombreux défis intéressants, les crises étant un exemple parfait de système socio-technique complexe, où le raisonnement se déroule sous de sévères contraintes en temps et en ressources, et où l’environnement évolue rapidement. Ces dernières années le travail de MAGMA sur les agents compagnons et les robots a augmenté; dans l’avenir nous aimerions étendre ce travail au domaine de la robotique de sauvetage.

Projets marquants

MOCA : Mon petit MOnde de Compagnons Artificiels. 2012-2015 (ANR CONTINT).

Membres MAGMA : Sylvie Pesty (coordinateur), Wafa Johal, Carole Adam, Humbert Fiorino, Damien Pellier, Julie Dugdale.

Autres collaborateurs : IIIHM-LIG, Lab-STICC, Telecom ParisTech LTCI, LIMSI.

Le projet MOCA est un projet de recherche fondamentale qui étudie les compagnons artificiels (personnages virtuels et robots personnels) et leur valeur pour des usagers dans des situations de la vie de tous les jours.

SMART ENERGY : Approches et outils pour la gestion de l’énergie. 2012-2015.

Un projet fédérant 4 laboratoires et financé par G-INP.

Membres MAGMA : Julie Dugdale (coordinateur), Humbert Fiorino, Ayesha Kashif.

Autres collaborateurs : IIIHM-LIG, PRIMA-LIG, GIPSA, G2E Lab, G-SCOP.

Ce projet porte sur la nouvelle génération des systèmes énergétiques. Nous développons des démonstrateurs multi-échelles et multi-formes, des systèmes de simulation qui modélisent l’ampleur et la diversité des acteurs impliqués dans la chaîne d’approvisionnement et de consommation d’énergie.

COCHISE. Coopération entre chiens d’assistance et robots (ANR 2013-2016)

(Membres MAGMA : Sylvie Pesty (coordinateur), Humbert Fiorino, Damien Pellier. Autres collabora-

DOSSIER “PRÉSENTATION D’ÉQUIPES DE RECHERCHE EN I.A.”

teurs : CHU Clermont-Ferrand, Université d’Angers, Lab-STICC).

Les objectifs du projet Cochise sont d’accroître les connaissances à propos des interactions dans un système animal / robot et de parvenir à un développement expérimental qui pourrait rapidement devenir un produit grand public.

GEOMEDIA : Observatoire des flux géomédias-tiques internationaux (ANR Corpus 2012-2015)

(Membres MAGMA : Yves Demazeau (coordinateur).
Autres collaborateurs : MESCAL-LIG, CIST, Géographie-Cités, GERiiCO, IDEES Riate, Density Design Lab Milan, INED).

L’objectif du projet est de mener une recherche conjointe sur les échanges contemporains de flux d’information médiatique entre pays du monde à travers les flux d’informations diffusés par les organes de presse.

VUSIM-AMS : Vulnérabilité urbaine aux tremblements de terre à Mendoza (2014-2015)

(Membres MAGMA : Julie Dugdale, Carole Adam.
Autres collaborateurs : PACTE Lab.)

Le projet porte sur l’acquisition de données pour la modélisation de comportements humains juste après un tremblement de terre. Il s’agit de déterminer la vulnérabilité physique des bâtiments de gestion de crise, et de développer un modèle multi-agent de mobilité humaine, de contribuer au développement de plans d’évacuation plus efficaces, d’évaluer les effets d’une campagne d’information du public, etc.

10 publications significatives

- Robin Lamarche-Perrin, Yves Demazeau, and Jean-Marc Vincent. Building Optimal Macroscopic Representations of Complex Multi-Agent Systems. Application to the Spatial and Temporal Analysis of International Relations through News Aggregation. *Transactions on Computational Collective Intelligence*, 15, pp. 1-27, LNCS 8670, Springer Verlag, 2014.
- Ayesha Kashif, Julie Dugdale, and Stéphane Ploix. Simulating Occupants’ Behaviour for Energy Waste Reduction in Dwellings : A Multi Agent Methodology. *Advances in Complex Systems*, 16(04n05) :1350022 :1-37, May 2013.
- Ayesha Kashif, Stéphane Ploix, Julie Dugdale, and Le Xuan Hoa Binh. Simulating the dynamics of occupant behaviour for power management in residential buildings. *Energy and Buildings*, 56 :85-93, January 2013.
- Bartel Van DeWalle and Julie Dugdale. Information management and humanitarian relief coordination : findings from the Haiti earthquake response. *International Journal of Business Continuity and Risk Management*, 3(4) :278-305, 2012.
- Guillaume Piolle and Yves Demazeau. Representing privacy regulations with deontico-temporal operators. *Web Intelligence and Agent Systems*, 9(3) :209-226, IOS Press, July 2011.
- Wafa Johal, Gaëlle Calvary, and Sylvie Pesty. A Robot with Style, because you are Worth it! In *ACM CHI Conference on Human Factors in Computing Systems (CHI’14)*, Toronto, Canada, 2014. ACM. Toronto, Canada.
- Adrien Dulac, Damien Pellier, Humbert Fiorino, and David Janiszek. Learning Useful Macroactions for Planning with N-Grams. In *IEEE International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 803-810, Herndon, Etats-Unis, 2013.
- Damien Pellier, Humbert Fiorino, and Marc Métivier. A new approach for continual planning. In *AAMAS*, pages 1115-1116, Saint-Paul, Etats-Unis, 2013.
- Jérémy Rivière, Carole Adam, and Sylvie Pesty. A Reasoning Module to Select ECA’s Communicative Intention. In Ana Paiva Marilyn Walker Yukiko Nakano, Michael Neff, editor, *IVA 2012 - Proceedings of the 12th International Conference on Intelligent Virtual Agents*. Volume 7502 of *Lecture Notes in Computer Science (LNCS)*, pages 447-454, Santa Cruz, CA, Etats-Unis, 2012. Springer Berlin / Heidelberg.
- Carole Adam, Benoit Gaudou, Dominique Longin, and Emiliano Lorini. Logical modeling of emotions for Ambient Intelligence. In Fulvio Mastrogiovanni and Nak-Young Chong, editors, *Handbook of Research on Ambient Intelligence : Trends and Perspectives*. IGI Global Publisher, 2011.

Mnemosyne - Bordeaux

(*Mnemonic Synergy*)

Equipe-Projet Commune Inria - LaBRI (CNRS, Université de Bordeaux), hébergée à l’IMN - Institut des

DOSSIER “PRÉSENTATION D’ÉQUIPES DE RECHERCHE EN I.A.”

Maladies Neurodégénératives, 146 rue Léo Saignat,
33076 Bordeaux

<https://team.inria.fr/mnemosyne/fr/>

Responsable : Frédéric ALEXANDRE

mel : Frederic.Alexandre@inria.fr

tel : (33/0) 5 47 30 42 60

Chercheurs :

- Frédéric Alexandre, DR Inria
- André Garenne, MCF U. Bordeaux
- Nicolas Rougier, CR Inria
- Thierry Viéville, DR Inria

Mots-clés :

- Réseaux de neurones artificiels, approches neuromimétiques
- Sciences cognitives et Intelligence Artificielle, cognition
- Robotique, vision par ordinateur, capteurs intelligents, systèmes physiques

Thématique de recherche

Aux frontières des neurosciences intégratives et computationnelles, nous proposons de modéliser le cerveau comme un système de mémoires actives en synergie et en interaction avec les mondes interne et externe et de le simuler comme un tout et en situation.

Les principales fonctions cognitives et comportementales (ex : attention, reconnaissance, planification, décision) émergent de boucles sensorimotrices impliquant le monde extérieur, le corps et le cerveau. Nous étudions, modélisons et implémentons ces boucles et leurs interactions dans la perspective d’un comportement complètement autonome. Avec cette approche “systémique”, nous indiquons que de tels systèmes complexes ne peuvent être vraiment appréhendés que comme un tout et dans des situations comportementales naturelles. Pour mettre au point les caractéristiques fonctionnelles et adaptatives de tels modèles au niveau de leur circuiterie neuronale et pour les implémenter dans des systèmes interagissant avec le monde, nous combinons les principes, les méthodes et les outils de différents domaines scientifiques.

- Nous modélisons les principales structures cérébrales et les principaux flux d’information du cer-

veau (comme en neurosciences intégratives et cognitives) en insistant sur les liens entre le cerveau, le corps et l’environnement (cognition incarnée).

- Nous utilisons des formalismes de calcul distribué nous permettant d’implanter des modèles à différentes échelles de description (comme en neurosciences computationnelles).
- Nous déployons nos modèles à large échelle (calcul à haute performance), les incarnons dans des corps en interaction avec l’environnement (robotique autonome) et nous les simulons de façon interactive avec les événements rencontrés par un robot virtuel/réel.

En retour, nous pensons pouvoir proposer à ces différents domaines une approche intégrative innovante et pensons pouvoir aussi contribuer à d’autres domaines connexes dans les sciences du vivant (neurosciences, médecine) et les sciences du numérique (informatique, apprentissage automatique).

Thèmes de recherche :

Par notre approche systémique, nous sommes donc principalement intéressés par le développement de boucles entre le cerveau, le corps et le monde extérieur et de leurs interactions pour la réalisation d’un comportement autonome. Nous portons particulièrement attention au fait que différentes formes de mémoires sont élaborées dans ces boucles et, par conséquent, que les phénomènes observés proviennent aussi d’échanges d’information entre ces mémoires.

- Dans ce que nous appelons “Boucles Perception”, nous explorons les circuits cérébraux qui permettent d’identifier des phénomènes dans les mondes internes (corporels, neuronaux) et externes (sensoriels), afin d’évaluer l’état courant. Différentes formes d’apprentissage sont impliquées dans l’acquisition et la représentation de l’information permettant de discriminer des états (indices, configurations, séquences, etc.) qui pourraient participer à la sélection des actions futures. Les structures modélisées correspondent à l’amygdale, à l’hippocampe et le cortex sensoriel.
- Dans ce que nous appelons “Boucles Action”, l’analyse perceptive évoquée plus haut est exploitée par des territoires moteurs du cerveau pour déclencher des actions dans les mondes interne (corporel, neuro-hormonal, décisionnel) et externe (moteur). Ces territoires incluent les ganglions de la base et le cortex préfrontal.

De façon importante, il doit être souligné que les actions déclenchées par les Boucles Action influencent les Boucles Perception indirectement par leurs conséquences

DOSSIER "PRÉSENTATION D'ÉQUIPES DE RECHERCHE EN I.A."

dans le monde externe et aussi directement par les activations neuronales des Boucles Perception (attention, activité soutenue, inhibition de contrôle). Réciproquement, les décisions prises par les Boucles Actions sont en partie fondées sur des identifications et des évaluations (prédiction de punition et de récompense, identification de catégories sensorielles, rappels d'épisodes) dans les Boucles Perception, formant ainsi les bases principales des interactions entre ces boucles.

Projets marquants

- modèle d'amygdale, pour le conditionnement répondant (collaboration avec l'INSERM)
- modèle d'hippocampe pour une mémoire épisodique résistante aux interférences (collaboration avec le CNRS)
- modèle de ganglions de la base pour la sélection de l'action (projet Cefipra-CNRS avec les IIT de Hyderabad et de Madras, Inde)
- modèle du système de vision précoce, rétine, thalamus, cortex (projet ANR international et équipe commune avec l'université de Valparaiso et l'Université Santa Maria, Chili)
- modèle de cortex préfrontal pour l'organisation de l'action (collaboration avec l'université de Boulder, Colorado, USA)

10 publications significatives

- L. ALECU, H. FREZZA-BUET, F. ALEXANDRE. Can Self-Organization Emerge through Dynamic Neural Fields Computation?, in "Connection Science", 2011, vol. 23, no 1, p. 1-31
- A. O. ARDILES, C. C. TAPIA-ROJAS, M. MANDAL, F. ALEXANDRE, A. KIRKWOOD, N. C. INESTROSA, A. G. PALACIOS. Postsynaptic

dysfunction is associated with spatial and object recognition memory loss in a natural model of Alzheimer s disease., in "Proceedings of the National Academy of Sciences", August 2012, vol. 109, no 34, p. 13835-40

- M. CARRERE, F. ALEXANDRE. Émergence de catégories par interaction entre systèmes d'apprentissage, in "Conférence Francophone sur l'Apprentissage Automatique (CAP)", Lille, France, P. PREUX, M. TOMMASI (editors), July 2013.,
- C. CARVAJAL, T. VIÉVILLE, F. ALEXANDRE. Impact of the Konio pathway in the thalamocortical visual system : a modeling study, in "Twenty Second Annual Computational Neuroscience Meeting : CNS 2013", Paris, France, July 2013,
- G. DETORAKIS, N. P. ROUGIER. A Neural Field Model of the Somatosensory Cortex : Formation, Maintenance and Reorganization of Ordered Topographic Maps, in "PLoS ONE", July 2012, vol. 7, no 7, e40257
- J. FIX, N. P. ROUGIER, F. ALEXANDRE. A dynamic neural field approach to the covert and overt deployment of spatial attention, in "Cognitive Computation", 2011, vol. 3, no 1, pp. 279-293.
- M. GUTHRIE, A. LEBLOIS, A. GARENNE, T. BORAUD. Interaction Between Cognitive and Motor Cortico- Basal Ganglia Loops During Decision Making : A Computational Study, in "Journal of Neurophysiology", March 2013.,
- N. P. ROUGIER, J. FIX. DANA : Distributed (asynchronous) Numerical and Adaptive modelling framework, in "Network : Computation in Neural Systems", December 2012, vol. 23, no 4, pp. 237-253.
- N. ROUGIER, A. HUTT. Synchronous and Asynchronous Evaluation of Dynamic Neural Fields, in "J. Diff. Eq. Appl.", 2009.
- W. TAOUALI, T. VIÉVILLE, N. ROUGIER, F. ALEXANDRE. No clock to rule them all, in "Journal of Physiology", 2011, vol. 105, no 1-3, pp. 83-90.

Prix de thèse AFIA 2014 (1er *ex aequo*)

*Conjunctive Query Answering under Existential Rules
Decidability, Complexity and Algorithms*

Michaël Thomazo¹ (TU Dresden, Germany)

La publication massive de données laisse entrevoir leur utilisation par tous. Celle-ci a déjà commencé, avec de nombreuses applications, que ce soit pour la gestion de flux automobiles (TrafficLarKC, [27]) ou pour la création d'évaluations par des particuliers (Revyu, [11]), pour n'en citer que deux. Cependant, la simple publication des données n'est pas suffisante pour permettre une exploitation efficace de celles-ci. En effet, chaque entité publiant des données utilise un vocabulaire spécifique, rendant difficile l'interopérabilité entre de multiples jeux de données. De plus, les données publiées sont probablement incomplètes, et ce pour de multiples raisons. Premièrement, l'éditeur de données peut avoir une connaissance seulement imparfaite du domaine qu'il cherche à décrire. Mais même dans le cas idéal où le domaine d'intérêt aurait été parfaitement décrit, cela n'empêche pas un utilisateur de vouloir exploiter ces données dans un cadre initialement non prévu – la définition d'un nouveau vocabulaire rend inévitablement les données initiales incomplètes.

Dans ce double objectif de permettre l'utilisation simultanée de plusieurs jeux de données et de compléter des données potentiellement incomplètes, des *ontologies* ont été introduites. Ce mot possède de multiples acceptions, et il sera utilisé exclusivement dans le sens suivant : une *ontologie* est une théorie logique représentant les connaissances d'un domaine. Le problème que nous considérons est le suivant : soit F un ensemble de données, q une requête, et \mathcal{O} une ontologie, est-il vrai que q est une conséquence logique de F et de \mathcal{O} ?

Restreignons un peu le cadre. Quelles requêtes étudier ? Quelles ontologies considérer ? Les requêtes conjonctives étant les requêtes de bases dans la communauté base de données, elles forment donc une cible naturelle. La question pour les ontologies est un peu plus complexe. En effet, l'approche principale pour la représentation d'ontologies dans la communauté représentation des connaissances et raisonnement est celle des logiques de description. Leurs éléments de base sont des concepts (prédicats unaires)

et des relations (binaires). Jusqu'à peu, les problèmes étudiés à propos des DLs étaient centrés sur l'ontologie elle-même : un concept est-il satisfiable, c'est-à-dire, peut-il avoir une interprétation non vide ? Est-il vrai que toute instance du concept A est également instance du concept B ? Alors que ces logiques sont bien connues, elles imposent des restrictions sur les vocabulaires utilisés. La restriction d'arité, déjà mentionnée, rend l'intégration directe à une base de données plus complexe. Plus gênant, la forme des connaissances représentables est restreinte à des formes arborescentes : il n'y a pas de possibilité d'exprimer des cycles entre variables. Ceci est dommageable dans certaines modélisations, comme par exemple en chimie [10], où les structures cycliques tiennent une place importante. Pour ces raisons, nous utilisons ici un autre formalisme pour représenter les ontologies : les règles existentielles.

Règles existentielles

Celle-ci ont été introduites sous différents noms, et proviennent de plusieurs communautés. Sous le nom de tuple-generating dependencies, elles forment la base d'un cadre appelé Datalog \pm [8], avec une vue base de données. Elle dérivent également des règles pour les graphes conceptuels [23, 1]. Une règle existentielle est une formule logique de la forme :

$$\forall \mathbf{x} \forall \mathbf{y} (\mathbf{B}[\mathbf{x}, \mathbf{y}] \rightarrow \exists \mathbf{z} \mathbf{H}[\mathbf{y}, \mathbf{z}]),$$

où \mathbf{B} et \mathbf{H} sont des conjonctions d'atomes, respectivement appelées corps (\mathbf{B} pour *body*) et tête (\mathbf{H} pour *head*) de la règle. D'un point de vue syntactique, ces règles ressemblent à des règles Datalog : l'unique différence est la présence de variables existentiellement quantifiées en tête de règle. Ces variables ont une importance capitale. D'un point de vue modélisation, elles permettent d'intro-

1. La thèse présentée ici a été réalisée à l'Université Montpellier 2, sous la direction de Jean-François Baget et Marie-Laure Mugnier, au sein de l'équipe INRIA GraphIK.

duire des individus qui n'étaient initialement pas présents dans la base de données (*value invention*). D'un point de vue computationnel, elles complexifient le problème d'interrogation.

Approches classiques

À quel point ? Le problème devient en fait indécidable [5], même lorsque des contraintes fortes sont imposées sur les règles [2]. Il existe cependant des procédures de semi-décidabilité. Le point commun entre ces deux procédures est de se ramener à un problème classique d'interrogation sur une base de données.

Matérialisation

La première consiste à ajouter des données à l'aide des règles, jusqu'à ce que la requête soit conséquence des faits matérialisés, ou que tout ajout soit redondant. Illustrons l'approche sur un exemple simple. Nous savons qu'un chat, *Garfield*, dort sur un canapé. Autrement dit, nous avons les données suivantes :

$$\exists x \text{ dortSur}(\text{Garfield}, x) \wedge \text{canapé}(x) \wedge \text{chat}(\text{Garfield}).$$

De plus, nous savons que tout chat est un animal, ce qui se traduit par la règle suivante :

$$\text{chat}(x) \rightarrow \text{animal}(x).$$

Nous voulons savoir s'il existe un animal qui dort sur un canapé. En appliquant la règle sur le fait, nous pouvons ajouter $\text{animal}(\text{Garfield})$. Nous pouvons maintenant répondre par l'affirmative à notre requête.

Réécriture

La deuxième approche consiste en une réécriture de la requête. Au lieu de modifier les données, en les complétant grâce aux règles, nous modifions la requête. De manière évidente, nous pouvons répondre par l'affirmative à notre requête si dans les données il existe un animal dormant sur un canapé. Mais c'est également le cas s'il est indiqué qu'un chat dort sur un canapé. La requête initiale

$$\exists x \exists y \text{ animal}(x) \wedge \text{dortSur}(x, y) \wedge \text{canapé}(y)$$

est donc complétée par une deuxième requête :

$$\exists x \exists y \text{ chat}(x) \wedge \text{dortSur}(x, y) \wedge \text{canapé}(y).$$

Comparaison rapide des deux approches

Ces deux approches ne s'appliquent pas aux mêmes ensembles de règles – l'une peut terminer alors que l'autre non. Même dans les cas où les deux approches sont applicables, le choix n'est pas forcément évident. En effet, la matérialisation de nouveaux atomes peut prendre du temps et nécessite des droits d'écritures. De plus, lorsque les données sont changées, les méthodes de mise à jour ne sont pas évidentes. Ces difficultés sont éludées par la réécriture, mais celle-ci génère typiquement des réécritures de grande taille.

Contributions de la thèse

Les deux principales contributions de la thèse s'attachent chacune à l'une de ces procédures de semi-décidabilité.

Matérialisation et treewidth

Il est connu que le problème d'interrogation est décidable lorsque les faits générés par matérialisation ont une structure simple, c'est-à-dire ici de treewidth bornée [22]. Cependant, aucun algorithme n'est connu dans ce cas général. La première contribution a donc été de définir un nouveau critère, couvrant les cas pratiques connus, et de proposer un algorithme générique pour ces cas [3]. L'idée centrale est de travailler directement sur une décomposition arborescente des faits, et de détecter des régularités au sein de cette structure, infinie, mais simple. En "factorisant" la structure selon ces régularités, nous obtenons une représentation finie de celle-ci. Cela induit quelques difficultés d'ordre technique au moment de l'interrogation : alors que l'algorithme proposé possède de bonnes propriétés théoriques [26], du travail est encore à faire pour obtenir un algorithme efficacement implémentable [25].

Réécriture et disjonction

La deuxième contribution [24] concerne l'approche de réécriture. La plupart des algorithmes et des prototypes développés pour la réécriture de requêtes conjonctives l'ont été pour des classes de règles restreintes, et dans le but de générer des unions de requêtes conjonctives. Ces réécritures peuvent être de taille exponentielle dans le pire des cas, même en utilisant des formules logiques du premier ordre. Cependant, utiliser des unions de requêtes

conjonctives comme langage de réécriture rend ce pire des cas courant : en effet, les ontologies existantes consistent généralement en un large “squelette” de hiérarchies de classes ou de relations. Or, ces hiérarchies font exploser la taille des réécritures lorsque des unions de requêtes conjonctives sont utilisées. L’algorithme proposé est donc le suivant : au lieu d’utiliser des requêtes conjonctives comme requête de base, nous ajoutons un niveau de disjonction. En lieu et place des deux requêtes conjonctives présentées ci-dessus, nous utilisons la requête suivante :

$\exists x \exists y (chat(x) \vee animal(x)) \wedge dortSur(x, y) \wedge canapé(y).$

Adaptant l’algorithme de réécriture à base de pièces [12], une opération d’unification a été définie pour ce type de requêtes. Outre des réécritures plus petites, obtenues plus rapidement, l’algorithme proposé est assuré, tout comme son parent [12], de terminer dès qu’une réécriture existe, indépendamment de contraintes syntaxiques sur les règles utilisées. Il est donc en ce sens générique.

Développements futurs

De nombreux développements sont envisageables à partir de ces travaux. Pour la matérialisation, il serait intéressant d’exploiter les structures définies dans la thèse en présence de contraintes négatives. La sémantique de la logique du premier ordre étant inadaptée dans ce cas, de nouvelles sémantiques, plus complexes, ont été définies : comment combiner ces sémantiques et des règles expressives ? Pour la réécriture, des langages de réécriture plus expressifs permettraient de pouvoir utiliser cette approche aussi en présence de relations transitives. Enfin, les deux approches présentées sont duales : une meilleure compréhension de celles-ci devraient permettre leur unification.

Références

- [1] J.-F. BAGET et M.-L. MUGNIER. “The Complexity of Rules and Constraints”. In : *J. Artif. Intell. Res. (JAIR)* 16 (2002), p. 425–465.
- [2] J.-F. BAGET et al. “On Rules with Existential Variables : Walking the Decidability Line”. In : *Artif. Intell.* 175.9-10 (2011), p. 1620–1654.
- [3] J.-F. BAGET et al. “Walking the Complexity Lines for Generalized Guarded Existential Rules”. In : *IJCAI*. 2011, p. 712–717.
- [5] C. BEERI et M.Y. VARDI. “A Proof Procedure for Data Dependencies”. In : *Journal of the ACM* 31.4 (1984), p. 718–741.
- [8] A. CALÌ, G. GOTTLÖB et T. LUKASIEWICZ. “A general Datalog-based framework for tractable query answering over ontologies”. In : *J. Web Sem.* 14 (2012).
- [10] J. HASTINGS et al. “Structure-based classification and ontology in chemistry”. In : *J. Cheminformatics* 4 (2012), p. 8. DOI : [10.1186/1758-2946-4-8](https://doi.org/10.1186/1758-2946-4-8). URL : <http://dx.doi.org/10.1186/1758-2946-4-8>.
- [11] Tom HEATH et Enrico MOTTA. “Revyu : Linking reviews and ratings into the Web of Data”. In : *J. Web Sem.* 6.4 (2008), p. 266–273. DOI : [10.1016/j.websem.2008.09.003](https://doi.org/10.1016/j.websem.2008.09.003). URL : <http://dx.doi.org/10.1016/j.websem.2008.09.003>.
- [12] M. KÖNIG et al. “A Sound and Complete Backward Chaining Algorithm for Existential Rules”. In : *RR*. 2012, p. 122–138.
- [22] N. ROBERTSON et P. D. SEYMOUR. “Graph Minors. II. Algorithmic Aspects of Tree-Width”. In : *J. Algorithms* 7.3 (1986), p. 309–322.
- [23] J. F. SOWA. *Conceptual Structures : Information Processing in Mind and Machine*. Addison-Wesley, 1984. ISBN : 0-201-14472-7.
- [24] M. THOMAZO. “Compact Rewriting for Existential Rules”. In : *IJCAI*. 2013.
- [25] M. THOMAZO et S. RUDOLPH. “Mixing Materialization and Query Rewriting for Existential Rules”. In : *ECAI*. 2014, p. 897–902. DOI : [10.3233/978-1-61499-419-0-897](https://doi.org/10.3233/978-1-61499-419-0-897). URL : <http://dx.doi.org/10.3233/978-1-61499-419-0-897>.
- [26] M. THOMAZO et al. “A Generic Querying Algorithm for Greedy Sets of Existential Rules”. In : *KR*. 2012.
- [27] Emanuele Della VALLE et al. “Semantic Traffic-Aware Routing Using the LarKC Platform”. In : *IEEE Internet Computing* 15.6 (2011), p. 15–23. DOI : [10.1109/MIC.2011.107](https://doi.org/10.1109/MIC.2011.107). URL : <http://doi.ieeecomputersociety.org/10.1109/MIC.2011.107>.

Prix de thèse AFIA 2014 (1er *ex aequo*)

Solving Games and All That

Abdallah Saffidine (University of New South Wales, Sydney, Australia)

La résolution des jeux dans sa compréhension la plus générale, c'est à dire la *détermination des comportements optimaux et rationnels d'agents d'un système dynamique donné*, constitue un défi majeur pour la communauté d'Intelligence Artificielle.

Malgré des limitations computationnelles théoriques, des algorithmes très efficaces en pratique sont connus pour certains fragments de ce problème : la technique de l'élagage *alpha-beta*, cruciale dans l'élaboration de joueurs d'échecs artificiels d'un niveau supérieur aux meilleurs humains ; l'algorithme *Proof Number Search (PNS)* utilisé lors de la résolution des dames anglaises ; l'algorithme *Monte Carlo Tree Search (MCTS)* à l'origine de progrès importants dans le domaine du jeu de Go et sa variante pour la résolution des jeux *MCTS Solver*.

Si ces algorithmes ont déjà été utilisés dans des cadres non ludiques tels la synthèse chimique ou l'optimisation combinatoire, leur applicabilité reste limitée par les hypothèses très restrictives imposées au domaine d'application. Ainsi, ces trois algorithmes supposent un jeu séquentiel, à deux agents, à somme nulle et à information parfaite, et les deux derniers requièrent en plus qu'il n'y ait que deux issues possible à une partie : la *partie nulle*, les victoires partielles ou les scores intermédiaires étant impossibles.

L'objectif principal de la thèse est d'élargir le champs d'application de la recherche sur les jeux en montrant comment en généraliser les algorithmes principaux. Comme pour l'exploration de graphes, on peut distinguer des algorithmes de jeux en meilleur d'abord, tels A^* , PNS et MCTS, et des algorithmes en profondeur d'abord, tels alpha-beta et minimax. Les Chapitres 2 à 4 de la thèse se concentrent sur le premier type tandis que le Chapitre 5 s'intéresse à la deuxième catégorie.

Chapitre 2 : jeux à deux issues

Je m'attache ici à définir un cadre formel unificateur pour les algorithmes de résolution en meilleur d'abord

2. À titre d'exemple, la première hypothèse est $\top \cap \perp = \emptyset$.

(BFS) dans les jeux à deux issues. Pour cela, je définis le concept de *schéma d'information* et propose un algorithme en meilleur d'abord générique paramétré par un tel schéma d'information. Résoudre un jeu nécessite d'explorer l'arbre des futures actions possibles et le schéma d'information représente le type de données que l'on veut mémoriser au fur et à mesure de cette exploration.

Mathématiquement, un schéma d'information est un tuple $\langle V, \top, \perp, \preceq, H, \zeta \rangle$ tel que V est un ensemble de *valeurs d'information* ; $\top \subset V$ et $\perp \subset V$ sont deux sous-ensembles distingués, appelés *top* et *bottom* ; \preceq est un pré-ordre total sur V appelé *relation de sélection* ; H est une *fonction d'hérédité* agrégeant un multi-ensemble de valeurs d'information en une valeur d'information résultat ; et ζ est une *fonction d'initialisation* associant à chaque état du jeu une valeur d'information.

Pour un schéma d'information paramètre donné, l'algorithme BFS correspondant développe un arbre d'exploration à partir d'un état argument s . Afin d'orienter la recherche vers la preuve d'une victoire ou d'une défaite pour le premier joueur, on associe des valeurs d'information aux nœuds de l'arbre. Partant de la racine, une feuille est sélectionnée en utilisant la relation de sélection nœud après nœud. Une fois la feuille développée et ses enfants ajoutés à l'arbre, sa valeur d'information est mis à jour par la fonction d'hérédité. Une fois la valeur d'information de tous les nœuds traversés mis à jour, le processus recommence. Tant que la racine n'est pas résolue, c'est à dire tant que la valeur associée n'appartient ni à \top ni à \perp , l'arbre agrandi afin d'obtenir plus d'information.

Bien que BFS soit défini pour tout schéma d'information paramètre, il n'y a pas de garantie générale de convergence ou de correction de l'algorithme. Un schéma d'information est *bien formé* s'il satisfait un certain nombre d'hypothèses omises ici.² Il est alors possible d'obtenir le résultat suivant. *Si le schéma d'information est bien formé et le jeu fini et acyclique, alors l'algorithme termine ; toute position à valeur top offre une stratégie ga-*

gnante au premier joueur et toute valeur bottom offre une stratégie gagnante à l'adversaire.

Pour illustrer la valeur de ce cadre formel, je montre que les algorithmes PNS, MCTS Solver et *Product Propagation* (PP) peuvent être vus comme des applications de BFS à des schémas d'information bien choisis, et prouve que ces schémas d'information sont bien formés. Cela constitue une démonstration concise de la correction de ces algorithmes classiques.

En 2007, Stern *et al* ont proposé d'utiliser PP pour la résolution de jeux à deux issues mais en raison de résultats préliminaires décevants au Go, cet algorithme a reçu peu d'attention dans la communauté des jeux. La projection de PP dans un le même formalisme que PNS et MCTS suggère d'adapter quelques heuristiques et optimisations populaires avec ces algorithmes. Je montre expérimentalement que la performances de PP s'en trouve grandement augmentée dans 3 domaines. Dans chaque cas, PP se révèle plus efficace que les algorithmes plus populaires PNS, MCTS Solver, et alpha-beta.

Chapitre 3 : jeux multi-issues

La méthode la plus simple pour adapter un algorithme A destiné aux jeux à deux issues au cas multi-issues est d'appliquer A de manière itérative afin d'effectuer une recherche binaire parmi les issues. L'objectif de ce chapitre est de développer des techniques plus efficace que cette méthode naïve où les recherches successives ne partagent pas d'information. De même que dans le fragment à deux issues, il est possible de définir directement les concepts de schéma d'information, de schéma d'information bien formé, ainsi qu'un algorithme BFS paramétré dans les jeux multi-issues. La première contribution de ce chapitre est donc un cadre formel pour les algorithmes de résolution en meilleur d'abord dans les jeux multi-issues. Tout schéma d'information bien formé, non seulement génère un algorithme BFS convergent et correct, mais donne lieu à de l'élagage alpha-beta, ce qui permet une résolution plus rapide et une économie de mémoire. Cette analyse constitue la première étude de l'élagage alpha-beta en dehors du cadre de la recherche en profondeur d'abord.

Si de nombreux algorithmes de type *en meilleur d'abord* avaient été inventés pour les jeux à deux issues, ce n'est pas le cas pour la classe plus générale des jeux à issues multiples. Je propose alors une méthode systématique associant à tout schéma d'information du cadre deux issues un schéma d'information du cadre multi-issues. Le principal résultat théorique étant que *si le schéma d'information deux issues est bien formé, alors le schéma*

multi-issues correspondant est bien formé aussi. En s'appuyant sur les algorithmes existants et mentionnés au Chapitre 2, cette méthode donne immédiatement accès à un nombre d'algorithmes en meilleur d'abord, corrects par construction, pour les jeux multi-issues.

Afin de valider l'approche suggérée dans ce chapitre, je me penche sur le nouvel algorithme *Multiple-Outcome Proof Number Search* (MOPNS) dérivé automatiquement à partir du schéma d'information de PNS. Puisque le schéma d'information sous-tendant PNS est bien formé, il résulte que MOPNS est convergent et correct. J'établis ensuite expérimentalement que MOPNS est plus efficace que l'application itérative de PNS dans deux domaines, le jeu Puissance 4 qui présente 3 issues : {Gagné, Nulle, Perdu} et le jeu Woodpush qui admet comme issue un score entier entre -5 et 5.

Chapitre 4 : logique modale K

Ce chapitre explore le lien entre espace d'états d'un jeu à coups séquentiels et structure de Kripke. L'existence d'une stratégie gagnante en un nombre de coups borné peut être spécifiée comme une formule de logique modale K multi-agents (MMLK). Le problème de model checking en MMLK constitue donc une généralisation du domaine des jeux à deux issues, et les concepts de schéma d'information et d'algorithme en meilleur d'abord s'y prêtent tout à fait. De la même manière qu'avec les jeux multi-issues, il est possible de s'inspirer des algorithmes de résolutions des jeux à deux issues pour dériver des algorithmes de model checking en MMLK. Par exemple, les nouveaux algorithmes de model checking *Proof Number Proof Search* et *Monte Carlo Proof Search* généralisent PNS et MCTS Solver.

La résolution d'un problème de model checking consiste à renvoyer une preuve que le modèle satisfait la spécification ou alors un contre-exemple. Obtenir une preuve ou un contre-exemple de petite taille est parfois souhaité par le modélisateur. À cet effet, je propose un nouvel algorithme, *Minimal Proof Search* (MPS) et prouve que MPS est admissible et optimal. À ce titre MPS peut être interprété comme la généralisation de A* au model checking MMLK. Autrement dit, MPS renvoie une preuve ou un contre-exemple de taille minimale, et tout algorithme admissible explore un espace d'états au moins aussi grand que MPS.

Après avoir montré comment les intuitions, heuristiques et algorithmes issus de la communauté jeux pouvaient être utiles en logique modale, je m'attache à la relation inverse. La puissance expressive de MMLK per-

met en effet de spécifier d'autres concepts de solution typiques que l'existence d'une stratégie gagnante dans un jeu séquentiel. Entre autres, je donne des formules correspondant à l'existence de menaces, d'un mat réflexe aux Échecs, d'une victoire paranoïaque dans un jeu multi-joueurs, ou encore d'une échelle au Go ou au Hex. D'une part, cette contribution permet de définir de ces notions de manière formelle et précise dans un langage mathématique, d'autre part découpler le concept de solution de l'algorithme de recherche donne une perspective unificatrice des algorithmes de recherche dans les jeux à information parfaite. Ainsi, quatre algorithmes de model checking combinés à des spécifications MMLK appropriées suffisent à simuler 16 algorithmes traditionnels, de Lambda Search à A* en passant par Best Reply Search. Le model checking en MMLK permet aussi d'obtenir automatiquement quatre nouveaux algorithmes pour trouver, par exemple, les mats réflexes d'une position d'Échecs.

Chapitre 5 : jeux à coups simultanés

Inventée dès les années cinquante, la technique de l'élagage alpha-beta évite d'explorer certaines parties inutiles de l'espace d'états. L'algorithme traditionnel suppose que les agents effectuent leurs actions de manière séquentielle. Une telle restriction n'est cependant pas toujours possible ou souhaitable. Le Chapitre 5 s'attaque aux domaines où les agents peuvent être amenés à choisir certaines de leurs actions simultanément.

Dans un premier temps, une généralisation de l'élagage alpha-beta est proposée pour de tels domaines. Une expérimentation sur le jeu Goofspiel démontre que si

l'élagage permet d'explorer un nombre d'états significativement plus petit qu'avec le raisonnement rétrograde (*backward induction*), le temps de calcul n'est pas diminué pour autant. En effet, le calcul des bornes alpha et beta est coûteux car il utilise la programmation linéaire. Un raffinement de l'algorithme est proposé afin de trouver un équilibre entre calculer toutes les bornes et explorer des états potentiellement inutiles. Combiné à une heuristique sur l'ordre des actions jointes, ce paramétrage permet la résolution du jeu Goofspiel plus rapidement qu'avec le raisonnement rétrograde.

Les phases de combats dans les jeux de stratégie en temps réel, tels Starcraft, mettent également en jeu des actions simultanées. Jusqu'à présent, les joueurs artificiels des compétitions annuelles de Starcraft utilisaient des méthodes réactives scriptées. Les contraintes de ressources, en particulier la nécessité de choisir une action en moins de 50 ms, ne permettant pas la résolution exacte de ces phases de combat, je propose l'algorithme *Alpha-Beta Considering Durations (ABCD)* afin d'approximer la solution dans le temps imparti. Cette approximation est fondée sur les concepts mini-max et maxi-min et s'inspire des méthodes de Monte-Carlo pour calculer une fonction d'évaluation. Une expérimentation fidèle aux conditions des tournois (5 ms par décision et moins de 100 Mo de mémoire) établit l'intérêt de l'algorithme ABCD. Un agent fondé sur ABCD gagnant de 87% à 96% des combats contre un agent reproduisant les comportements scriptés observés lors des récentes compétitions. Par ailleurs, si des méthodes de modélisation de l'adversaire en amont permettent de déterminer que l'adversaire utilise un tel comportement scripté, il est possible d'obtenir de 95% à 98% de victoires dans les mêmes conditions.

Accessit au prix de thèse AFIA 2014

Apprentissage de vote de majorité pour la classification supervisée et l'adaptation de domaine : approches PAC-Bayésiennes et combinaison de similarités

Emilie Morvant (Laboratoire Hubert Curien, Saint-Étienne, France)

L'apprentissage automatique est un des champs d'étude de l'intelligence artificielle et se situe à la frontière de l'informatique, des statistiques et de l'optimisation. La notion d'apprentissage englobe toute méthode permettant de construire un modèle le plus proche possible de la réalité à partir d'observations. Son objectif est d'extraire et d'exploiter automatiquement l'information présente dans des jeux de données, en s'intéressant au développement, à l'analyse et à l'implémentation de méthodes capables de s'améliorer à l'aide des observations. De ce fait, selon les données et les objectifs, de nombreuses applications existent, notamment en multimédia, en biologie, en traitement automatique de la langue, etc.

Dans cette thèse, nous nous plaçons dans le cadre de l'apprentissage statistique introduit par Vapnik et Chervonenkis, lié à la théorie d'études statistiques sur les processus empiriques. En particulier, nous nous intéressons au paradigme de la classification supervisée qui s'oppose à celui de la classification non supervisée (*e.g.*, *clustering*) pour laquelle nous ne disposons d'aucune supervision sur la classe des observations. Prenons l'exemple d'un système de filtrage ham/spam. Ici, les observations sont des messages déjà étiquetés, prenant la forme de couples (\mathbf{x}, y) , où \mathbf{x} est la représentation d'un message et où $y \in \{-1; +1\}$ est la l'étiquette affectée au message \mathbf{x} (-1 ="spam", $+1$ ="ham"). Le but est alors de d'apprendre un modèle qui attribue automatiquement une de ces deux classes à un nouveau message, en commettant le moins d'erreurs possible. On parle alors de classification supervisée puisque l'apprentissage se base sur des observations déjà étiquetées. La classification supervisée se formalise alors comme l'apprentissage d'une fonction, souvent appelée hypothèse ou classifieur, à partir d'un échantillon de données d'apprentissage composé d'observations indépendantes et identiquement distribuées selon une distribution de probabilité fixée et inconnue. Évidemment, le modèle appris doit se montrer performant sur les observations mais aussi—et surtout—sur les nouvelles données. Autrement dit, puisque la distribution des données est supposée inconnue, la question majeure

en apprentissage statistique concerne l'apprentissage d'un classifieur se comportant le mieux possible sur l'ensemble de la distribution : il doit donc avoir de bonnes propriétés en généralisation (voir la figure 1).



Fig 1. Intuition des problématiques liées à l'apprentissage statistique : avec quelles garanties peut-on apprendre un modèle performant ?

Dans la littérature, de nombreux algorithmes ont été proposés pour résoudre ces tâches de classification. Dans le contexte de cette thèse, nous nous intéressons à la problématique de l'étiquetage automatique de vidéos, dans une situation d'indexation sémantique, dans le but de suggérer de nouvelles vidéos. De telles données multimédia sont habituellement représentées sous plusieurs formes. Basiquement, lorsque l'on parle de vidéos, nous pouvons à la fois décrire le son et l'image (et ce de différentes manières). Les données sont donc définies par un couple (\mathbf{x}, y) , où y est l'étiquette du document vidéo et où cette fois-ci $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ est décomposé en n représentations distinctes de plusieurs informations du document. Autrement dit, la tâche d'apprentissage peut être vue comme un apprentissage de n classifieurs associés aux n couples (\mathbf{x}_i, y) , que l'on va devoir combiner. De plus, la personnalisation des suggestions soulève une seconde question importante. En effet, à l'instar d'un système de filtrage de spams, des vidéos intéressantes pour un utilisateur donné ne le seront pas nécessairement pour un nouvel utilisateur. Dans le même esprit, l'expansion d'Internet induit une très grande quantité de sources de données et donc de corpora différents. Il faut donc être capable d'adapter le modèle d'un utilisateur à un autre

ou d'un corpus à un autre. Deux grandes questions nous intéressent donc :

1. La combinaison de classifieurs : Comment peut-on apprendre une combinaison de classifieurs issus de différentes connaissances *a priori* prenant la forme, par exemple, de différentes sources d'informations ou de représentations, et ce avec de bonnes garanties en généralisation ?
2. L'adaptation de domaine : Comment peut-on apprendre à partir de données sources—pour lesquelles on a des étiquettes—un classifieur performant sur des données cibles—pour lesquelles peu (ou pas) d'étiquettes sont disponibles ? D'un point de vue statistique, le cadre de l'adaptation de domaine suppose que la distribution des données sources diffère de celle des nouvelles données cibles (voir la figure 2).

Notre objectif est donc d'apprendre une combinaison performante et robuste tout en tirant bénéfice des différents classifieurs avec ou sans adaptation.

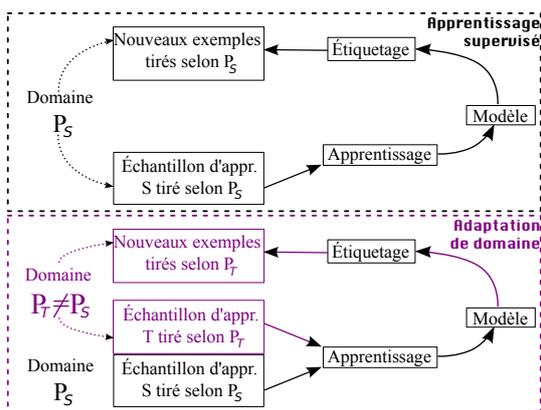


Fig 2. Distinction entre l'apprentissage supervisé et l'adaptation de domaine.

Dans un premier temps, en se plaçant dans un cadre non adaptatif, nous nous interrogeons sur l'existence d'un cadre pertinent pour combiner des classifieurs. C'est ainsi que nous nous tournons vers la théorie PAC-Bayésienne [16]. En effet, cette théorie offre un cadre naturel et élégant pour apprendre une combinaison de classifieurs prenant la forme d'un vote de majorité pondéré en considérant une connaissance *a priori*.

Tout d'abord, nous généralisons un algorithme récemment introduit, MinCq [14], qui apprend un tel vote en optimisant les poids associés aux votants. Son principe repose sur la minimisation du risque du vote de majorité (la C -borne [13]), dans le cadre de la théorie PAC-Bayésienne.

3. En apprentissage automatique, on parle aussi d'apprentissage multimodal ou multivue.

Une limite de MinCq vient du fait qu'il ne peut tirer avantage d'une connaissance *a priori* sur la performance des votants. Nous introduisons alors P-MinCq [6], une extension de MinCq, afin de considérer une contrainte *a priori* sur la distribution des poids des votants. De plus, nous généralisons les preuves de convergence aux schémas de compression permettant de considérer des votants dépendant des données d'apprentissage. Appliqué à un vote de majorité sur un ensemble de classifieurs PPV et évalué sur vingt jeux de données, nous montrons que P-MinCq est significativement plus performant qu'un PPV classique et MinCq lui-même. En outre, combiné avec un algorithme d'apprentissage de métrique, P-MinCq se montre encore meilleur. Enfin, nous spécialisons l'approche à la tâche de fusion de classifieurs³ [20] dans un contexte d'indexation sémantique de documents multimédia, lorsque les votants à combiner sont appris à partir de différentes descriptions des données.

Alors que les contributions précédentes se placent dans le cadre de la classification binaire, nous présentons ensuite deux travaux en théorie PAC-Bayésienne pour la classification multiclassée lorsque le nombre de classes est fini et supérieur à 2. Nous démontrons [19] tout d'abord une borne PAC-Bayésienne sur le risque du classifieur stochastique de Gibbs, qui correspond à la moyenne, selon une distribution, des erreurs des classifieurs issus d'un même ensemble. L'originalité et l'intérêt de ce résultat est de considérer une mesure de risque basée sur la matrice de confusion des classifieurs. En effet, en classification multiclassée, la matrice de confusion offre un outil plus approprié et plus riche que le simple taux d'erreurs, en prenant en compte les probabilités qu'une donnée de classe y soit classé en une classe différente $y' \neq y$. À notre connaissance, il s'agit de la première borne PAC-Bayésienne pour la matrice de confusion. Nous étudions ensuite les relations qui existent entre l'erreur du classifieur de Gibbs et l'erreur du vote de majorité [21]. Tout d'abord, nous montrons que la relation directe repose sur un facteur égal au nombre de classes. Puis, nous généralisons la C -borne. Nous en donnons plusieurs formulations reposant sur des notions toutes équivalentes en classification binaire.

Les contributions précédentes se positionnent dans le cadre usuel de la classification supervisée où les données d'apprentissage sont représentatives des données à traiter. Cependant, dans de nombreuses applications, cette hypothèse idéale ne peut être vérifiée. Dans de telles situations, une des stratégies vise à adapter un classifieur d'une distribution source vers une distribution cible différente. Notre première contribution [17, 18] en adaptation de domaine s'inspire des travaux fondateurs qui suggèrent de

rapprocher les deux distributions dans un nouvel espace tout en gardant de bonnes garanties sur les observations sources [15, 7]. L'idée est d'utiliser le cadre offert par les fonctions de similarité (ϵ, γ, τ) -bonnes proposé par [4]. Plus précisément, nous travaillons sur le ϕ^R -espace de projection, défini par les similarités $K(\cdot, \mathbf{x}'_j)$ à des points raisonnables d'un ensemble $R = \{\mathbf{x}'_j\}_{j=1}^r$. Cet espace est défini par :

$$\phi^R : \begin{cases} X & \rightarrow \mathbb{R}^r \\ \mathbf{x} & \mapsto (K(\mathbf{x}, \mathbf{x}'_1), \dots, K(\mathbf{x}, \mathbf{x}'_r))^\top. \end{cases}$$

En utilisant la flexibilité de ce cadre, notre méthode construit un espace où les distributions source et cible sont proches tout en gardant de bonnes garanties en généralisation sur la distribution source. Ceci est réalisé à l'aide d'une repondération des similarités, contrôlée par un terme de régularisation, pour rapprocher les exemples sources des exemples cibles. Nous proposons deux algorithmes : un algorithme spécifique à l'adaptation de domaine dite non-supervisée (*i.e.*, sans étiquette cible), que nous généralisons ensuite à l'adaptation de domaine semi-supervisée (*i.e.*, avec des étiquettes cibles). Nous réalisons, de plus, une large étude empirique de ces algorithmes. Au delà de ces aspects purement techniques, l'intérêt de cette contribution est de proposer une approche permettant de rapprocher les distributions marginales tout en optimisant l'erreur sur la distribution source ce qui, en général, se réalise en deux étapes distinctes.

Finalement, pour répondre simultanément aux deux problématiques principales de cette thèse, nous étudions l'adaptation de domaine avec un point de vue PAC-Bayésien dans le but de construire une combinaison de classifieurs capable de s'adapter à de nouvelles données. D'une part, nous savons que la théorie PAC-Bayésienne permet de proposer des analyses en moyenne du processus d'apprentissage et induit, en ce sens, des bornes en généralisation plus précises. D'autre part, elle se focalise sur l'apprentissage de votes de majorité qui est au cœur de cette thèse. Nous proposons donc la première analyse PAC-Bayésienne du problème de l'adaptation de domaine [9]. Notre contribution repose sur trois points principaux. Premièrement, nous définissons une divergence entre les distributions appropriée à l'analyse PAC-Bayésienne, bien plus simple à estimer et à minimiser que les divergences classiques, et dont le processus de minimisation empirique est consistant. À partir de cette mesure, nous démontrons une borne d'adaptation de domaine que nous analysons selon l'approche PAC-Bayésiennes. Nous en dérivons, un premier algorithme PAC-Bayésien pour l'adaptation de domaine en spécialisant ce résultat aux classifieurs

linéaires. Cet algorithme montre des comportements empiriques très prometteurs, ouvrant de nouvelles directions pour s'attaquer à la problématique de l'adaptation de domaine.

Pour conclure, nous avons étudié l'apprentissage de combinaisons de classifieurs ou de similarités pour des tâches de classification. D'une part, nous sommes restés dans le cadre usuel où les données d'apprentissage suivent la même distribution de probabilité que les données à classer. D'autre part, nous avons abordé la situation de l'adaptation de domaine pour laquelle ces deux distributions sont différentes. Ceci a donné lieu à des résultats originaux, à la fois théoriques et empiriques, et permettent de faire le lien avec des problématiques actuelles importantes en apprentissage automatique. Un premier aspect concerne la prédiction structurée qui peut être reliée au multiclasse. Il serait donc intéressant de généraliser nos résultats multiclassés à la prédiction structurée. Un second aspect est le *deep learning* qui se base sur l'apprentissage de représentation. Or, une donnée peut être décrite selon différentes descriptions dont certaines peuvent être plus pertinentes pour des tâches ou des données particulières. Il faut donc être capable de définir de bonnes représentations ou de les "apprendre" correctement. Ainsi, tirer avantage de ce cadre pour définir de nouvelles méthodes d'adaptation ou de combinaison de votants est une perspective des plus prometteuses.

Références

- [4] M. F. BALCAN, A. BLUM et N. SREBRO. "Improved Guarantees for Learning via Similarity Functions". In : *Proc. of Annual Conference on Computational Learning Theory*. 2008, p. 287–298.
- [6] Aurélien BELLET et al. "Learning A Priori Constrained Weighted Majority Votes". In : *Machine Learning Journal* 97.1-2 (2014), p. 129–154.
- [7] S. BEN-DAVID et al. "A theory of learning from different domains". In : *Machine Learning Journal* 79.1-2 (2010), p. 151–175.
- [9] P. GERMAIN et al. "PAC-Bayesian Domain Adaptation Bound with Specialization to Linear Classifiers". In : *Proc. of International Conference on Machine Learning*. 2013.
- [13] A. LACASSE et al. "PAC-Bayes Bounds for the Risk of the Majority Vote and the Variance of the Gibbs Classifier". In : *Proc. of conference on Neural Information Processing Systems*. 2007.

- [14] F. LAVIOLETTE, M. MARCHAND et J.-F. ROY. “From PAC-Bayes Bounds to Quadratic Programs for Majority Votes”. In : *Proc. of International Conference on Machine Learning*. 2011.
- [15] Y. MANSOUR, M. MOHRI et A. ROSTAMIZADEH. “Domain Adaptation : Learning Bounds and Algorithms”. In : *Proc. of Conference on Learning Theory*. 2009.
- [16] D. A. MCALLESTER. “Some PAC-Bayesian Theorems”. In : *Machine Learning Journal* 37 (3 1999), p. 355–363.
- [17] E. MORVANT, A. HABRARD et S. AYACHE. “Parsimonious Unsupervised and Semi-Supervised Domain Adaptation with Good Similarity Functions”. In : *Knowledge and Information Systems* 33.2 (2012), p. 309–349.
- [18] E. MORVANT, A. HABRARD et S. AYACHE. “Sparse Domain Adaptation in Projection Spaces based on Good Similarity Functions”. In : *Proc. of International Conference on Data Mining series*. 2011.
- [19] E. MORVANT, S. KOÇO et L. RALAIVOLA. “PAC-Bayesian Generalization Bound on Confusion Matrix for Multi-Class Classification”. In : *Proc. of International Conference on Machine Learning*. Omnipress, 2012, p. 815–822.
- [20] Emilie MORVANT, Amaury HABRARD et Stéphane AYACHE. “Majority Vote of Diverse Classifiers for Late Fusion”. In : *Joint Workshops on Statistical Techniques in Pattern Recognition and Structural and Syntactic Pattern Recognition*. 2014.
- [21] Emilie MORVANT et al. “On the Generalization of the C-Bound to Multiclass Setting”. In : *arXiv :1408.1336* (2014). Research Report.

Compte-rendu de la journée IA, IC et Santé

rédigé par Sandra Bringay, Lynda Tamine Lechani et Nathalie Souf

Synthèse générale de l'atelier

Au cours des dernières années, le nombre de données médicales n'a cessé d'augmenter suite à l'informatisation croissante des professionnels de santé. Ces données restent trop souvent sous-exploitées. L'atelier, qui s'est déroulé le 13 mai, a eu pour thème la présentation de méthodes d'ingénierie des connaissances permettant de traiter toute la richesse potentielle de ces bases. Des questions classiques d'acquisition de données, de modélisation et de mise en place de chaînes de traitements ont été évoquées. L'atelier a été également l'occasion de débattre des verrous liés à la structuration, l'extraction, la recherche et la fouille de données de santé de tous types, notamment issues des réseaux sociaux.

Nous avons reçu 10 soumissions. Les articles ont été relus par 3 relecteurs au moins. Nous avons retenu 7 articles suite aux relectures. L'atelier a rassemblé plus de 20 personnes (4 membres de l'AFIA et 2 membres de l'AIM) de disciplines variées et les discussions ont démontré une facette de l'ingénierie des connaissances appliquée à la santé, issues de l'expérience et de l'activité des professionnels de santé et qui inspire largement des approches interdisciplinaires nouvelles.

Le programme des interventions avec les articles associés est disponible sur le site web de l'atelier (<https://www.lirmm.fr/ic-sante/>).

Présentation du programme de la journée

Deux conférenciers locaux ou proches de Clermont Ferrand ont été invités en début de matinée et début d'après-midi afin de décrire des résultats concrets obtenus dans le cadre de deux projets ANR TECSAN réalisés dans le domaine de l'informatique médicale. Les autres exposés de la journée ont été organisés en deux sessions : une première session a traité des réseaux sociaux et de leur exploitation pour en extraire de nouvelles connaissances médicales et une deuxième session plus classique, a été dédiée au traitement des données médicales. Dans la suite de ce compte-rendu, chaque intervention sera accompagnée d'une annotation personnelle des organisatrices de la journée.

Exposé invité

Marinette BOUET, LIMOS, Clermont-Ferrand : *Spécification et pré-planification des soins à domicile. ANR TECSAN 2010 "PlasOsoins"*. **Notes.** Le projet PlasOsoins a comme objectif de modéliser et proposer une plateforme d'aide au suivi et à la coordination des activités de soins à domicile. L'exposé a dans un premier temps décrit le contexte de cette étude. L'objectif a ensuite été de présenter de façon plus précise les travaux menés par le LIMOS dans le cadre d'une pré-planification du plan de soins utilisée pour d'une part proposer aux infirmières coordinatrices des organisations et regroupements de soins envisageables pour construire les tournées et d'autre part pour alimenter les modules de planification fine des tournées des soignants. Cette présentation a permis de fixer le vocabulaire relatif à la mise en place des protocoles de soins et de montrer l'importance de leur modélisation et de leur planification. Une attention particulière a porté sur la modélisation de la représentation temporelle attachée aux différentes activités à effectuer pour les soins.

Session 1 : Réseau sociaux.

Thomas Opitz, Sandra Bringay, Jérôme Azé, Cyrille Joutard, Christian Lavergne and Caroline Mollevi : *Roles de patients dans les forums de santé : une perspective originale sur la qualité de la vie.* **Notes.** L'orateur présente une méthode permettant de faire émerger des sujets d'intérêt des patientes s'exprimant dans les forums traitant du cancer du sein et qui sont méconnus des oncologues. La méthode repose sur une méthode de recherche d'information permettant d'identifier des thèmes en contexte.

Amine Abdaoui, Jérôme Azé, Sandra Bringay, Pascal Poncelet and Natalia Grabar : *Analyse des messages des patients et des médecins dans les fora de santé.* **Notes.** L'orateur propose de distinguer automatiquement les patients et les professionnels de santé dans les forums de santé via une méthode basée sur de l'apprentissage. L'utilité de l'approche d'un point de vue médical est discutée.

Ghazar Chahbandarian, Mustapha Mojahid and Nathalie Bricon-Souf : *Contextual presentation of medical*

forum's discussions. **Notes.** L'orateur présente un outil de visualisation permettant de représenter des fils de discussions en prenant en considération différents niveaux de contextes médicaux. Plusieurs champs applicatifs sont évoqués.

Exposé invité.

Thierry GARAIX, LIMOS et Ecole des Mines de Saint-Etienne : *Planification des ressources humaines pour l'évitement des tensions au service d'urgences. ANR TECSAN 2011 "Host".* **Notes.** Le projet HOST Hôpital : Optimisation, Simulation et évitement des Tensions a pour objectif d'étudier et de modéliser un système d'Aide à la Décision pour anticiper la tension de l'hôpital et établir des propositions d'évitement de ces tensions. L'orateur, après avoir présenté le contexte de ce projet, a décrit les travaux menés dans son équipe de recherche pour simuler différentes situations représentant différents scénarios de flux de patients entrant aux urgences et confrontés à différentes disponibilités de soignants. Les problèmes d'optimisation de la prise en charge dans de tels cadres ont été évoqués et en particulier, des résultats portant sur l'impact d'optimisations locales sur le processus général ont été discutés.

Session 2 : Traitement des données médicales.

Laure Martin, Delphine Battistelli, Thierry Charnois, Marie-Christine Jaulent and Laure Marelle : *Mise en place d'une méthode de reconnaissance des symptômes dans le contexte des maladies rares.* **Notes.** Afin de capitaliser des connaissances à partir des documents produits pour les maladies rares, l'oratrice propose de détecter des symptômes dans les textes médicaux, en se basant sur

une approche linguistique originale pour reconnaître les entités nommées associées aux symptômes.

Frédérique Segond, Aleksandra Ponomareva, Do-moina Rabarijaona, André Bittar, Luca Dini, Ivan Ker-gourlay, Stefan Darmoni, Quentin Gicquel and Marie-Helene Metzger : *Bien représenter pour mieux raisonner : deux approches pour le dossier patient.* **Notes.** L'oratrice nous a montré dans son exposé comment lier l'analyse linguistique, la représentation des connaissances, les ontologies médicales et le raisonnement afin de fournir un système générique permettant d'extraire, de structurer et d'exploiter l'information provenant des dossiers patients informatisés.

Meriem Maaroufi, Rémy Choquet, Paul Landais and Marie-Christine Jaulent : *Formalisation des correspondances pour l'optimisation des alignements automatisés de schémas de données : Application au domaine des maladies rares.* **Notes.** L'oratrice a décrit une méthode permettant de formaliser les correspondances et ainsi optimiser l'alignement automatisé des schémas de données. Cette formalisation intègre notamment l'élément de donnée et celui de la valeur qui permet la déduction automatisée des correspondances exprimées sous forme de règles.

Yves Parès, Xavier Aimé, Jean Charlet and Marie-Christine Jaulent : *Vers une harmonisation automatique de la représentation de comptes rendus médicaux pour évaluer leurs similarities.* **Notes.** L'orateur décrit une méthode pour normaliser la représentation de comptes rendus textuels en foetopathologie de manière à constituer une base qui sera utilisée pour raisonner à partir de cas. Cette méthode se base sur la transformation des comptes rendus en arbres.

Adhésion individuelle et abonnement		<input type="checkbox"/> Demande	<input type="checkbox"/> Renouvellement
Nom :	Prénom :		
Affiliation :			
Adresse postale :			
N° de téléphone :	N° de télécopie :		
Adresse électronique :			
Activité (à titre professionnel / à titre privé (<i>raier la mention inutile</i>)) :			
Type d'adhésion			
<input type="checkbox"/> Simple :			30 €
<input type="checkbox"/> Étudiant (sur justificatif) :			15 €
<input type="checkbox"/> Soutient :			Sans objet
<input type="checkbox"/> Adhésion au collège SMA : gratuite <input type="checkbox"/> Adhésion au collège IC : gratuite <input type="checkbox"/> Adhésion au collège <i>FERA</i> (<i>Apprentissage</i>) : gratuite			

Adhésion personne morale		<input type="checkbox"/> Demande	<input type="checkbox"/> Renouvellement
Organisme :			
Adresse postale commune aux bénéficiaires couverts par cette adhésion :			
Nom et prénom du représentant :		Fonction :	
Mél :	Tél :	Fax :	
Adresse postale :			
L'adhésion morale donne droit à 5 adhésions pour les universitaires et à 15 adhésions pour les non universitaires.			
Coordonnées des bénéficiaires :			
Nom, prénom	Mél.	Tél.	Fax
		Tarif de base fixe :	Tarif par bénéficiaire :
<input type="checkbox"/> Laboratoire universitaires/PME	150 €	Gratuit pour 5 personnes (30 € par bénéficiaire supplémentaire)	
<input type="checkbox"/> Personnes morales non universitaires	450 €	Gratuit pour 15 personnes (30 € par bénéficiaire supplémentaire)	
<input type="checkbox"/> Adhésion de soutien	600 €	Sans objet	
<input type="checkbox"/> j'accepte que les renseignements ci-dessus apparaissent dans l'annuaire de l'AFIA			
<input type="checkbox"/> j'accepte que les renseignements ci-dessus soient transmis à l'ECCAI pour constituer un fichier européen			
Veillez trouver un règlement (à l'ordre de l'AFIA) deEuros			

Trésorier AFIA en charge des adhésions : Davy MONTICOLO, ENSGSI, 8 rue Bastien Lepage, 54000 Nancy.

Mode d'adhésion :
De préférence, en ligne via le site Internet de l'AFIA : <http://www.afia.asso.fr>
A défaut, cette page doit être envoyée au trésorier.

Modes de paiement :

1. par Paypal
2. par bon de commande administratif, à l'ordre de l'AFIA, envoyé au trésorier ;
3. par virement bancaire sur le compte de l'AFIA : Société Générale, 1 place du Maréchal Foch, 35000 Rennes, France. Code banque 30003, code guichet 01902, numéro de compte 00037283856 clef RIB 39.
4. par chèque, à l'ordre de l'AFIA, envoyé au trésorier ;

SOMMAIRE DU BULLETIN N° 85

Editorial	3
Dossier “Présentation d’équipes de recherche en I.A.”	4
AMA - Grenoble	4
DUKe - Nantes	6
AIDR - Paris	8
MAGMA - Grenoble	11
Mnemosyne - Bordeaux	14
Prix de thèse AFIA 2014 (1er <i>ex aequo</i>)	17
Prix de thèse AFIA 2014 (1er <i>ex aequo</i>)	20
Accessit au prix de thèse AFIA 2014	23
Compte-rendu de la journée IA, IC et Santé	27

CALENDRIER DE PARUTION DU BULLETIN DE L’AFIA

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