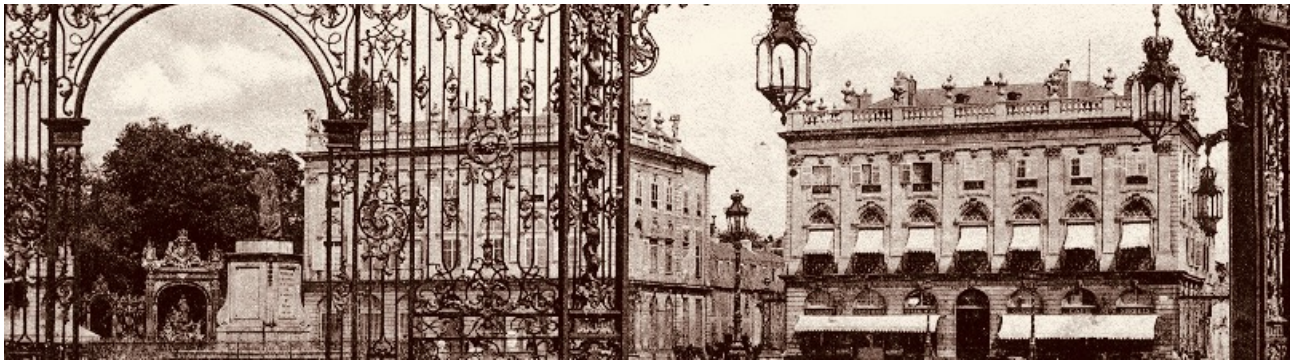


Conférence Nationale d'Intelligence Artificielle Année 2018

Sections Spéciales
Sandra Bringay
Olivier Buffet
Alain Dutech
Jérôme Euzenat
Juliette Mattioli
Sylvie Ranwez
François Schwarzentruber
Vincent Thomas

Coordination AfIA
Yves Demazeau
Dominique Longin

AfIA
Association française pour l'Intelligence Artificielle

Conférence Nationale d'Intelligence Artificielle Année 2018

Actes CNIA 2018

Conférence Nationale d'Intelligence Artificielle

Éditeurs : Président(e)s des CP de CNIA 2018

Table des matières

Mot des Présidents.
Nancy au cœur de l'IA de demain 5

Comités.
Comités 2018 de programmation, de programme, et d'organisation 7

Partenaires.
Partenaires PFIA 2018 13

I. Badache, S. Fournier, A.-G. Chifu (IC).
Prédire l'intensité de contradiction dans les commentaires : faible, forte ou très forte ? 15

G. Bono, J. Dibangoye, L. Matignon, F. Pereyron, O. Simonin (JFPDA).
Sur le gradient de la politique pour les systèmes multi-agents coopératifs 29

F. Chiaroni, M.-C. Rahal, F. Dufaux, N. Hueber (RJCIA).
Classification d'images en apprenant sur des échantillons positifs et non labélisés avec un réseau antagoniste génératif 41

C. Comparot, C. Trojahn, H. Arenas, N. Aussenac-Gilles (IC).
Ontologie pour l'intégration de données d'observation de la Terre et contextuelles basée sur les relations topologiques 49

T. Delacroix (CNIA).
Choisir un encodage CNF de contraintes de cardinalité performant pour SA 61

T. Deregnaucourt, C. Samir, A.-F. Yao (RJCIA).
Inférence bayésienne pour l'estimation de déformations larges par champs Gaussien : application au recalage d'images multi-modales 69

S. Despres (IC).
oogo : ontologie des outils utiles à la gestion d'ontologies 77

M.T. Dhouib, C. Faron-Zucker, A. Tettamanzi (APIA).
Categorization of B2B Service Offers : Lessons learnt from the Silex Use case 87

G. Gadek, A. Pauchet, S. Brunessaux, K. Khelif, B. Grilhères (APIA).
AI techniques to analyse a social network on text, user and group level : application on Galaxy2 95

A. Giboin (IC).
Étude de l'évolution du modèle de l'utilisateur des systèmes de construction collaborative d'ontologies 105

N. Gutowski, T. Amghar, O. Camp, F. Chhel (APIA).
Bandits-manchots contextuels : précision globale versus individuelle 119

G. Kassel (IC).
Une alternative à la distinction 'continuant' vs 'occurrent' 127

F. Koriche (JFPDA).
Sur la compilation de jeux de prédictions combinatoires 137

D. Lahoual, M. Fréjus (APIA).
Utilisation des systèmes d'interaction vocale dans l'habitat : quels retours d'expérience utilisateur et quels critères d'interaction à intégrer en conception ? 145

F. Maris, M. Valais, J. Vianey (JFPDA).
Codage SMT dans un espace de plans (liens causaux) pour la planification temporelle en temps continu 157

J. Mattioli, P.-O. Robic, T. Reydellet (APIA).
L'intelligence artificielle au service de la maintenance prévisionnelle 163

J.-C. Mensonides, S. Harispe, J. Montmain, V. Thireau (CNIA).
Réseaux de neurones récurrents multi-tâches pour l'analyse automatique d'arguments 167

Y. Mercadier, J. Azé, S. Bringay, V. Clavier, E. Cuenca, C. Paganelli, P. Poncelet, A. Sallaberry (IC).
#AIDS Analyse Information Dangers Sexualité : caractériser les discours à propos du VIH dans les forums de santé 175

V. Pozzoli, F. Ramond (APIA).
Optimisation en temps réel de la gestion des convergences ferroviaires 187

J. Reynaud, Y. Toussaint, A. Napoli (CNIA).
Trois approches pour classifier les données du web des données 195

E.A. Sidi Aly, M.L. Diakité, A. Giacometti, B. Markhoff, A. Soulet (CNIA).
Découverte de cardinalité maximale contextuelle dans les bases de connaissances 203

Nancy, au cœur de l'Intelligence Artificielle de demain

Armelle Brun et Davy Monticolo
Maîtres de Conférence Habilités à l'Université de Lorraine
Co-Présidents du comité d'organisation de la Plate-Forme IA 2018

Yves Demazeau
Directeur de Recherche au CNRS
Président de l'Association Française pour l'Intelligence Artificielle

L'AFIA et l'Université de Lorraine organisent à Nancy du 02 au 06 juillet 2018 la onzième Plate-Forme Intelligence Artificielle (PFIA). Les précédentes éditions se sont tenues à Palaiseau (1999), Grenoble (2001), Laval (2003), Nice (2005), Grenoble (2007), Hammamet (2009), Chambéry (2011), Lille (2013), Rennes (2015) et Caen (2017).

La plate-forme IA constitue un point de rencontre unique pour la communauté IA permettant de rapprocher les différentes disciplines qui la composent et d'établir des passerelles entre elles. A cette intention, la plate-forme IA s'adresse à l'ensemble de la communauté francophone d'Intelligence Artificielle pour aborder des problématiques communes.

Pour son édition 2018, la plate-forme IA héberge les cinq conférences suivantes, ainsi que leurs ateliers associés :

APIA	4ème Conférence Nationale sur les Applications Pratiques de l'IA
CNIA	22ème Conférence Nationale d'Intelligence Artificielle
IC	29èmes Journées Francophones d'Ingénierie des Connaissances
JFPDA	13ème Journées Francophone Planification, Décision et Apprentissage
RJCIA	16ème Rencontres des Jeunes Chercheurs en Intelligence Artificielle

Mais aussi les autres événements :

Education & IA	1ère Journée IA pour l'Education (avec MEN/DNE)
Ethique & IA	3ème Journée Ethique et IA (avec INRIA/COERLE)
France @ IJCAI	Présentation par leurs auteurs français d'articles acceptés à IJCAI 2018
Robotique & IA	3ème Journée Robotique et IA (avec CNRS/GdR Robotique)
Santé & IA	3ème Journée Santé et IA (avec AIM)
TAL & IA	4ème Journée Traitement Automatique des Langues et IA (avec ATALA)

Sept conférenciers ont également été invités pour des conférences plénières :

Nawal Benabbou	Prix de Thèse IA 2018	France
Daniele Calandriello	Prix de Thèse IA 2018	France
Aldo Gangemi	Université Paris Nord	France
Nicola Guarino	CNR Trento	Italie
Daniela Rus	MIT Boston	USA
Zhongzhi Shi	CAS Beijing	Chine
Moshe Vardi	Rice University	USA

L'ensemble des informations sur ces événements est disponible à partir du site de la Plate-Forme IA 2018 (http://pfia2018.loria.fr/) dont le contenu sera maintenu jusqu'à l'été 2022.

Nous remercions les présidents de comités de programme des conférences et journées hébergées, les membres de leurs comités de programme, les orateurs, les membres des comités de programmation et d'organisation, nos partenaires institutionnels et industriels, ainsi que tous les participants, pour leurs contributions précieuses à la réussite de cette plateforme.

©AfIA 2018

Comité de programmation

Conseil d'Administration de l'Association Française pour l'Intelligence Artificielle

Président : Yves Demazeau

- Sandra Bringay
- Yves Demazeau
- Catherine Faron-Zucker
- Dominique Longin
- Frédéric Maris
- Engelbert Nguifo Mephu

Comités de programme

4ème Conférence Applications Pratiques de l'Intelligence Artificielle (APIA)

Présidentes : Sandra Bringay LIRMM, (Université Paul-Valéry Montpellier 3)
Juliette Mattioli (Thales)

• Carole Adam,	LIG CNRS UMR 5217 - Université Grenoble-Alpes
• Florence Amardeilh,	créatrice d'entreprise en Ingénierie des Connaissances
• Ghislain Atemezing,	Mondeca
• Jérome Azé,	LIRMM-UM-CNRS
• Alain Berger,	Ardans
• Matthieu Boussard,	Craft AI
• Bertrand Braunschweig,	Inria Saclay
• Stéphan Brunessaux,	AIRBUS
• Moez Bouchouicha,	Université du Sud Toulon-Var
• Nathalie Chaignaud,	LITIS - INSA Rouen Normandie
• Caroline Chopinaud,	Craft AI
• Jan Ciger,	Reviatech SAS
• Laurent Cosserat,	Renault
• Jean-Marc David,	Renault
• Philippe David,	SNCF
• Nicolas Delaforge,	mnemotix
• Etienne de Sevin,	SANPSY - University of Bordeaux
• Yves Demazeau,	CNRS - LIG
• Christian de Sainte Marie,	IBM
• Sylvie Despres,	LIMICS - Université Paris 13
• Sébastien Destercke,	CNRS, UMR Heudiasyc
• Jean-Yves Donnart,	THALES
• Béatrice Fuchs,	LIRIS, IAE - Université Lyon 3
• Christophe Guettier,	Safran Electronique et Défense
• Céline Hudelot,	École Centrale Paris

©AfIA 2018

- Philippe Laborie, IBM
- Christophe Labreuche, Thales
- Arnaud Lallouet, Huawei
- Christine Largouët, Irisa /Agrocampus Ouest
- Vincent Lemaire, Orange
- Dominique Lenne, Heudiasyc - Université de Technologie de Compiègne
- Domitile Lourdeaux, Heudiasyc - Université de Technologie de Compiègne
- Sylvain Mahe, EDF Recherche et Développement
- Philippe Mathieu, University of Lille 1
- Nada Matta, University of Technology of Troyes
- Eunika Mercier-Laurent, CRESTIC- Université Reims Champagne Ardennes
- Christophe Menichetti, IBM
- Philippe Morignot, SAFRAN
- Selmin Nurcan, Université Paris 1 Panthéon - Sorbonne
- Jean-Marc Ogier, L3i - University of La Rochelle
- Benoit Ricard, EDF R&D
- Jean Rohmer, De Vinci Research Center (DVRC)
- Dominique Romary, ODΣFI
- Catherine Tessier, Onera
- Erwan Tranvouez, LSIS - Polytech'Marseille - Université d'Aix-Marseille
- Brigitte Trousse, Université Côte d'Azur, INRIA Sophia Antipolis - Méditerranée
- Amel Yessad, LIP 6 - Université Pierre et Marie Curie

21ème Conférence Nationale en Intelligence Artificielle
16ème Rencontre des Jeunes Chercheurs en Intelligence Artificielle (RJCIA)

Présidents CNIA : Jérôme Euzenat (INRIA & LIG, Grenoble)
RJCIA : François Schwarzentruber (ENS Rennes)

- Carole Adam LIG, CNRS – UJF
- Emmanuel Adam Univ. Lille Nord de France
- Audrey Baneyx Médialab, Sciences Po
- Nicolas Béchet IRISA
- Isabelle Bichindaritz State University of New York at Oswego
- Meghyn Bienvenu CNRS – Université de Montpellier – INRIA
- Isabelle Bloch ENST – CNRS LTCI
- Olivier Boissier Lab. Hubert Curien, Mines Saint-Etienne, Institut Henri Fayol
- Grégory Bonnet Université de Caen Normandie
- Sylvain Bouveret LIG, Grenoble INP – Université Grenoble Alpes
- Elena Cabrio I3S, Université Côte d'Azur – CNRS – INRIA
- François Charpillet LORIA, INRIA
- Mohamed Chetouani ISIR
- Laurence Cholvy ONERA-Toulouse
- Rémy Courdier LIM, Université de la Réunion
- Bruno Cremilleux Université de Caen Normandie
- Jérôme David LIG, Université Grenoble-Alpes – INRIA
- Tiago de Lima Université d'Artois – CNRS
- Cyril de Runz CRESTIC
- Nicolas Delestre LITIS, Université de Caen Normandie – INSA de Rouen
- Yves Demazeau LIG, CNRS
- Gaël Dias Université de Caen Normandie
- Catherine Faron Zucker Université Nice Sophia Antipolis
- Jérémy Fix SUPELEC
- Catherine Garbay LIG, CNRS
- Serge Garlatti ENST Bretagne
- Thomas Guyet IRISA

- Salima Hassas — Université Claude Bernard-Lyon 1
- Abir Beatrice Karami — LIRIS, CNRS
- Sébastien Konieczny — CRIL, CNRS
- Philippe Laborie — IBM
- Nicolas Lachiche — Université de Strasbourg
- Jean-Marie Lagniez — CRIL
- Arnaud Lallouet — Huawei Technologies Ltd
- Robin Lamarche-Perrin — Laboratoire d'Informatique de Paris 6
- Philippe Lamarre — LIRIS
- Jérôme Lang — LAMSADE, CNRS – Université Paris-Dauphine
- Fabien Lauer — LORIA, Université de Lorraine
- Florence Le Ber — iCube, Strasbourg
- Christophe Lecoutre — CRIL, Université d'Artois
- Marie Lefevre — LIRIS, Université Lyon 1
- Jean-Guy Mailly — LIPADE, Université Paris Descartes
- Pierre Marquis — CRIL, Université d'Artois – CNRS
- Philippe Mathieu — Université de Lille
- Nicolas Maudet — Université Pierre et Marie Curie
- Engelbert Mephu Nguifo — LIMOS, Université Clermont Auvergne
- Fabien Michel — LIRMM, Université de Montpellier
- Frédéric Migeon — IRIT, Toulouse
- Davy Monticolo — Université de Lorraine
- Maxime Morge — Computer Science Laboratory of Lille
- Abdel-Illah Mouaddib — Université de Caen Normandie
- Philippe Muller — IRIT, Toulouse University
- Amedeo Napoli — LORIA, CNRS – Inria – Université de Lorraine
- Alexandre Niveau — GREYC
- Antoine Nongaillard — Université de Lille
- Wassila Ouerdane — LGI, CentraleSupelec
- Odile Papini — LSIS, CNRS
- Alexandre Pauchet — LITIS, INSA Rouen – Normandy University
- Damien Pellier — LIG
- Frédéric Pennerath — Supelec, Metz
- Henri Prade — IRIT, CNRS
- Chedy Raïssi — INRIA
- Sylvie Ranwez — LGI2P, Ecole des mines d'Alès
- Regis Riveret — CSIRO
- Mathieu Roche — TETIS, Cirad
- Marie-Christine Rousset — Université Grenoble Alpes
- Catherine Roussey — Irstea
- Nicolas Sabouret — LIMSI, CNRS
- Thomas Schiex — INRA
- Karima Sedki — CRIL, Université d'Artois
- Christine Solnon — LIRIS, CNRS – INSA Lyon
- Isabelle Tellier — Lattice
- Alexandre Termier — Université de Rennes 1
- Konstantin Todorov — LIRMM
- Charlotte Truchet — LINA, CNRS – Université de Nantes
- Tim Van de Cruys — IRIT, CNRS
- Laurent Vercouter — LITIS, INSA de Rouen
- Srdjan Vesic — IRIT, Toulouse
- Serena Villata — I3S, Université Côte d'Azur – CNRS – Inria
- Christel Vrain — LIFO, Université d'Orléans

©AfIA 2018

- Bruno Zanuttini — GREYC, Université de Caen Normandie, CNRS, ENSICAEN
- Antoine Zimmermann — École des Mines de Saint-Étienne
- Pierre Zweigenbaum — LIMSI, CNRS – Université Paris-Saclay

29es Journées Francophones Ingénierie des Connaissances (IC)

Présidente : Sylvie Ranwez, LGI2P, IMT Mines Alès

- Marie-Hélène Abel — HEUDIASYC, Université de Technologie de Compiègne
- Xavier Aimé — LIMICS, Université Paris 13
- Yamine Ait Ameur — IRIT, INP Toulouse
- Florence Amardeilh — Créatrice d'entreprise en Ingénierie des Connaissances
- Bruno Bachimont — COSTECH, Université de Technologie de Compiègne
- Jean-Paul Barthès — HEUDIASYC, Université de Technologie de Compiègne
- Aurélien Bénel — ICD, Université de Technologie de Troyes
- Nacéra Bennacer — LRI, Centrale Supélec campus de Gif-sur-Yvette
- Pierre Bourhis — LIFL, CNRS
- Bertrand Braunschweig — Inria Saclay-Île-de-France
- Nathalie Bricon-Souf — IRIT, Université Paul Sabatier Toulouse 3
- Sandra Bringay — LIRMM, Université Paul-Valéry Montpellier 3
- Patrice Buche — IATE, INRA Montpellier
- Davide Buscald — LIPN, Université Paris 13
- Elena Cabri — I3S, INRIA Sophia Antipolis - Méditerranée
- Sylvie Calabretto — LIRIS, INSA de Lyon
- Gaoussou Camara — Université Alioune Diop de Bambey
- Pierre-Antoine Champin — LIRIS, Université Claude Bernard Lyon 1
- Jean-Pierre Chanet — TSCF, Irstea de Clermont Ferrand
- Jean Charlet — LIMICS, APHP INSERM Paris
- Olivier Corby — Université Côte d'Azur, Inria
- Amélie Cordier — Hoomano, Université de Lyon
- Mathieu D'Aquin — Insight Centre for Data Analytics, NUI Galway, Ireland
- Jérôme David — LIG, INRIA Grenoble
- Sylvie Despres — LIMICS, Université Paris 13
- Rim Djedidi, — LIMICS, Université Paris 13
- Jean-Pierre Evain — EBU Suisse
- Gilles Falquet, — CUI, Université de Genève
- Catherine Faron-Zucker — I3S, Université Nice Sophia Antipolis
- Cécile Favre — ERIC, Université Lumière Lyon 2
- Béatrice Fuchs — LIRIS, IAE - université Lyon 3
- Frédéric Fürst — MIS, Université de Picardie Jules Verne
- Jean-Gabriel Ganascia — LIP6, Université Pierre et Marie Curie
- Serge Garlatti — IMT Atlantique
- Alain Giboin — I3S, INRIA Sophia Antipolis - Méditerranée
- Nathalie Guin — LIRIS, Université Claude Bernard Lyon 1
- Ollivier Haemmerlé — IRIT, Université Toulouse le Mirail
- Sébastien Harispe — LGI2P, IMT Mines Alès
- Mounira Harzallah — LS2N, Université de Nantes
- Nathalie Hernandez — IRIT, Université Toulouse le Mirail
- Liliana Ibanescu — MIA, INRA AgroParistech Paris
- Sébastien Iksal — LIUM, IUT Le Mans
- Antoine Isaac — Europeana & VU University Amsterdam
- Clement Jonquet — LIRMM, Université de Montpellier
- Mouna Kamel — IRIT, Université de Perpignan Via Domitia
- Gilles Kassel — MIS, Université de Picardie Jules Verne
- Pascale Kuntz — LS2N, Université de Nantes
- Florence Le Ber — ICUBE, Université de Strasbourg / ENGEES

- Michel Leclère — LIRMM, Université de Montpellier
- Maxime Lefrançois — laboratoire Hubert Curien, IMT Mines Saint-Étienne
- Alain Léger — Orange Labs, France Telecom Rennes
- Dominique Lenne — HEUDIASYC, Université de Technologie de Compiègne
- Moussa Lo — Université Gaston Berger de Saint Louis
- Cédric Lopez — Emvista, Montpellier
- Nada Matta — ICD, Université de Technologie de Troyes
- Pascal Molli — LS2N, Université de Nantes
- Alexandre Monnin — Origens Medialab et ESC Clermont
- Isabelle Mougenot — UMR Espace dev, Université de Montpellier
- Fleur Mougin — ERIAS/INSERM BPH U1219, Université de Bordeaux
- Amedeo Napoli — LORIA, CNRS Nancy
- Emmanuel Nauer — LORIA, Université de Metz
- Jérôme Nobécourt — LIMICS, Université Paris 13
- Nathalie Pernelle — LRI, Université Paris Sud
- Yannick Prié — LS2N, Université de Nantes
- Cédric Pruski — Luxembourg Institute of Science and Technology
- Chantal Reynaud — LRI, Université Paris Sud
- Catherine Roussey — TSCF, Irstea
- Fatiha Saïs — LRI, Université Paris Sud
- Pascal Salembier — ICD, Université de Technologie de Troyes
- Karim Sehaba — LIRIS, Université Lumière Lyon 2
- Hassina Seridi-Bouchelaghem — LabGED, Badji Mokhtar University
- Andrea Tettamanzi — I3S, Université Nice Sophia Antipolis
- Konstantin Todorov — LIRMM, Université de Montpellier
- Francky Trichet — LS2N, Université de Nantes
- Raphaël Troncy — Data Science, EURECOM
- Serena Villata — I3S, CNRS
- Amel Yessad — LIP6, Université Pierre et Marie Curie
- Haïfa Zargayouna — LIPN, Université Paris 13
- Pierre Zweigenbaum — LIMSI, CNRS, Université Paris-Saclay, Orsay

13es Journées Planification, Décision et Apprentissage (JFPDA)

Présidents : Olivier Buffet, LORIA, INRIA
Alain Dutech, LORIA, INRIA
Vincent Thomas, LORIA, Université de Lorraine

- Aurélie Beynier — LIP6, Université Pierre et Marie Curie (Paris 6)
- Martin Cooper — IRIT, Université Paul Sabatier
- Jilles Dibangoye — CITI, INSA-Lyon, Université de Lyon, INRIA
- Humbert Fiorino — LIG, Université Grenoble Alpes
- Matthieu Geist — LIEC, Université de Lorraine
- Laurent Jeanpierre — GREYC, Université Caen-Normandie
- Emilie Kaufmann — CRIStAL, CNRS, Lille
- Frédéric Maris — IRIT, Université Paul Sabatier - Toulouse III
- Marc Metivier — LIPADE, Université Paris Descartes (Paris 5)
- Abdel-Illah Mouaddib — GREYC, Université Caen-Normandie
- Damien Pellier — LIG, Université Grenoble Alpes
- Philippe Preux — CRIStAL, INRIA, Université de Lille
- Emmanuel Rachelson — ISAE-SUPAERO, Toulouse
- Bruno Zanuttini — GREYC, Université Caen-Normandie

©AfIA 2018

Comité d'organisation

Laboratoire LORIA

Présidents : Armelle Brun et Davy Monticolo

• Anne Bonneau	CNRS
• Anne Boyer	Université de Lorraine
• Olivier Buffet	Inria
• Sylvain Castagnos	Université de Lorraine
• Vincent Chevrier	Université de Lorraine
• Alain Dutech	Inria
• Audrey Knauf	Université de Lorraine
• David Langlois	Université de Lorraine
• Slim Ouni	Université de Lorraine
• Azim Roussanaly	Université de Lorraine
• Romain Serizel	Université de Lorraine
• Vincent Thomas	Université de Lorraine
• Olivia Brenner	Université de Lorraine
• Chantal Chrétien	Université de Lorraine

Partenaires PFIA 2018

©AfIA 2018

Prédire l'intensité de contradiction dans les commentaires : faible, forte ou très forte ?

Ismail Badache Sébastien Fournier Adrian-Gabriel Chifu

Aix Marseille Univ, Université de Toulon, CNRS, LIS, Marseille, France

{ismail.badache, sebastien.fournier, adrian.chifu}@lis-lab.fr

Résumé

Les commentaires sur des ressources Web (ex. des cours, des films) deviennent de plus en plus exploitées dans des tâches d'analyse de texte (ex. détection d'opinion, détection de controverses). Cet article étudie l'intensité de contradiction dans les commentaires en exploitant différents critères tels que la variation des notations et la variation des polarités autour d'entités spécifiques (ex. aspects, sujets). Premièrement, les aspects sont identifiés en fonction des distributions des termes émotionnels à proximité des noms les plus fréquents dans la collection des commentaires. Deuxièmement, la polarité est estimée pour chaque segment de commentaire contenant un aspect. Seules les ressources ayant des commentaires contenant des aspects avec des polarités opposées sont prises en compte. Enfin, les critères sont évalués, en utilisant des algorithmes de sélection d'attributs, pour déterminer leur impact sur l'efficacité de la détection de l'intensité des contradictions. Les critères sélectionnés sont ensuite introduits dans des modèles d'apprentissage pour prédire l'intensité de contradiction. L'évaluation expérimentale est menée sur une collection contenant 2244 cours et leurs 73873 commentaires, collectés à partir de coursera.org. Les résultats montrent que la variation des notations, la variation des polarités et la quantité de commentaires sont les meilleurs prédicteurs de l'intensité de contradiction. En outre, J48 est l'approche d'apprentissage la plus efficace pour cette tâche.

Mots Clef

Analyse de sentiments, Détection d'aspects, Évaluation des critères, Intensité de contradiction.

Abstract

Reviews on web resources (e.g. courses, movies) become increasingly exploited in text analysis tasks (e.g. opinion detection, controversy detection). This paper investigates contradiction intensity in reviews exploiting different features such as variation of ratings and variation of polarities around specific entities (e.g. aspects, topics). Firstly, aspects are identified according to the distributions of the emotional terms in the vicinity of the most frequent nouns in the reviews collection. Secondly, the polarity of each review segment containing an aspect is estimated. Only resources containing these aspects with opposite polarities are considered. Finally, some features are evaluated, using feature selection algorithms, to determine their impact on the effectiveness of contradiction intensity detection. The selected features are used to learn some state-of-the-art learning approaches. The experiments are conducted on the Massive Open Online Courses data set containing 2244 courses and their 73,873 reviews, collected from coursera.org. Results showed that variation of ratings, variation of polarities, and reviews quantity are the best predictors of contradiction intensity. Also, J48 was the most effective learning approach for this type of classification.

Keywords

Sentiments analysis, Aspects detection, Criteria evaluation, Contradiction intensity.

1 Introduction

Au cours des dernières années, le web 2.0 est devenu un espace ouvert où les gens peuvent exprimer leurs opinions en laissant des traces (par exemple, un commentaire, une notation, un j'aime) sur les ressources Web. De nombreux services, tels que les blogs et les réseaux sociaux, représentent une source riche de ces données sociales, qui peuvent être analysées et exploitées dans diverses applications et contextes [2, 3, 4, 6, 17]. En particulier la détection d'opinion et l'analyse de sentiments [15], par exemple, pour connaître l'attitude d'un client vis-à-vis d'un produit ou de ses caractéristiques, ou pour révéler la réaction des gens à un événement. De tels problèmes nécessitent une analyse rigoureuse des aspects couverts par le sentiment pour produire un résultat représentatif et ciblé.

Une autre problématique concerne la diversité des opinions sur un sujet donné. Certains travaux l'abordent dans le contexte de différents domaines de recherche, avec une notion différente dans chaque cas. Par exemple, Wang et Cardie [40] visent à identifier des sentiments au niveau d'une phrase exprimée au cours d'une discussion et à les utiliser comme des caractéristiques dans un classifieur qui prédit

1. https://www.coursera.org/learn/dog-emotion-and-cognition

 ©AfIA 2018

Ressource	Commentaire (gauche)	Aspect	Commentaire (droite)	Polarité	Notation
Cours [1]	The lecturer was an annoying	speaker	and very repetitive.	-0.9	1
	Passionate	speaker	and truly amazing things to learn	+0.7	4

TABLE 1 – Exemple de deux opinions contradictoires sur le "Speaker" d'un cours coursera

la dispute dans la discussion. Qiu *et al.* [29] identifient automatiquement les débats entre des utilisateurs à partir du contenu textuel (interactions) dans les forums, en se basant sur des modèles de variables latentes. Il y a eu d'autres travaux dans l'analyse des interactions avec les utilisateurs, par exemple, l'extraction des expressions de type *agreement* et *disagreement* [24] et d'en déduire les relations de l'utilisateur en regardant leurs échanges textuels [1].
Cet article étudie les entités (par exemple, les aspects, les sujets) pour lesquelles des contradictions peuvent apparaître dans les commentaires associés à une ressource Web (par exemple des films, des cours) et comment estimer leur intensité. L'intérêt d'estimer l'intensité de la contradiction dépend du cadre d'application. Par exemple, suivre des événements ou des crises politiques controversés tels que la reconnaissance par les États-Unis de Jérusalem comme capitale d'Israël. Cela a généré des opinions (avis) contradictoires, dans les réseaux sociaux, entre différentes communautés à travers le monde. L'estimation de l'intensité de ce conflit peut être utile pour mieux analyser la tendance et les conséquences de cette décision politique. Dans le cas de la recherche d'information sociale, pour certains besoins d'information, mesurer l'intensité de la contradiction peut être utile pour identifier et classer les documents les plus controversés (par exemple les nouvelles, les événements, etc.). Dans notre cas, connaître l'intensité des opinions contradictoires sur un aspect spécifique (par exemple, *Lecturer, Speaker, Slide, Quiz*) d'un cours en ligne (en anglais) peut être utile pour savoir s'il y a certains éléments à améliorer dans ce cours. Table 1 présente une instance de commentaires contradictoires à propos de l'aspect *Speaker* (conférencier) d'un cours donné.
Afin de concevoir notre approche, des tâches fondamentales sont effectuées. Premièrement, l'extraction automatique des aspects caractérisants ces commentaires. Deuxièmement, l'identification des opinions opposées autour de chacun de ces aspects en utilisant un modèle d'analyse des sentiments. Enfin, nous allons évaluer l'impact de certains critères (par exemple, le nombre de commentaires négatifs, le nombre de commentaires positifs) sur l'estimation de l'intensité de contradiction. Plus précisément, nous tentons de sélectionner les critères les plus efficaces et de les combiner avec des approches d'apprentissage pour prédire l'intensité de contradiction. Les principales contributions abordées dans cet article sont doubles :

- **(C1).** Une contradiction dans des commentaires liés à une ressource Web donnée signifie des opinions contradictoires exprimées sur un aspect spécifique, qui est une forme de diversité de sentiments autour de l'aspect au sein de la même ressource. Mais en plus de détecter la contradiction, il est souhaitable d'estimer son intensité. Par conséquent, nous essayons de répondre aux questions de recherche suivantes :
  - ◇ **QR1.** Comment estimer/prédire l'intensité de la contradiction ?
  - ◇ **QR2.** Quel est l'impact de la prise en compte des polarités et des notations sur la prédiction de l'intensité des commentaires contradictoires ?
- **(C2).** La construction d'une collection de test issue du site Web des MOOC [2] `coursera.org`. Cette collection est utile pour l'évaluation des systèmes mesurant l'intensité de contradiction. Des études expérimentales orientées utilisateurs - *user studies* - ont été menées pour collecter les jugements de l'intensité de contradiction (*Not Contradictory, Very Low, Low, Strong* and *Very Strong*).

L'article est organisé comme suit. La section 2 présente certains travaux connexes. La section 3 détaille notre approche pour la prédiction de l'intensité des contradictions autour de certains aspects spécifiques. L'évaluation expérimentale est présentée dans la section 4. Enfin, la section 5 conclut l'article en annonçant des perspectives.

2 Vue d'ensemble : État de l'art

La détection de contradictions est un processus complexe qui nécessite souvent l'utilisation de plusieurs méthodes. Plusieurs travaux ont été proposés pour ces méthodes (détection des aspects, analyse de sentiments) mais à notre connaissance, très peu de travaux traitent de la détection et de la mesure de l'intensité de la contradiction. Dans cette section, nous allons brièvement présenter quelques approches de détection de controverses proches de nos travaux puis nous allons présenter les approches liées à la détection des aspects et l'analyse de sentiments, qui sont utiles pour introduire notre approche.

2.1 Approches de détection des contradictions et des controverses

Les études les plus liées à notre approche incluent [14, 9, 36] et [37], qui tentent de détecter une contradiction dans le texte. Il y a deux approches principales, où les contradictions sont définies comme une forme d'inférence textuelle (par exemple, *entailment identification*) et analysées en utilisant des technologies linguistiques. Harabagiu *et al.* [14] ont proposé une approche d'analyse des contradictions en exploitant des caractéristiques linguistiques et sémantiques

2. Massive Open Online Course

(ex. typologie de verbes), ainsi que des informations syntaxiques telles que la négation (ex. *I love you - I do not love you*) ou l'antonyme (des mots qui ont des significations opposées, c.-à-d. *hot-cold* ou *light-dark*). Leur travail définit les contradictions comme une implication textuelle (*textual entailment*[3]) qui est fausse, lorsque deux phrases expriment des informations mutuellement exclusives sur le même sujet. L'antonymie peut donner lieu à une contradiction lorsque les gens utilisent ces mots pour décrire un sujet. Poursuivant l'amélioration des travaux dans ce sens, De Marneffe *et al.* [9] a introduit une classification des contradictions consistant en 7 types qui se distinguent par les caractéristiques qui contribuent à une contradiction, par exemple, l'antonyme, la négation, les discordances numériques qui peuvent être causées par des données erronées : «*there are 7 wonders of the world - the number of wonders of the world are 9*». Ils ont défini les contradictions comme une situation où il est extrêmement improbable que deux phrases soient vraies lorsqu'elles sont ensemble. Tsytsarau *et al.* [36, 37] ont proposé une solution automatique et évolutive pour le problème de détection de contradictions en utilisant l'analyse des sentiments. L'intuition de leur approche est que lorsque la valeur agrégée des sentiments (sur un sujet et un intervalle de temps spécifiques) est proche de zéro, alors que la diversité des sentiments est élevée, la contradiction devrait être élevée.

Un autre thème lié à notre travail concerne la détection des controverses et des disputes. Dans la littérature, la détection des controverses a été abordée à la fois par des méthodes supervisées comme dans [27, 7, 41] ou par des méthodes non supervisées comme dans [5, 10, 11, 19]. Pour détecter les événements controversés sur Twitter (par exemple, l'accusation de viol de David Copperfield entre 2007 et 2010)[4], Popescu et Pennacchiotti [27] ont proposé un classifieur basé sur un apprentissage par arbre de décision et un ensemble de caractéristiques telles que les parties du discours, la présence de mots issus du lexique d'opinion ou de controverse, et les interactions des utilisateurs (*retweet* et *reply*). Balasubramanyan *et al.* [7] ont étendu le modèle LDA (Latent Dirichlet Allocation) supervisé pour prédire comment les membres des différentes communautés politiques réagiront émotionnellement au même sujet c-à-d. la prédiction du niveau de controverse associé à ce sujet. Des classifieurs de type machine à vecteurs de support et régression logistique ont également été proposés par Wang *et al.* [41] et par Wang et Cardie [40] pour détecter les disputes dans les discussions sur la page de Wikipedia. Par exemple dans le cas des commentaires sur les modifications des pages Wikipedia.

D'autres travaux ont également exploité Wikipédia pour détecter et identifier des sujets controversés sur le Web [10, 18, 19]. Dori-Hacohens et Allan [10] et Jang et Allan [18] ont proposé d'aligner les pages Web aux pages de Wikipedia en supposant qu'une page traite un sujet controversé si la page Wikipedia décrit un sujet lui-même controversé. La nature controversée ou non controversée d'une page Wikipedia est automatiquement détectée sur la base des métadonnées et des discussions associées à la page. Jang *et al.* [19] ont construit un modèle de langage des sujets controversés appris sur des articles de Wikipédia et utilisé ensuite pour identifier si une page Web est controversée.

La détection des controverses dans les médias sociaux a également été abordée sans supervision en se basant sur les interactions entre les différents utilisateurs [11]. Garimella *et al.* [11] ont proposé d'autres approches de mesure de contradiction basées sur la topologie du réseau, telles que la marche aléatoire (*random walk*), la centralité intermédiaire (*betweenness centrality*) et le plongement de graphe à faible dimension (*low-dimensional graph embeddings*). Les auteurs ont testé des méthodes basées sur le contenu et ont noté leur inefficacité par rapport aux méthodes basées sur un graphe utilisateur. D'autres études tentent de détecter des controverses sur des domaines spécifiques, p. ex. dans les news [35] ou dans l'analyse du débat [29].

Cependant, à notre connaissance, aucun travail antérieur n'a abordé, de manière explicite et concrète, l'intensité de la contradiction ou de la controverse. Dans cet article, contrairement aux travaux antérieurs, plutôt que d'identifier seulement la controverse autour d'un sujet choisi au préalable (par exemple, aspect lié aux nouvelles politiques), nous nous concentrons également sur l'estimation de l'intensité des opinions contradictoires autour de sujets spécifiques. Nous proposons de mesurer l'intensité de la contradiction en utilisant certaines caractéristiques (par exemple, la notation et la polarité).

2.2 Approches de détection des aspects

Les premières tentatives de détection d'aspects ont été basées sur l'approche classique d'extraction d'information (IE) en exploitant les phrase nominales fréquentes [16]. De telles approches fonctionnent bien dans la détection des aspects qui sont sous la forme d'un seul nom, mais sont moins efficaces lorsque les aspects sont de faible fréquence. Dans le contexte de la détection d'aspects, bon nombre de travaux utilisent les CRF (*Conditional Random Fields*) ou les HMM (*Hidden Markov Models*). Parmi ces travaux, nous pouvons citer [13] qui utilisent les CRF. D'autres méthodes sont non supervisées et ont prouvé leur efficacité tel que dans [34] qui construisent un modèle thématique à grains multiples (Multi-Grain Topic Model). Nous pouvons aussi citer le modèle HASM (*unsupervised Hierarchical Aspect Sentiment Model*) proposé par Kim *et al.* [20] qui permet de découvrir une structure hiérarchique du sentiment fondée sur les aspects dans les avis en ligne non labellés. Dans nos travaux, nous nous sommes inspirés de la méthode non supervisée développée par Poria *et al.* [28] basée sur l'utilisation de règles d'extraction pour les avis sur les produits. Cette méthode est en cohérence avec

3. https://en.wikipedia.org/wiki/Textual_entailment

4. http://news.bbc.co.uk/2/hi/entertainment/8456070.stm

©AfIA 2018

nos données expérimentales issues de *coursera.org*.

2.3 Approches d'analyse de sentiments

L'analyse du sentiment a fait l'objet de très nombreuses recherches antérieures. Comme dans le cas de la détection d'aspects, les approches supervisées et non supervisées ont chacune leurs solutions. Ainsi, dans les approches non supervisées, nous pouvons citer les approches basées sur les lexiques telles que l'approche développée par Turney [38] ou bien des méthodes basées sur des corpus comme les travaux de Mohammad *et al.* [23]. Au rang des approches supervisées, nous pouvons citer [25] qui comme nombre de travaux perçoivent la tâche d'analyse de sentiments comme une tâche de classification et utilisent donc des méthodes comme les SVM (*Support Vector Machines*) ou les réseaux bayésiens. D'autres travaux récents sont basés sur les RNN (*Recursive Neural Network*) tels que les travaux de Socher *et al.* [32]. Comme le propos de cet article est de mesurer l'intensité de contradiction et que l'analyse de sentiments n'est qu'une étape du processus, nous avons utilisé l'approche proposée par Radford *et al.* [31], dont son implémentation est publiquement disponible [5]. Nous décrivons cette méthode dans la section 3.1.

3 Notre approche : Prédiction de l'intensité des contradictions

Notre approche est basée à la fois sur la détection d'aspects dans les commentaires ainsi que sur l'analyse des sentiments du texte autour de ces aspects. En plus de la détection de contradiction, notre objectif est de prédire le niveau d'intensité de la contradiction en utilisant certains critères et caractéristiques. Ces caractéristiques sont liées à la notation et à la polarité des *commentaires-aspect* (texte autour d'un aspect donné).

3.1 Pré-traitement : Identification des polarités autour des aspects

Le pré-traitement est une étape clé pour l'analyse des commentaires (aspects et sentiments). Le module de pré-traitement se compose de trois étapes principales : d'une part, le marquage des termes (identification des noms, verbes, etc), par une analyse syntaxique, au sein des commentaires. Deuxièmement, les noms les plus fréquents dans l'ensemble des commentaires des différents documents sont extraits. Troisièmement, uniquement les noms entourés par des termes émotionnels sont considérés comme des aspects. Nous détaillons ces étapes dans ce qui suit.

Extraction des aspects. Dans notre étude, un aspect est une entité nominale fréquente dans les commentaires et entourée par des termes émotionnels. Afin d'extraire les aspects à partir du texte des commentaires, nous nous sommes basés sur le travail de Poria *et al.* [28]. Cette méthode correspond à nos données expérimentales (commentaires issus de *coursera*). De plus, les traitements suivants sont appliqués :

1. Calcul fréquentiel des termes constituant le corpus des commentaires,
2. Catégorisation des termes (Part-of-speech tagging) de chaque commentaire en utilisant *Stanford Parser* [6],
3. Sélection des termes ayant la catégorie nominale (NN, NNS selon le *Part of Speech Tags*) [7],
4. Sélection des noms avec des termes émotionnels dans leur voisinage de 5 mots (en utilisant *SentiWordNet*) [8]. Le choix de 5 mots a été fait après plusieurs expérimentations,
5. Extraction des termes les plus fréquents dans le corpus parmi ceux sélectionnés dans l'étape précédente. Ces termes seront considérés comme des aspects.

Exemple : Soit $C = \{c_1, c_2, c_3\}$ un ensemble de 3 commentaires associés à un document D. Nous voulons extraire les aspects à partir de chacun des commentaires en appliquant les étapes décrites ci-dessus.
Nous avons c_1 ="*The lecturer was an annoying speaker and very repetitive. I just couldn't listen to him. . . I'm sorry. There was also so much about human development etc that I started to wonder when the info about dogs would start. . . . I found the formatting so different from other courses I've taken, that it was hard to get started and figure things out. Adding to that, was the constant interruption of the "paid certificate" page. If I answer "no" once, please leave me alone ! I also think it's a bit suspect for a prof to be plugging his own book for one of these courses."*
La table 2 récapitule les 5 étapes. Premièrement, nous calculons les fréquences des termes dans l'ensemble des commentaires (à titre d'exemple, les termes "course", "material", "assignments", "content", "lecturer" apparaissent 44219, 3286, 3118, 2947, 2705, respectivement). Deuxièmement, nous étiquetons grammaticalement chaque mots (par exemple, "NN", "NNS" signifient nom en singulier et nom en pluriel, respectivement [9]). Troisièmement, seul les termes de catégorie nominale sont sélectionnés. Quatrièmement, nous gardons uniquement les noms entourés par des termes appartenant au dictionnaire *SentiWordNet* (The *lecturer* was an annoying *speaker* and very repetitive). Enfin, nous considérons comme aspects utiles uniquement les noms qui figurent parmi les noms les plus fréquents dans le corpus des commentaires (l'aspect utile dans ce commentaire est *lecturer*).
Une fois la liste des aspects qui caractérisent notre collection de données est définie, nous devons estimer la polarité

5. https://github.com/openai/generating-reviews-discovering-sentiment

6. http://nlp.stanford.edu:8080/parser/

7. https://cs.nyu.edu/grishman/jet/guide/PennPOS.html

8. http://sentiwordnet.isti.cnr.it/

9. http://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html

Étape	Description
(1)	course : 44219, material : 3286, assignments : 3118, content : 2947, lecturer : 2705, ... terme$_i$
(2)	The/DT **lecturer**/NN was/VBD an/DT annoying/VBG **speaker**/NN and/CC very/RB repetitive/JJ ./. I/PRP just/RB could/MD n't/RB listen/VB to/TO him/PRP .../ : I/PRP 'm/VBP sorry/JJ ./. There/EX was/VBD also/RB so/RB much/JJ about/IN human/JJ **development**/NN **etc**/NN that/IN I/PRP started/VBD to/TO wonder/VB when/WRB the/DT **info**/NN about/IN **dogs**/NNS would/MD start/VB .../ : ./. I/PRP found/VBD the/DT **formatting**/NN so/RB different/JJ from/IN other/JJ **courses**/NNS I/PRP 've/VBP taken/VBN ,/, that/IN it/PRP was/VBD hard/JJ to/TO get/VB started/VBN and/CC figure/VB **things**/NNS out/RP ./. Adding/VBG to/TO that/DT ,/, was/VBD the/DT constant/JJ **interruption**/NN of/IN the/DT "/" paid/VBN **certificate**/NN "/" **page**/NN ./. If/IN I/PRP answer/VBZ "/" no/UH "/" once/RB ,/, please/VB leave/VB me/PRP alone/RB !/. I/PRP also/RB think/VBP it/PRP 's/VBZ a/DT bit/RB suspect/JJ for/IN a/DT **prof**/NN to/TO be/VB plugging/VBG his/PRP$ own/JJ **book**/NN for/IN one/CD of/IN these/DT **courses**/NNS ./.
(3)	lecturer, speaker, development, dogs, formatting, courses, interruption, certificate, page, prof
(4)	lecturer, speaker
(5)	lecturer

TABLE 2 – Les différentes étapes pour extraire les aspects dans un commentaire

des sentiments autour de ces aspects. La section suivante présente notre modèle d'analyse de sentiments.

Analyse de sentiments. Le sentiment porté par un commentaire sur un aspect donné (commentaire-aspect) est estimé en utilisant la méthode appelée *SentiNeuron* [10]. *SentiNeuron* est un modèle non supervisé proposé par Radford *et al.* [31] pour détecter les signaux de sentiment dans les commentaires. Cette approche est basée sur les réseaux de neurones récurrent (*recurrent neural network*) de type mLSTM (multiplicative Long Short-Term Memory). [31] ont également trouvé qu'une unité dans le mLSTM correspond directement au sentiment de la sortie. Les auteurs ont mené une série d'expérimentations sur plusieurs collections de tests telles que les collections des commentaires issues d'Amazon [22] et d'IMDb [11]. Cette approche fournie une précision de 91.8%, et surpasse de manière significative plusieurs approches de l'état de l'art telles que celles présentées dans [21]. Nous notons que le terme polarité signifie sentiment, c'est une valeur comprise entre -1 et 1.

3.2 Collection de test : coursera.org

Données collectées. A notre connaissance, il n'existe pas à ce jour de collection de test standard, contenant des informations comme les aspects, les notations et les polarités des commentaires, pour évaluer l'efficacité des systèmes de détection de contradictions dans les commentaires. De ce fait, dans le but d'expérimenter l'efficacité de notre approche, nous avons collecté 2244 ressources en anglais extraites du site "coursera.org" via son API [12], entre le 10 et le 14 octobre 2016. Chaque ressource décrit un cours et est représentée par un ensemble de métadonnées. Pour chaque cours, nous avons collecté également ses commentaires et ses notations via le *parsing* des pages web des cours (voir les statistiques sur la table 3).

Champ	Nombre
Cours	2244
Cours notés	1115
Commentaires	73873
notations	298326
Commentaires avec notation ★☆☆☆☆	1705
Commentaires avec notation ★★☆☆☆	1443
Commentaires avec notation ★★★☆☆	3302
Commentaires avec notation ★★★★☆	12202
Commentaires avec notation ★★★★★	55221

TABLE 3 – Les chiffres des données coursera.org

Nous avons pu capturer automatiquement 22 aspects utiles à partir de l'ensemble des commentaires (voir table 4). La table 4 présente les statistiques sur les 22 aspects détectés, par exemple, pour l'aspect *Slide* nous avons enregistré : 56 notations d'une étoile, 64 notations de deux étoiles, 81 notations de trois étoiles, 121 notations de quatre étoiles, 115 notations de cinq étoiles, 131 commentaires avec une polarité négative, 102 commentaires avec une polarité positive ainsi que 192 commentaires et 41 cours concernant cet aspect.

Jugements par les utilisateurs (contradictions et sentiments). Afin d'obtenir des jugements de contradictions et de sentiments pour un aspect donné : 1) nous avons demandé à trois utilisateurs d'évaluer la classe de sentiment pour chaque commentaires-aspect de 1100 cours ; 2) trois autres utilisateurs ont évalué le degré de contradiction entre les commentaires-aspect. En moyenne 60 commentaires-aspect par cours sont jugés manuellement pour chaque aspect (totalement : 66104 commentaires-aspect de 1100 cours, c'est-à-dire 50 cours pour chaque aspect). Nous notons que chaque aspect a été jugé par 3 utilisateurs.
Pour évaluer les sentiments et les contradictions dans les

10. https://github.com/openai/generating-reviews-discovering-sentiment
11. https://www.cs.cornell.edu/personnes/pabo/film-review-data/
12. https://building.coursera.org/app-platform/catalog

©AfIA 2018

Aspects	Not1	Not2	Not3	Not4	Not5	Négatif	Positif	Comment	Cours
Assignment	204	208	333	840	1726	1057	1763	2384	186
Content	176	179	341	676	1641	505	1496	1883	207
Exercise	29	46	94	290	693	195	531	673	58
Information	100	123	238	523	1389	299	1165	1359	143
Instructor	129	106	122	302	1514	295	1107	1322	140
Knowledge	74	72	121	400	1604	905	791	1243	178
Lecture	185	206	290	613	1762	763	1508	1988	208
Lecturer	32	41	48	85	461	55	193	236	39
Lesson	40	59	75	224	712	187	420	554	84
Material	191	203	328	722	2234	784	1693	2254	237
Method	19	23	40	125	404	53	187	224	31
Presentation	46	50	75	142	413	93	196	274	54
Professor	76	74	129	452	3001	331	2234	2369	151
Quality	55	53	51	110	372	113	170	262	54
Question	94	98	172	284	356	311	289	502	104
Quiz	151	155	221	401	581	481	475	824	128
Slide	56	64	81	121	115	131	102	192	47
Speaker	17	15	34	70	170	34	72	103	24
Student	140	105	171	383	1035	519	709	1066	172
Teacher	62	46	82	293	2180	248	1481	1642	119
Topic	67	89	176	437	1154	236	951	1066	130
Video	228	238	356	707	1614	941	1421	2058	245
Nombre total : 22 aspects détectés									

TABLE 4 – Statistiques sur les aspects issus des commentaires de Coursera.org

FIGURE 1 – Interface du système d'évaluation

commentaires-aspect de chaque cours, nous utilisons une échelle de notation de 3 points pour les sentiments : (*Negative, Neutral, Positive*) ; et une échelle de 5 points pour les contradictions : *Not Contradictory, Very Low, Low, Strong* et *Very Strong* (voir la figure 1).

Nous avons analysé le degré d'accord entre les évaluateurs des contradictions pour chaque aspect avec la mesure Kappa Cohen k [8]. Cet indicateur prend en compte la proportion d'accord entre les évaluateurs et la proportion de l'accord attendu entre les évaluateurs par hasard. La mesure de Kappa est égale à 1 si les évaluateurs sont complètement d'accord, 0 s'ils ne sont d'accord que par hasard. k est négatif si l'accord entre évaluateurs est pire que l'aléatoire. Comme nous avons trois évaluateurs par aspect, la valeur Kappa a été calculée pour chaque paire d'évaluateurs, puis leur moyenne a été calculée.

La figure 2 montre la distribution de la mesure kappa pour chaque aspect. Nous constatons que la mesure de l'accord varie de 0.60 à 0.91. La mesure moyenne d'accord entre les évaluateurs est de 80%, ce qui correspond à un accord fort. Concernant l'analyse du degré d'accord entre les évaluateurs des sentiments, nous avons trouvé un accord de Kappa $k = 0.78$, qui correspond aussi à un accord fort.

3.3 Identification des critères les plus efficaces

Dans cette étude, nous nous sommes appuyés sur des algorithmes de sélection d'attributs pour déterminer les critères

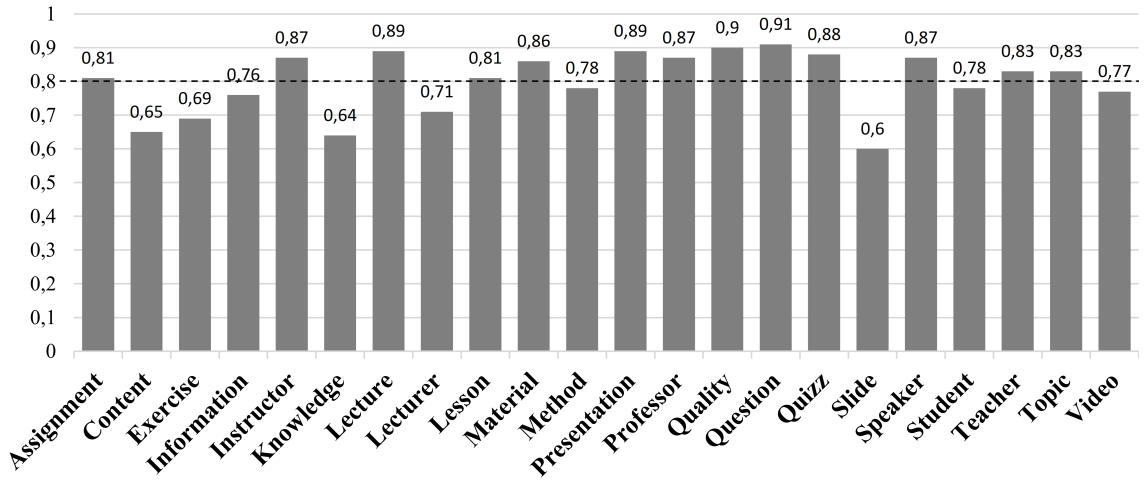


FIGURE 2 – Répartition de la mesure Kappa k par aspect. < 0 désaccord, $0.0 - 0.2$ accord très faible, $0.21 - 0.40$ accord faible, $0,41 - 0,6$ accord modéré, $0,61 - 0,80$ accord fort, $0,81 - 1$ accord parfait.

c_i	Critère	Description
c_1	NegCom	Nombre de commentaires négatifs sur le document
c_2	PosCom	Nombre de commentaires positifs sur le document
c_3	TotalCom	Nombre total des commentaires sur le document
c_4	Not1	Nombre de commentaires avec notation ★☆☆☆☆
c_5	Not2	Nombre de commentaires avec notation ★★☆☆☆
c_6	Not3	Nombre de commentaires avec notation ★★★☆☆
c_7	Not4	Nombre de commentaires avec notation ★★★★☆
c_8	Not5	Nombre de commentaires avec notation ★★★★★
c_9	VarNot	Variation des notations (écart type selon [26])
c_{10}	VarPol	Variation des polarités (écart type selon [26])

TABLE 5 – Liste des critères exploités

les plus importants pour la tâche de prédiction d'intensité de contradiction. Les algorithmes de sélection d'attributs visent à identifier et supprimer le maximum d'information inutile, redondante et non pertinente en amont d'un processus à base d'apprentissage [12]. Ils permettent également de sélectionner de manière automatique les sous ensembles de critères permettant d'avoir les meilleurs résultats. Nous avons utilisé Weka [13] (dernière version stable 2018 : 3.8.2), un outil open-source écrit entièrement en Java et qui rassemble un bon ensemble de techniques d'apprentissage et des techniques de sélection d'attributs.

La table 5 présente les 10 critères que nous avons considérés pour prédire l'intensité de contradiction dans les commentaires. La nature des critères c_1 jusqu'au critère c_8 est un simple comptage, par exemple les critères c_1 et c_2 liés à la polarité représentent le nombre de commentaires négatifs et positifs sur le document, respectivement. Les critères c_4, c_5, c_6, c_7 et c_8 sont liés à la notation. La notation est une note sur une échelle de 1 à une valeur max de 5, où 3 signifie "moyen" et 5 signifie "excellent". Concernant les deux derniers critères c_9 et c_{10}, ils représentent la variation des notations et des polarités des commentaires pour un aspect donné associés à un document (un cours dans notre cas). Ces deux critères sont calculé en se basant sur la formule de l'écart type proposée par Pearson et Stephens [26] :

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x - \bar{x})^2}{n}} \quad (1)$$

Où x est la valeur du critère (notation, polarité), $\bar{x}$ est la moyenne de l'échantillon du critère concerné, et n est la taille de l'échantillon.

Dans cette étude, nous avons procédé ainsi : 50 cours avec leurs commentaires pour chaque aspect (22 aspects) de la collection *coursera* ont été extraits aléatoirement. Ensuite, nous avons considéré l'échelle des 4 points comme des classes reflétant l'intensité des contradictions autour d'un aspect spécifique : *Very Low, Low, Strong* et *Very Strong*, selon les jugements des évaluateurs. L'ensemble résultant contient 1100 cours (instances) répartis selon leur classe d'intensité de contradiction comme suit :

— 230 Very Low
— 264 Low
— 330 Strong

13. http://www.cs.waikato.ac.nz/ml

©AfIA 2018

— 276 Very Strong

Les classes de ces ensembles sont déséquilibrées, or lorsque le nombre d'éléments d'une classe dans une collection d'apprentissage dépasse considérablement les autres échantillons des autres classes, un classifieur tend à prédire les échantillons de la classe majoritaire et peut ignorer complètement les classes minoritaires [42]. Pour cette raison, nous avons appliqué une approche de sous-échantillonnage (en réduisant le nombre d'échantillons qui ont la classe majoritaire) pour générer des collections équilibrées composées de :

— 230 Very Low

— 230 Low

— 230 Strong

— 230 Very Strong

Les classes *Low, Strong* et *Very Strong* ont été sélectionnées aléatoirement. Enfin, nous avons appliqué les algorithmes de sélection d'attributs sur les quatre ensembles obtenus, pour 5 itérations de validation croisée (5-folds cross-validation).

Dans notre cas, les algorithmes de sélection d'attributs consistent à attribuer un score à chaque critère en fonction de sa signification vis-à-vis la classe d'intensité de contradiction (*Very Low, Low, Strong* et *Very Strong*). Ces algorithmes fonctionnent différemment : certains retournent un classement d'importance des critères (par exemple, *FilteredAttributeEval*), tandis que d'autres retournent le nombre de fois qu'un critère donné a été sélectionné par un algorithme dans une validation croisée (par exemple, *FilteredSubsetEval*). Nous notons que nous avons utilisé pour chaque algorithme le paramétrage par défaut fourni par Weka.

Nous avons appliqué une validation croisée à 5 itérations pour 10 critères, c'est-à-dire $n = 10$. La table 6 présente les critères sélectionnés par les algorithmes de sélection d'attributs. Nous avons utilisé deux types de ces algorithmes : a) ceux qui utilisent des méthodes de classement pour ordonner les critères sélectionnés (la métrique dans la table est [Rank]) ; et b) ceux qui utilisent des méthodes de recherche qui indiquent combien de fois le critère a été sélectionné pendant la tâche de la validation croisée (la métrique dans la table est [Folds]). Un critère fortement préféré (choisi) par l'algorithme de sélection est un critère bien classé, c'est-à-dire *Rank* $= 1$ et fortement sélectionné, c'est-à-dire *Folds* $= 5$.

La table 6 montre que les critères c_{10} : *VarPol*, c_9 : *VarNot*, c_1 : *NegCom* et c_2 : *PosCom* sont les plus sélectionnés et les mieux classés par rapport aux autres critères. Les critères c_3 : *TotalCom*, c_4 : *Not1* et c_8 : *Not5* sont modérément favorisés par les algorithmes de sélection d'attributs, à l'exception de l'algorithme *CfsSubsetEval* qui n'a pas sélectionné c_4 et c_8. Les critères c_5, c_6 et c_7 ne sont pas sélectionnés à la fois par les algorithmes *CfsSubsetEval* et *WrapperSubsetEval*. Enfin, les critères les plus faibles et les plus désavantagés sont c_5 : *Not2* et c_6 : *Not3*, ils sont ordonnés au rang 9 et 10, respectivement.

3.4 Apprentissage des critères pour prédire l'intensité de contradiction

D'autres expérimentations ont été menées en exploitant ces critères dans des approches supervisées basées sur des modèles d'apprentissage. Nous avons utilisé les instances (les cours) des 22 aspects de la collection *coursera.org* comme ensembles d'apprentissage. Nous avons ensuite utilisé trois algorithmes d'apprentissage. Ce choix s'explique par le fait qu'ils ont souvent montré leur efficacité dans les tâches d'analyse de texte : SVM [39], J48 (implémentation C4.5) [30] et Naive Bayes [43]. L'entrée de chaque algorithme est un vecteur de critères (voir table 5), soit tous les critères ou seulement les critères sélectionnés par un algorithme de sélection précis. Les algorithmes d'apprentissage prédisent la classe d'intensité de contradiction pour les cours (*Very Low, Low, Strong* et *Very Strong*). Enfin, nous avons appliqué une validation croisée pour 5 itérations (5-folds cross-validation). La figure 3 illustre le processus d'apprentissage que nous avons mis en place pour l'évaluation des critères. Nous rappelons que la phase des algorithmes de sélection d'attributs a fait ressortir les ensembles de critères suivants (voir la table 7).

La question à ce stade est liée à la spécification du vecteur de critères d'entrée pour les algorithmes d'apprentissage, soit on prend tous les critères, soit on garde uniquement ceux sélectionnés par les techniques de sélection d'attributs. Dans ce cas, avec quels algorithmes d'apprentissage ces derniers seront combinés.

Afin de prendre en compte les critères choisis par les algorithmes de sélection dans des modèles d'apprentissage, nous nous sommes basés sur les travaux de Hall et Holmes [12]. Ils ont étudié l'efficacité de certaines techniques de sélection d'attributs en les confrontant avec les techniques d'apprentissage. Étant donné que la performance des critères diffère d'une technique d'apprentissage à une autre, ils ont identifié les meilleures techniques de sélection d'attributs permettant de retrouver les critères les plus performants en fonction des techniques d'apprentissage à utiliser. En se basant sur leur étude, nous avons utilisé les mêmes couples des techniques d'apprentissage et des techniques de sélection d'attributs :

— L'ensemble des critères sélectionnés par *CfsSubsetEval* (CFS) et *WrapperSubsetEval* (WRP) sont appris par le modèle Naïve Bayes.

— L'ensemble des critères sélectionnés par *ReliefFAttributeEval* (RLF) sont appris par le modèle J48 (C4.5 implementation).

— L'ensemble des critères sélectionnés par *SVMAttributeEval* (SVM) sont appris par le modèle SVM à multi-classes (appelé *SMO function* sur Weka).

Afin de vérifier la significativité des résultats par rapport aux résultats de Naïve Bayes (considérés comme références - résultats de base), nous avons effectué le test de

Algorithm	Metric	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}
CfsSubsetEval	[Folds]	5	5	2	0	0	0	0	0	5	5
WrapperSubsetEval	[Folds]	4	4	4	2	0	0	0	2	5	5
ConsistencySubsetEval	[Folds]	5	5	4	2	1	1	2	2	5	5
FilteredSubsetEval	[Folds]	5	5	4	3	2	2	3	3	5	5
	Moyenne	4.75	4.75	3.5	1.75	0.75	0.75	1.25	1.75	5	5
ChiSquaredAttributeEval	[Rank]	3	4	5	7	9	10	8	6	2	1
FilteredAttributeEval	[Rank]	4	3	5	7	9	10	8	6	2	1
GainRatioAttributeEval	[Rank]	3	4	5	7	9	10	8	6	2	1
InfoGainAttributeEval	[Rank]	3	4	5	7	9	10	8	6	1	2
OneRAttributeEval	[Rank]	4	3	5	7	9	10	8	6	2	1
ReliefFAttributeEval	[Rank]	4	3	6	8	9	10	7	5	1	2
SVMAttributeEval	[Rank]	4	3	5	7	9	10	8	6	2	1
SymetricalUncertEval	[Rank]	3	4	5	7	9	10	8	6	2	1
	Moyenne	3.5	3.5	5.12	7.12	9	10	7.87	5.87	1.75	1.25

TABLE 6 – Les critères sélectionnés par les algorithmes de sélection d'attributs

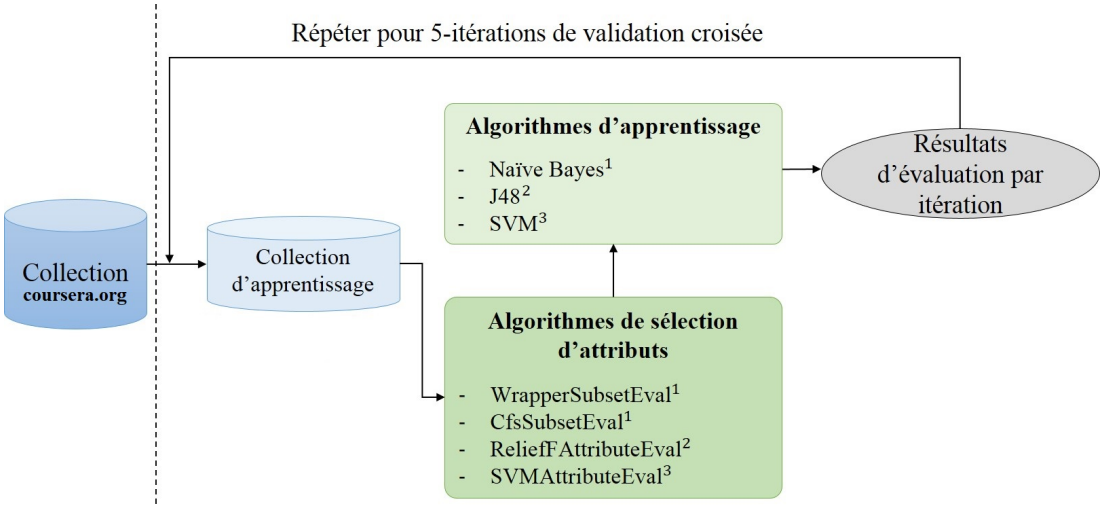


FIGURE 3 – Processus d'apprentissage en utilisant les algorithmes de sélection

Algorithmes de sélection	Critères
CfsSubsetEval	c_1, c_2, c_3, c_9, c_{10}
WrapperSubsetEval	c_1, c_2, c_3, c_4, c_8, c_9, c_{10}
Other algorithms	c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, c_9, c_{10}

TABLE 7 – Ensembles des critères sélectionnés

Student [33]. Nous avons attaché * (forte signification) et ** (très forte signification) aux résultats de la table 8 quand *p-value <0.05* et *p-value <0.01*, respectivement.

La table 8 présente les résultats des trois algorithmes d'apprentissage des critères ressortis de l'étude utilisant les techniques de sélection d'attributs. Les résultats sont discutés ci-dessous pour chaque algorithme d'apprentissage.

Résultats obtenus par Naïve Bayes (Baseline). Les résultats en termes de précision obtenus en utilisant des algorithmes de sélection CFS et WRP avec NaiveBayes, sont de 0.68 et 0.72, respectivement. Ces résultats dépassent ceux obtenus en utilisant tous les critères (précision : 0.60). En effet, nous avons enregistré des taux d'amélioration moyens de 14% et 20% pour Naïve Bayes en utilisant seulement les critères sélectionnés par CFS (0.68) et WRP (0.72), respectivement, par apport au résultat obtenu en utilisant tous les critères (0.60). Par conséquent, les approches d'apprentissage automatique peuvent donner une meilleure efficacité (précision) quand ils sont combinés avec des approches de sélection d'attributs. Les meilleures précisions

©AfIA 2018

Classifieurs	Classes (Niveaux d'intensité)	Techniques de sélection	Tous les critères
Naïve Bayes (Baseline)	Very Low	0.81 (CFS)	0.71
	Low	0.38 (CFS)	0.34
	Strong	0.75 (CFS)	0.66
	Very Strong	0.78 (CFS)	0.69
	Moyenne	0.68 (CFS)	0.60
	Very Low	0.86 (WRP)	0.72
	Low	0.46 (WRP)	0.38
	Strong	0.76 (WRP)	0.63
	Very Strong	0.80 (WRP)	0.67
	Moyenne	0.72 (WRP)	0.60
SVM	Very Low	0.88^{*} (SVM)	0.88^{*}
	Low	0.72^{**} (SVM)	0.72^{**}
	Strong	0.78^{*} (SVM)	0.78^{*}
	Very Strong	0.90^{**} (SVM)	0.90^{**}
	Moyenne	0.82^{**} (SVM)	0.82^{**}
J48	Very Low	0.97^{**} (RLF)	0.97^{**}
	Low	0.92^{**} (RLF)	0.92^{**}
	Strong	0.97^{**} (RLF)	0.97^{**}
	Very Strong	0.98^{**} (RLF)	0.98^{**}
	Moyenne	0.96^{**} (RLF)	0.96^{**}

TABLE 8 – Les résultats de précision Weka pour les techniques d'apprentissage automatique

sont obtenues pour les classes *Very Strong, Strong* et *Very Low*. Il semble que la classe *Low* est difficile à prédire avec Naïve Bayes, tout en utilisant à la fois les deux algorithmes de sélection CFS (0.38) et WRP (0.46).

Résultats obtenus par SVM. Les résultats obtenus par SVM en utilisant l'algorithme de sélection *SVMAttributeEval*, où tous les critères ont été sélectionnés, sont meilleurs par rapport à ceux obtenus par Naïve Bayes. Nous avons enregistré des taux d'amélioration moyens de 21% et 14% pour SVM par rapport à Naïve Bayes en utilisant CFS et WRP, respectivement. Nous avons également remarqué que SVM était capable de prédire la classe *Low* avec une meilleure précision que celle fournie par Naïve Bayes. Même si l'algorithme SVM est un peu coûteux en termes de temps d'exécution par rapport à Naïve Bayes, il reste favoriser pour obtenir des résultats significatifs en termes de précision.

Résultats obtenus par J48. Les résultats confirment que l'arbre de décision J48 est le modèle le plus approprié, il prend en compte tous les critères de manière plus efficace que les autres configurations. Les taux d'amélioration moyens par rapport à Naïve Bayes (en utilisant CFS et WRP) et SVM sont 41%, 33% et 17%, respectivement. En outre, les améliorations sont également fortement significatives pour chaque classe par rapport à SVM et Naïve Bayes. La classe *Low*, difficile à prédire avec les configurations précédentes, a été prédite avec une très forte précision de 92%. Comparées à Naïve Bayes (en utilisant CFS et WRP) et SVM, les améliorations enregistrées concernant la classe *Low* sont de 142%, 100% et 28%, respectivement.

Enfin, tous ces résultats expérimentaux montrent clairement que l'approche proposée permet de détecter de manière significative l'intensité de la contradiction dans les commentaires. Nous avons constaté que les résultats obtenus, par les deux algorithmes CFS et WRP, confirment l'hypothèse lancé par Hall et Holmes. C'est en effet les deux seuls cas pour lequel les résultats de précision obtenus avec la sélection d'attributs, soient 0.68 (CFS) et 0.72 (WRP), dépassent l'utilisation de tous les critères, 0.60 en termes de précision. Ces améliorations montrent l'intérêt de combiner les algorithmes de sélection d'attributs avec les modèles d'apprentissage. En outre, le modèle J48 a

donné les meilleures améliorations par rapport à toutes les autres configurations. Nous concluons que les ressources (cours) ayant des opinions plus diversifiées (commentaires positifs et négatifs), sont susceptibles d'avoir des contradictions avec différents niveaux d'intensité.

4 Conclusion

Cet article propose une approche supervisée exploitant un ensemble de critères permettant de prédire l'intensité de la contradiction, en attirant l'attention sur les aspects dans lesquels les utilisateurs ont des opinions contradictoires. L'intuition derrière l'approche proposée est que les notations et les sentiments associés aux commentaires sur un aspect spécifique peuvent être considérés comme des critères (ex. diversité des sentiments et des notations en fonction de l'écart-type) pour mesurer l'intensité de contradiction. L'évaluation expérimentale menée sur la collection issue de *coursera.org* montre que les critères *NegCom, PosCom, VarNot* et *VarPol* sont les plus fructueux pour prédire l'intensité de la contradiction. De plus, les algorithmes d'apprentissage basés sur les critères les plus pertinents selon les algorithmes de sélection d'attributs sont généralement mieux comparés à ceux obtenus lorsque les algorithmes de sélection d'attributs sont ignorés. L'algorithme J48 apporte les meilleurs résultats par rapport à Naïve Bayes et SVM. Enfin, nous notons que nous sommes conscients que l'évaluation de notre approche est encore limitée. La principale faiblesse de notre approche est sa dépendance à la qualité des modèles de sentiments et d'extraction d'aspect. D'autres expérimentations à plus grande échelle sur d'autres types de collections sont également envisagées. Ceci étant même avec ces éléments simples, les premiers résultats obtenus nous encouragent à investir davantage cette piste.

Références

[1] Ahmed Hassan Awadallah, Amjad Abu-Jbara, and Dragomir R. Radev. Detecting subgroups in online discussions by modeling positive and negative relations among participants. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, EMNLP-CoNLL 2012, July 12-14, 2012, Jeju Island, Korea*, pages 59–70, 2012.

[2] Ismail Badache and Mohand Boughanem. Harnessing social signals to enhance a search. In *Proceedings of the 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT) - Volume 01*, WI-IAT '14, pages 303–309, Washington, DC, USA, 2014.

[3] Ismail Badache and Mohand Boughanem. Emotional social signals for search ranking. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '17, pages 1053–1056, New York, NY, USA, 2017. ACM.

[4] Ismail Badache and Mohand Boughanem. Fresh and diverse social signals : Any impacts on search ? In *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval*, CHIIR '17, pages 155–164, New York, NY, USA, 2017. ACM.

[5] Ismail Badache, Sébastien Fournier, and Adrian-Gabriel Chifu. Finding and quantifying temporal-aware contradiction in reviews. In *Information Retrieval Technology - 13th Asia Information Retrieval Societies Conference, AIRS 2017, Jeju Island, South Korea, November 22-24, 2017, Proceedings*, pages 167–180, 2017.

[6] Ismail Badache, Sébastien Fournier, and Adrian-Gabriel Chifu. Harnessing ratings and aspect-sentiment to estimate contradiction intensity in temporal-related reviews. In *Knowledge-Based and Intelligent Information & Engineering Systems : Proceedings of the 21st International Conference KES-2017, Marseille, France, 6-8 September 2017.*, pages 1711–1720, 2017.

[7] Ramnath Balasubramanyan, William W. Cohen, Douglas Pierce, and David P. Redlawsk. Modeling polarizing topics : When do different political communities respond differently to the same news ? In *Proceedings of the Sixth International Conference on Weblogs and Social Media, Dublin, Ireland, June 4-7, 2012*, 2012.

[8] Jacob Cohen. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1) :37–46, 1960.

[9] Marie-Catherine de Marneffe, Anna N. Rafferty, and Christopher D. Manning. Finding contradictions in text. In *ACL 2008, Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics, June 15-20, 2008, Columbus, Ohio, USA*, pages 1039–1047, 2008.

[10] Shiri Dori-Hacohen and James Allan. Automated controversy detection on the web. In *Advances in Information Retrieval - 37th European Conference on IR Research, ECIR 2015, Vienna, Austria, March 29 - April 2, 2015. Proceedings*, pages 423–434, 2015.

[11] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. Quantifying controversy in social media. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining, San Francisco, CA, USA, February 22-25, 2016*, pages 33–42, 2016.

[12] Mark A. Hall and Geoffrey Holmes. Benchmarking attribute selection techniques for discrete class data mining. *IEEE Trans. Knowl. Data Eng.*, 15(6) :1437–1447, 2003.

[13] Hussam Hamdan, Patrice Bellot, and Frédéric Béchet. Lsislif : CRF and logistic regression for opinion target extraction and sentiment polarity analysis.

©AfIA 2018

In *Proceedings of the 9th International Workshop on Semantic Evaluation, SemEval@NAACL-HLT, Colorado, USA, June 4-5, 2015*, pages 753–758, 2015.

[14] Sanda M. Harabagiu, Andrew Hickl, and V. Finley Lacatusu. Negation, contrast and contradiction in text processing. In *Proceedings, The Twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference, July 16-20, 2006, Boston, Massachusetts, USA*, pages 755–762, 2006.

[15] Amal Htait, Sébastien Fournier, and Patrice Bellot. LSIS at semeval-2016 task 7 : Using web search engines for english and arabic unsupervised sentiment intensity prediction. In *Proceedings of the 10th International Workshop on Semantic Evaluation, SemEval@NAACL-HLT 2016, San Diego, CA, USA, June 16-17, 2016*, pages 469–473, 2016.

[16] Minqing Hu and Bing Liu. Mining and summarizing customer reviews. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, Washington, USA, August 22-25, 2004*, pages 168–177, 2004.

[17] Melanie Imhof, Ismail Badache, and Mohand Boughanem. Multimodal social book search. In *Working Notes of CLEF 2015 - Conference and Labs of the Evaluation forum, Toulouse, France, September 8-11, 2015.*, 2015.

[18] Myungha Jang and James Allan. Improving automated controversy detection on the web. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval, SIGIR 2016, Pisa, Italy, July 17-21, 2016*, pages 865–868, 2016.

[19] Myungha Jang, John Foley, Shiri Dori-Hacohen, and James Allan. Probabilistic approaches to controversy detection. In *Proceedings of the 25th ACM International Conference on Information and Knowledge Management, CIKM 2016, Indianapolis, IN, USA, October 24-28, 2016*, pages 2069–2072, 2016.

[20] Suin Kim, Jianwen Zhang, Zheng Chen, Alice H. Oh, and Shixia Liu. A hierarchical aspect-sentiment model for online reviews. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence, July 14-18, 2013, Bellevue, Washington, USA.*, 2013.

[21] Moshe Looks, Marcello Herreshoff, DeLesley Hutchins, and Peter Norvig. Deep learning with dynamic computation graphs. *CoRR*, abs/1702.02181, 2017.

[22] Julian J. McAuley, Rahul Pandey, and Jure Leskovec. Inferring networks of substitutable and complementary products. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia, August 10-13, 2015*, pages 785–794, 2015.

[23] Saif Mohammad, Svetlana Kiritchenko, and Xiaodan Zhu. Nrc-canada : Building the state-of-the-art in sentiment analysis of tweets. In *Proceedings of the 7th International Workshop on Semantic Evaluation, SemEval@NAACL-HLT 2013, Atlanta, Georgia, USA, June 14-15, 2013*, pages 321–327, 2013.

[24] Arjun Mukherjee and Bing Liu. Mining contentions from discussions and debates. In *The 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '12, Beijing, China, August 12-16, 2012*, pages 841–849, 2012.

[25] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs up ? sentiment classification using machine learning techniques. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing, EMNLP 2002, Philadelphia, PA, USA, July 6-7, 2002*, 2002.

[26] ES Pearson and MA Stephens. The ratio of range to standard deviation in the same normal sample. *Biometrika*, 51(3/4) :484–487, 1964.

[27] Ana-Maria Popescu and Marco Pennacchiotti. Detecting controversial events from twitter. In *Proceedings of the 19th ACM Conference on Information and Knowledge Management, CIKM 2010, Toronto, Ontario, Canada, October 26-30, 2010*, pages 1873–1876, 2010.

[28] Soujanya Poria, Erik Cambria, Lun-Wei Ku, Chen Gui, and Alexander F. Gelbukh. A rule-based approach to aspect extraction from product reviews. In *Proceedings of the Second Workshop on Natural Language Processing for Social Media, SocialNLP@COLING, Dublin, Ireland, August 24, 2014*, pages 28–37, 2014.

[29] Minghui Qiu, Liu Yang, and Jing Jiang. Modeling interaction features for debate side clustering. In *22nd ACM International Conference on Information and Knowledge Management, CIKM'13, San Francisco, CA, USA, October 27 - November 1, 2013*, pages 873–878, 2013.

[30] J. Ross Quinlan. *C4.5 : Programs for Machine Learning*. Morgan Kaufmann, 1993.

[31] Alec Radford, Rafal Józefowicz, and Ilya Sutskever. Learning to generate reviews and discovering sentiment. *CoRR*, abs/1704.01444, 2017.

[32] Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Christopher D Manning, Andrew Ng, and Christopher Potts. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 conference on empirical methods in natural language processing*, volume 1631, pages 1631–1642, 2013.

[33] Student. The probable error of a mean. *Biometrika*, 6(1) :1–25, 1908.

[34] Ivan Titov and Ryan T. McDonald. Modeling online reviews with multi-grain topic models. In *Proceedings of the 17th International Conference on World Wide Web, WWW 2008, Beijing, China, April 21-25, 2008*, pages 111–120, 2008.

[35] Mikalai Tsytsarau, Themis Palpanas, and Malú Castellanos. Dynamics of news events and social media reaction. In *The 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '14, New York, NY, USA - August 24 - 27, 2014*, pages 901–910, 2014.

[36] Mikalai Tsytsarau, Themis Palpanas, and Kerstin Denecke. Scalable discovery of contradictions on the web. In *Proceedings of the 19th International Conference on World Wide Web, WWW 2010, Raleigh, North Carolina, USA, April 26-30, 2010*, pages 1195–1196, 2010.

[37] Mikalai Tsytsarau, Themis Palpanas, and Kerstin Denecke. Scalable detection of sentiment-based contradictions. *DiversiWeb, WWW*, 1 :9–16, 2011.

[38] Peter D. Turney. Thumbs up or thumbs down ? semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, July 6-12, 2002, Philadelphia, PA, USA.*, pages 417–424, 2002.

[39] Jan Vosecky, Kenneth Wai-Ting Leung, and Wilfred Ng. Searching for quality microblog posts : Filtering and ranking based on content analysis and implicit links. In *Database Systems for Advanced Applications - 17th International Conference, DASFAA 2012, Busan, South Korea, April 15-19, 2012, Proceedings, Part I*, pages 397–413, 2012.

[40] Lu Wang and Claire Cardie. A piece of my mind : A sentiment analysis approach for online dispute detection. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014, June 22-27, 2014, Baltimore, MD, USA, Volume 2*, pages 693–699, 2014.

[41] Lu Wang, Hema Raghavan, Claire Cardie, and Vittorio Castelli. Query-focused opinion summarization for user-generated content. In *COLING 2014, 25th International Conference on Computational Linguistics, August 23-29, 2014, Dublin, Ireland*, pages 1660–1669, 2014.

[42] Show-Jane Yen and Yue-Shi Lee. Under-sampling approaches for improving prediction of the minority class in an imbalanced dataset. pages 731–740, 2006.

[43] Quan Yuan, Gao Cong, and Nadia Magnenat-Thalmann. Enhancing naive bayes with various smoothing methods for short text classification. In *Proceedings of the 21st World Wide Web Conference, WWW 2012, Lyon, France, April 16-20, 2012 (Companion Volume)*, pages 645–646, 2012.

©AfIA 2018

Sur le Gradient de la Politique pour les Systèmes Multi-Agents Coopératifs

G. Bono[1] J. Dibangoye[1] L. Matignon[1,2] F. Pereyron[3] O. Simonin[1]

[1] Univ Lyon, INSA Lyon, INRIA, CITI, F-69621 Villeurbanne, France
[2] Univ Lyon, Université Lyon 1, LIRIS, CNRS, UMR5205, Villeurbanne, F-69622, France
[3] Volvo Group, Advanced Technology and Research

guillaume.bono@inria.fr

Résumé

L'apprentissage par renforcement (RL) pour les processus décisionnels de Markov partiellement observables décentralisés (Dec-POMDPs) accuse un certain retard par rapport aux progrès spectaculaires du RL mono-agent. Ceci s'explique en partie par un certain nombre d'hypothèses valables dans le cadre mono-agent, mais invalides dans les systèmes multi-agents. Pour combler ce retard, nous explorons les fondements mathématiques des méthodes par ascension du gradient de la politique dans le paradigme de l'entraînement centralisé pour un contrôle décentralisé (CTDC). Dans ce paradigme, l'apprentissage peut avoir lieu de façon centralisée tout en gardant la contrainte d'une exécution décentralisée. En partant de cette intuition, nous établissons dans ce document une extension multi-agents du théorème du gradient de la politique et du théorème de compatibilité des fonctions d'approximation de la valeur. Nous en tirons des méthodes « acteur critique » (AC) qui parviennent (i) à estimer le gradient de la politique à partir d'expériences collectives mais aussi (ii) à préserver le contrôle décentralisé du système à l'exécution. Nos expérimentations montrent que nos méthodes ne souffrent pas de la comparaison avec les techniques standard en RL sur un ensemble de bancs de test de la littérature.

Mots Clef

Contrôle décentralisé et stochastique – Processus Décisionnel de Markov Partiellement Observable – Systèmes Multi-Agents – Méthodes Acteur Critique

Abstract

Reinforcement Learning (RL) for decentralized partially observable Markov decision processes (Dec-POMDPs) is lagging behind the spectacular breakthroughs of single-agent RL. That is because assumptions that hold in single-agent settings are often obsolete in decentralized multi-agent systems. To tackle this issue, we investigate the foundations of policy gradient methods within the centralized training for decentralized control (CTDC) paradigm. In this paradigm, learning can be accomplished in a centralized manner while each agent can still execute its policy independently at deployment. Using this insight, we establish a new policy gradient theorem and compatible function approximations for decentralized multi-agent systems. Resulting actor critic methods preserve the decentralized control at the execution phase, but can also estimate the policy gradient from collective experiences guided by a centralized critic at the training phase. Experiments demonstrate our policy gradient methods compare favorably against standard RL techniques in benchmarks from the literature.

Keywords

Decentralized and Stochastic Control – Partially Observable Markov Decision Processes – Multi-Agent Systems – Actor Critic Methods.

1 Introduction

Ces dernières années, la capacité d'agents artificiels à apprendre des comportements par eux-mêmes en interagissant avec leur environnement s'est développée de façon spectaculaire [19, 20], promettant de grandes avancées dans la société et l'industrie. Une partie de ces progrès peut être imputée à l'apprentissage par renforcement (RL) mono-agent, particulièrement le RL profond. Il s'agit d'une branche de l'apprentissage automatique où le monde dans lequel évolue l'agent est décrit comme un processus décisionnel de Markov (MDP) [25]. Si d'autres agents interviennent, ils sont intégrés à cette description du monde, et les hypothèses considérées pendant l'apprentissage et l'exécution sont identiques. Dans ce modèle, les méthodes par ascension du gradient de la politique et les algorithmes *acteur critique* (AC) et *acteur critique naturel* (NAC) ont eu un certain succès, avec de bonnes garanties de convergence [1, 7, 15, 26]. Ces méthodes cherchent directement dans un espace de politiques stochastiques paramétrées en ajustant la valeur des paramètres dans la direction du gradient de la politique. Malheureusement, les adaptations de ces méthodes au cadre multi-agents se sont focalisées soit sur des agents indépendants [22, 29], soit sur des systèmes

 ©AfIA 2018

où la connaissance du monde est commune [32]. Ces derniers peuvent en fait être ramenés à des systèmes mono-agent.

Au lieu de cela, nous considérons dans cet article un système multi-agents coopératif où l'apprentissage est centralisé, mais où l'exécution reste décentralisée. Grâce à ce paradigme, nous pouvons relâcher la contrainte de contrôle décentralisé pendant la phase d'apprentissage. Nous préservons néanmoins l'indépendance des politiques apprises pour leur exécution. Dans de nombreux cas d'application des systèmes multi-agents, les conditions lors de l'entraînement ne sont pas aussi strictes que les conditions à l'exécution : pendant les répétitions d'une pièce de théâtre, les acteurs peuvent lire le script, prendre des pauses ou échanger avec le metteur en scène ; pour gagner un match de football, les joueurs appliquent des tactiques développées avec leur entraîneur bien avant la rencontre, etc. Il est donc naturel de se demander si les méthodes d'ascension du gradient de la politique peuvent tirer profit de ce paradigme.

Le paradigme CTDC a été utilisé avec succès à des méthodes de planification pour les MDP partiellement observables décentralisés (Dec-POMDP), un modèle de choix pour la prise de décision séquentielle par un ensemble d'agents coopérant à la poursuite d'un objectif commun [4, 9, 10, 11, 14, 21, 27, 28]. Dans la littérature de la théorie des jeux, les Dec-POMDP sont des jeux stochastiques partiellement observables à gains identiques. Les Dec-POMDP généralisent d'autres modèles multi-agents collaboratifs, comme les MDP multi-agents [6], ou les jeux stochastiques à gains identiques [24] par exemple. L'hypothèse principale qui rend les Dec-POMDP fondamentalement différents et plus complexes que les MDP pourrait n'être pertinente que pendant la phase d'exécution : les agents n'ont pas accès au véritable état du monde, et ne peuvent pas transmettre aux autres agents leurs observations locales bruitées. Rien n'empêche les agents de partager leurs informations locales pendant l'entraînement à la condition qu'à l'exécution, ils ne basent leurs décisions que sur leurs expériences individuelles. De façon assez surprenante, cette intuition a été relativement peu exploitée, et le traitement formel du paradigme CTDC a fait l'objet de peu d'attention au sein de la communauté RL [17]. Quand cet entraînement centralisé a lieu en laboratoire ou se base sur un simulateur, il est possible d'exploiter l'information disponible plus riche que celle collectée lors de l'exécution, par exemple les états cachés, les informations locales d'autres agents, etc. Même si les travaux récents en RL multi-agents – et RL profond – partent de ce paradigme pour développer des méthodes spécifiques à certains domaines [12, 13, 18], les fondements théoriques du RL multi-agents décentralisé en sont encore à leurs balbutiements.

Cet article explore les fondements théoriques des méthodes du gradient de la politique dans le paradigme CTDC. Dans ce paradigme, les algorithmes acteur critique peuvent être adaptées pour entraîner plusieurs acteurs (ou politiques) indépendants guidés par un critique centralisé (c.à.d. une approximation de la fonction de valeur état-action) [12]. Les méthodes de cette famille diffèrent seulement par la façon dont est représenté et mis à jour ce critique centralisé. Le résultat principal de ce papier est une généralisation du théorème du gradient de la politique et du théorème de compatibilité des approximations de la valeur pour les Dec-POMDP en s'inspirant de leurs équivalents en MDP. Nous montrons en particulier que le critique centralisé est compatible avec le gradient de la politique s'il se décompose en une somme de critiques individuels, eux-mêmes combinaisons linéaires des caractéristiques de la politique individuelle correspondante. De plus, nous dérivons des règles de mise à jour qui ajustent les paramètres des critiques individuels dans la direction du gradient du critique centralisé. Nos expérimentations tendent à montrer que notre approche est compétitive par rapport aux techniques issues des paradigmes classiques du RL sur un jeu de bancs de test de la littérature.

Le reste de l'article est structuré de la façon suivante. La Section 2 revient sur la définition formelle des MDP partiellement observables (POMDP) et Dec-POMDP, et rappelle un ensemble de propriétés pertinentes. Dans la Section 3, nous passons en revue les méthodes du gradient de la politique pour les POMDP (mono-agent), puis nous poursuivons par une analyse des travaux les utilisant dans le cadre multi-agents coopératif en Section 4. La Section 5 présente nos résultats sur les fondements théoriques des méthodes du gradient de la politique pour les Dec-POMDP, et en dérive des algorithmes acteur-critique. Nous concluons par la présentation et l'analyse de nos résultats empiriques dans la Section 6

2 Contexte

2.1 MDP partiellement observables

Considérons un agent jouant le rôle de coordinateur central chargé de contrôler le comportement d'un processus décisionnel de Markov partiellement observable (POMDP) au cours de son évolution dans le temps. Ce cadre sert souvent à formaliser les systèmes multi-agents coopératifs dans lesquels tous les agents se transmettent explicitement et instantanément leurs observations bruitées.

Définition 1. *Soit* $M_1 \doteq (\mathcal{X}, \mathcal{U}, \mathcal{Z}, p, R, T, s_0, \gamma)$ *un POMDP. On note* X_t*,* U_t*,* Z_t *et* R_t *les variables aléatoires qui prennent respectivement leurs valeurs dans* $\mathcal{X}$*,* $\mathcal{U}$*,* $\mathcal{Z}$ *et* $\mathbb{R}$ *et qui décrivent : l'état de l'environnement, les commandes appliquées par l'agent, les observations et le signal de récompense qu'il a reçu à l'instant* $t = 0, 1, \ldots, T$*. La dynamique de l'environnement est modélisée par les probabilités de transition et d'observation* $p(x', z|x, u) \doteq \mathbb{P}(X_{t+1} = x', Z_{t+1} = z|X_t = x, U_t = u)$*.* $R(x, u) \doteq \mathbb{E}[R_{t+1}|X_t = x, U_t = u]$ *dénote l'espérance de la récompense immédiate,* $s_0(x) = \mathbb{P}(X_0 = x)$ *la distribution initiale sur les états, et* $\gamma \in [0, 1]$ *le facteur d'amortissement des récompenses futures.*

On définit récursivement l'historique $o_t \doteq (o_{t-1}, u_{t-1}, z_t)$ avec $o_0 \doteq \emptyset$, contenant la séquence de commandes et d'observations que l'agent a effectuées et perçues jusqu'au temps t. On note $\mathcal{O}_t$ l'ensemble des historiques potentiellement expérimentés par l'agent au temps t.

Définition 2. *L'agent sélectionne à chaque instant une commande u_t selon une politique paramétrée $\pi \doteq (a_0, a_1, \ldots, a_T)$, où $a_t(u_t|o_t) \doteq \mathbb{P}(u_t|o_t;\theta_t)$ dénote la règle de décision appliquée au temps t, de paramètres $\theta_t \in \mathbb{R}^{\ell_t}$ avec $\ell_t \ll |\mathcal{O}_t|$, $\forall t \in \{0 \ldots T\}$.*

En pratique, la politique peut être construite comme par exemple un réseau de neurone profond, un contrôleur à états finis ou encore une simple structure linéaire normalisée par *softmax*. De telles représentations de la politique reposent sur différentes descriptions parfois incomplètes des historiques, appelées états internes de l'agent. Notons que si le modèle est connu ou estimé, il peut servir à calculer une forme d'états internes appelés croyances qui constituent une statistique exhaustive de l'historique [2]. Si on note $b^o \doteq \mathbb{P}(X_t|O_t = o)$ la croyance courante induit par l'historique o, avec une croyance initial $b^\emptyset \doteq s_0$, alors, après avoir effectué la commande $u \in \mathcal{U}$ et perçu l'observation $z \in \mathcal{Z}$, la nouvelle croyance se calcule ainsi : $\forall x' \in \mathcal{X}$,

$$\begin{aligned} b^{o,u,z}(x') &\doteq \mathbb{P}(X_{t+1} = x'|O_{t+1} = (o,u,z)) \\ &\propto \sum_{x\in\mathcal{X}} p(x', z|x, u) b^o(x). \end{aligned}$$

Par conséquent, utiliser les croyances au lieu des historiques dans la description des politiques préserve la capacité à agir de façon optimale, tout en réduisant significativement l'empreinte mémoire de la représentation de l'état interne. Cela permet aussi de se focaliser sur des politiques stationnaires, particulièrement intéressantes dans le cadre à horizon infini ($T = \infty$). Une politique π est dite stationnaire si $a_0 = a_1 = \ldots = a$, ou dit autrement $\theta_0 = \theta_1 = \ldots = \theta$; sinon, elle est non stationnaire.
En évoluant dans l'environnement guidé par la politique π, l'agent génère une trajectoire composée de récompenses, d'observations, de commandes et d'états $\omega_{t:T} \doteq (r_{t:T}, x_{t:T}, z_{t:T}, u_{t:T})$. Chaque trajectoire rapporte une récompense cumulée $R(\omega_{t:T}) \doteq \gamma^0 r_t + \cdots + \gamma^{T-t} r_T$. Les meilleures politiques sont celles qui produisent à partir de s_0 les meilleures récompense cumulée en espérance :

$$\begin{aligned} J(s_0;\theta_{0:T}) &\doteq \mathbb{E}_{\Omega_{0:T}\sim\mathbb{P}(\cdot|\pi,M_1)}[R(\Omega_{0:T})] \\ &= \sum_{\omega_{0:T}} \mathbb{P}(\omega_{0:T}|\pi, M_1) R(\omega_{0:T}) \end{aligned} \tag{1}$$

où on note $\mathbb{P}(\omega_{0:T}|\pi, M_1)$ la probabilité de générer la trajectoire $\omega_{0:T}$ en suivant π. Trouver le meilleur moyen pour l'agent d'interagir avec M_1 consiste à trouver le vecteur de paramètres $\theta^*_{0:T}$ solution de : $\theta^*_{0:T} \in \arg\max_{\theta_{0:T}} J(s_0;\theta_{0:T})$.

Il s'avère judicieux de subdiviser la performance d'une politique pour exploiter sa structure sous-jacente : en effet, les performances d'une politique π à partir du temps t ne dépendent des commandes antérieures à t qu'à travers les états et historiques courants. Cela nous conduit à définir les fonctions de valeur, Q-valeur et avantage de π.
La fonction de Q-valeur de π est défini par :

$$Q^\pi_t : (x, o, u) \mapsto \mathbb{E}[R(\Omega_{t:T})] \tag{2}$$

où $\Omega_{t:T} \sim \mathbb{P}(\cdot|X_t = x, O_t = o, U_t = u; \pi, M_1)$.
$Q^\pi_t(x, o, u)$ est l'espérance sur la récompense cumulée en exécutant au temps t la commande u à partir d'un état x et d'un historique o, puis en choisissant par π les futures commandes à partir de $t+1$. La fonction de valeur de π est donnée par :

$$V^\pi_t : (x, o) \mapsto \mathbb{E}_{U\sim a_t(\cdot|o)}[Q^\pi_t(x, o, U)] \tag{3}$$

où $V^\pi_t(x, o)$ est l'espérance sur la récompense cumulée en suivant la politique π à partir du temps t depuis un état x et un historique o. Enfin, la fonction avantage de π est simplement la différence des deux :

$$A^\pi_t : (x, o, u) \mapsto Q^\pi_t(x, o, u) - V^\pi_t(x, o) \tag{4}$$

où $A^\pi_t(x, o, u)$ est donc l'avantage relatif à exécuter u plutôt que de suivre π au temps t depuis un état x et un historique o, puis en revenant à la politique π par la suite. La propriété qui rend ces fonctions intéressantes est qu'elles vérifient la relation de récurrence suivante.

Lemme 1 (Équations de Bellman [3])**.** $\forall t = 0, 1, \ldots, T$, $\forall x \in \mathcal{X}, o \in \mathcal{O}_t, u \in \mathcal{U}$,

$$Q^\pi_t(x, o, u) = R(x, u) + \gamma\mathbb{E}[Q^\pi_{t+1}(X', O', U')] \tag{5}$$

où $(X', O', U') \sim \mathbb{P}(\cdot|X_t = x, O_t = o, U_t = u; a_{t+1}, M_1)$

(5) établit une relation temporelle entre les différents $V^\pi_{0:T}$, $Q^\pi_{0:T}$ et $A^\pi_{0:T}$, mais aussi $J(s_0;\theta_{0:T})$:

$$J(s_0;\theta_{0:T}) = \mathbb{E}_{X_0\sim\mathbb{P}(\cdot|s_0)}[V^\pi_0(X_0, o_0 = \emptyset)]. \tag{6}$$

Jusqu'à présent, nous nous sommes concentrés sur les systèmes contrôlés par un unique agent. La suite généralise aux systèmes décentralisés dans lesquels plusieurs agents doivent coopérer pour contrôler le même système.

2.2 POMDP Décentralisés

Considérons à présent un contexte où n agents coopèrent pour influencer l'évolution d'un système décrit comme un POMDP, mais où aucun d'eux ne peut percevoir l'état du monde, ni communiquer ses observations bruitées aux autres.

Définition 3. *Un Dec-POMDP $M_n \doteq (\mathcal{I}_n, \mathcal{X}, \mathcal{U}, \mathcal{Z}, p, R, T, \gamma, s_0)$ est défini tel que : $i \in \mathcal{I}_n$ dénote l'indice du $i^{ème}$ agent impliqué dans le processus ; $\mathcal{X}, \mathcal{U}, \mathcal{Z}, p, R, T, \gamma$ et s_0 sont définis tels que dans*

©AfIA 2018

M_1, à la particularité que $\mathcal{U}$ et $\mathcal{Z}$ se décomposent en ensembles de commandes et d'observations individuelles $\mathcal{U} = \mathcal{U}^1 \times \cdots \times \mathcal{U}^n$ et $\mathcal{Z} = \mathcal{Z}^1 \times \cdots \times \mathcal{Z}^n$, avec $\mathcal{U}^i$ et $\mathcal{Z}^i$ l'ensemble des commandes et des observations propres au $i^{ème}$ agent.

On appelle historique individuel de l'agent $i \in \mathcal{I}_n$ la séquence de commandes et d'observations individuelles exécutées et perçues jusqu'au temps $t = 0, 1, \ldots, T$, notée $o_t^i = (o_{t-1}^i, u_{t-1}^i, z_t^i)$ avec $o_0^i = \emptyset$. On note $\mathcal{O}_t^i$, l'ensemble des historiques individuels possible de l'agent i au temps t.

Définition 4. *L'agent $i \in \mathcal{I}_n$ sélectionne la commande u_t^i au temps t selon une politique paramétrée $\pi^i \doteq (a_0^i, a_1^i, \ldots, a_T^i)$ où $a_t^i(u_t^i|o_t^i) \doteq \mathbb{P}(u_t^i|o_t^i; \theta_t^i)$ est la règle de décision appliquée au temps t, de paramètres $\theta_t^i \in \mathbb{R}^{\ell_t^i}$, avec $\ell_t^i \ll |\mathcal{O}_t^i|$.*

De même que dans M_1, le nombre d'historiques individuels possibles grandit de façon exponentiel à chaque pas de temps. À ce jour, la seule statistique exhaustive connue pour les historiques individuels repose sur les états d'occupation définis par : $s_t(x, o) \doteq \mathbb{P}(x, o|\theta_{0:t-1}^{1:n}), \forall x \in \mathcal{X}, \forall o \in \mathcal{O}_t$. L'état d'occupation individuel induit par l'historique $o^i \in \mathcal{O}_t^i$ est une distribution de probabilité conditionnelle : $s_t^i(x, o^{-i}) \doteq \mathbb{P}(x, o^{-i}|o^i, s_t)$, où o^{-i} dénote l'historique joint des $n-1$ agents autres que i. Il est très difficile d'apprendre à projeter les historiques individuels vers des états internes proches des états d'occupation individuels, ce qui limite la capacité des algorithmes d'apprentissage à trouver des politiques optimales dans un temps raisonnable pour M_n. On peut alors se focaliser sur des politiques stationnaires pour lesquelles l'espace des historiques est projeté sur un ensemble fini d'états internes $\varsigma \doteq (\varsigma^1, \ldots, \varsigma^n)$, qui sont des représentations – souvent à pertes – des états d'occupations individuels. Les ς^i peuvent par exemple être les nœuds d'un contrôleur à états finis, ou encore les états internes d'un réseau de neurone récurrent (RNN). On note $\psi(\varsigma'|\varsigma, u, z)$ la probabilité de transition d'un état interne joint ς à l'état interne suivant ς' en fonction de la dernière commande effectuée u et de l'observation reçue z. Cette loi de transition est décomposable en lois individuelles $\psi^i(\varsigma'^{,i}|\varsigma^i, u^i, z^i)$. Dans la suite de cet article, nous considérerons la loi ψ fixée à priori, même si en général, elle fait partie de la politique et possède des paramètres à optimiser. Résoudre M_n revient à trouver une politique jointe $\pi \doteq (\pi^1, \ldots, \pi^n)$ – un tuple de n politiques individuelles – qui donne la meilleure récompense cumulée à partir d'une distribution d'états initiaux s_0. Autrement dit, $\theta_{0:T}^{*,1:n} \in \arg\max_{\theta_{0:T}^{1:n}} J(s_0; \theta_{0:T}^{1:n})$, avec

$$\begin{aligned} J(s_0; \theta_{0:T}^{1:n}) &\doteq \mathbb{E}_{\Omega_{0:T} \sim \mathbb{P}(\cdot|\pi, M_n)}[R(\Omega_{0:T})] \\ &= \sum_{\omega_{0:T}} \mathbb{P}(\omega_{0:T}|\pi, M_n) R(\omega_{0:T}) \end{aligned} \tag{7}$$

où $\mathbb{P}(\omega_{0:T}|\pi, M_n)$ est la probabilité de générer la trajectoire jointe $\omega_{0:T}$ en suivant la politique π. Étant donné une politique jointe π, M_n hérite des mêmes définitions que M_1, en particulier pour les fonctions $V_{0:T}^{\pi}$, $Q_{0:T}^{\pi}$ et $A_{0:T}^{\pi}$.

3 Gradient de la politique

Dans cette section, nous passons en revue la littérature des méthodes du gradient de la politique pour le contrôle centralisé de systèmes mono-agent. Dans ce contexte, l'approche du gradient de la politique se base sur des algorithmes centralisés qui cherchent à optimiser la récompense cumulée directement dans l'espace des paramètres de la politique $\theta_{0:T}$. Bien que nous n'abordions ici que des politiques non stationnaires, les méthodes citées peuvent facilement être étendues au cas de politiques stationnaires pour lesquelles $a_t = a$ pour tout t. En considérant que la politique π est différentiable par rapport à ses paramètres $\theta_{0:T}$, l'algorithme met à jour $\theta_{0:T}$ dans la direction du gradient :

$$\Delta\theta_{0:T} = \alpha \frac{\partial J(s_0; \theta_{0:T})}{\partial \theta_{0:T}} \tag{8}$$

avec α un taux d'apprentissage. En itérant sur cette règle de mise à jour, pour peu que l'estimation du gradient soit correcte, $\theta_{0:T}$ converge vers un optima local. Malheureusement, estimer correctement le gradient peut se révéler impossible. Pour dépasser cette difficulté, le gradient peut être remplacé par une estimation non biaisée, ce qui revient à restreindre (8) à un gradient stochastique :

$$\Delta\theta_{0:T} = \alpha R(\omega_{0:T}) \frac{\partial \log \mathbb{P}(\omega_{0:T}|\pi, M_1)}{\partial \theta_{0:T}} \tag{9}$$

On calcule $\frac{\partial}{\partial \theta_{0:T}} \log \mathbb{P}(\omega_{0:T}|\pi, M_1)$ sans connaissance à priori de la distribution sur les trajectoires $\mathbb{P}(\omega_{0:T}|\pi, M_1)$. En effet, les propriétés d'indépendance conditionnelle du processus donnent :

$$\mathbb{P}(\omega_{0:T}|\pi, M_1) \doteq s_0(x_0) \prod_{t=0}^{T} p(x_{t+1}, z_{t+1}|x_t, u_t) a_t(u_t|o_t) \tag{10}$$

ce qui implique :

$$\frac{\partial \log \mathbb{P}(\omega_{0:T}|\pi, M_1)}{\partial \theta_{0:T}} = \sum_{t=0}^{T} \frac{\partial \log a_t(u_t|o_t)}{\partial \theta_t} \tag{11}$$

3.1 Méthode du rapport de vraisemblance

Les méthodes du rapport de vraisemblances, *e.g.,* `Reinforce` [30], exploitent la séparabilité du vecteur de paramètres $\theta_{0:T}$, qui mènent à la règle de mise à jour suivante : $\forall t = 0, 1, \ldots, T$

$$\Delta\theta_t = \alpha \mathbb{E}_{\mathcal{D}} \left[R(\omega_{0:T}) \frac{\partial \log a_t(u_t|o_t)}{\partial \theta_t} \right] \tag{12}$$

où $\mathbb{E}_{\mathcal{D}}[\cdot]$ dénote une approximation de l'espérance par une moyenne empirique prise sur un ensemble de trajectoires $\mathcal{D} = \{\omega_{0:T,j}\}_{1 \leq j \leq m}$ échantillonnées à partir de π et M_1. Le principal problème rencontré en utilisant cette règle de mise à jour centralisé est la variance importante de $R(\omega_{0:T})$, qui peut fortement ralentir la convergence. Pour

compenser partiellement cette variance, deux constatations rapides permettent d'améliorer la règle de mise à jour. Premièrement, les futures actions ne dépendent pas des récompenses passées, *c.à.d.* $\mathbb{E}_{\mathcal{D}}[R(\omega_{0:t-1})\frac{\partial}{\partial\theta_t}\log a_t(u_t|o_t)] = 0$, On peut donc utiliser $R(\omega_{t:T})$ au lieu de $R(\omega_{0:T})$ dans 12, ce qui réduit significativement la variance de notre estimation du gradient. Deuxièmement, l'estimation du gradient de la politique n'est pas biaisée par l'introduction d'une fonction de référence $\tilde{\beta}_t$ (en anglais : *baseline*) tant que celle-ci ne dépend pas de $\theta_{0:T}$. On peut ainsi remplacer $R(\omega_{t:T})$ par une récompense relative à cette référence $R(\omega_{t:T}) - \tilde{\beta}_t(x_t, o_t)$.

3.2 Acteur Critique

Pour améliorer la variance de l'estimation du gradient dans (12), le théorème du gradient de la politique [26] propose de remplacer $R(\omega_{t:T})$ par $Q_t^w(x_t, o_t, u_t)$, une approximation de la Q-valeur associée à la commande u_t dans l'état x_t et après un historique o_t puis en complétant la trajectoire en suivant la politique π : $Q_t^w(x_t, o_t, u_t) \approx Q_t^\pi(x_t, o_t, u_t)$, où $w_t \in \mathbb{R}^{l_t}$ est un vecteur de paramètres, avec $l_t \ll |\mathcal{X}||\mathcal{O}_t||\mathcal{U}|$. Ceci conduit au schéma algorithmique *acteur critique*, dans lequel un algorithme centralisé met à jour à la fois les paramètres $\theta_{0:T}$ de la politique et les paramètres $w_{0:T}$ de l'approximation de la Q-valeur :

$$\Delta w_t = \alpha\mathbb{E}_{\mathcal{D}}\left[\delta_t \frac{\partial \log a_t(u_t|o_t)}{\partial\theta_t}\right] \tag{13a}$$

$$\Delta\theta_t = \alpha\mathbb{E}\mathcal{D}\left[Q_t^w(x_t, o_t, u_t)\frac{\partial \log a_t(u_t|o_t)}{\partial\theta_t}\right] \tag{13b}$$

où on note $\delta_t \doteq \widehat{Q}_t^\pi(x_t, o_t, u_t) - Q_t^w(x_t, o_t, u_t)$, avec $\widehat{Q}_t^\pi(x_t, o_t, u_t)$ une estimation non-biaisée de la véritable Q-valeur $Q_t^\pi(x_t, o_t, u_t)$, utilisant par exemple une estimation Monte-Carlo (MC) de $R(\omega_{0:T})$ ou bien une différence temporelle (erreur TD) $r_t + \gamma Q_{t+1}^w(x_{t+1}, o_{t+1}, u_{t+1})$. Le choix des paramètres $w_{0:T}$ est crucial pour que l'estimation du gradient reste non-biaisé [26]. Il n'y a pas de biais si les approximations de Q-valeur $Q_{0:T}^w$ sont *compatibles* avec la politique π. Sans rentrer dans les détails formels, une approximation compatible $Q_{0:T}^w$ de la véritable fonction $Q_{0:T}^\pi$ peut être une combinaison linéaire des « caractéristiques » de la politique π, et ses paramètres $w_{0:T}$ constituent la solution du problème de régression linéaire qui estime $Q_{0:T}^\pi$ à partir de ces caractéristiques. En pratique, cette seconde condition n'est pas utilisée directement, et les paramètres $w_{0:T}$ sont mis à jour en utilisant par exemple des méthodes MC ou TD.

3.3 Acteur Critique Naturel

Suivre la direction du gradient n'est pas toujours le moyen le plus rapide pour converger. Le gradient naturel propose une mise à jour des paramètres $\theta_{0:T}$ selon la direction de plus rapide ascension par rapport à la métrique d'information de Fisher :

$$\mathbf{\Phi} \doteq \mathbb{E}_{\mathcal{D}}\left[\frac{\partial \log a_t(u_t|o_t)}{\partial\theta_t}\left(\frac{\partial \log a_t(u_t|o_t)}{\partial\theta_t}\right)^\top\right] \tag{14}$$

Cette métrique est invariante à une reparamétrisation de la politique. Combiner le théorème du gradient de la politique avec une approximation compatible de la Q-valeur, puis effectuer une mise à jour des paramètres dans la direction donnée par (14), donne naissance au schéma algorithmique *acteur-critique naturel*, qui remplace (13b) dans (13) par : $\Delta\theta_t = \alpha\mathbb{E}_{\mathcal{D}}[w_t]$.

4 Gradient de la politique pour les systèmes multi-agents

Dans cette section, nous passons en revue les extensions des méthodes du gradient de la politique mono-agent à des modèles multi-agents coopératifs. On distingue trois paradigmes, illustrés dans la Figure 1 : l'entraînement centralisé pour un contrôle centralisé (CTCC), l'entraînement distribué pour un contrôle décentralisé (DTDC), et l'entraînement centralisé pour un contrôle décentralisé (CTDC).

4.1 CTCC

Dans certaines applications des systèmes multi-agents coopératifs, on peut considérer des agents pouvant communiquer entre eux de façon parfaite, instantanée et illimitée. De telles applications peuvent être modélisées par des POMDP, ce qui rend possible d'utiliser les méthodes du gradient de la politique mono-agent (Section 3). Dans ce paradigme CTCC, illustré par la Figure 1 *(gauche)*, un seul acteur « coordinateur » guidé par un critique unique est utilisé. Le principal défaut de ce paradigme fait aussi sa force : ce besoin de communication parfaite, instantanée et illimitée entre les agents d'un bout à l'autre de l'exécution, à la fois pendant l'entraînement et le déploiement.

4.2 DTDC

Assez étonnamment, il s'avère que la première adaptation des méthodes du gradient de la politique au cadre multi-agent visait à apprendre de façon distribuée les politiques à exécuter de façon décentralisée, comme par exemple l'algorithme `Reinforce` distribué [22]. Dans ce paradigme DTDC, les agents apprennent simultanément mais indépendamment leurs politiques individuelles en utilisant l'algorithme Reinforce avec chacun un critique et un acteur individuels, comme l'illustre la Figure 1 *(droite)*. L'indépendance des vecteurs de paramètres $\theta_{0:T}^1, \theta_{0:T}^2, \ldots, \theta_{0:T}^n$ donne lieu à la règle de mise à jour suivante : $\forall t = 0, 1, \ldots, T, \forall i \in \mathcal{I}_n$,

$$\Delta\theta_t^i = \alpha\mathbb{E}_{\mathcal{D}}\left[R(\omega_{0:T})\frac{\partial \log a_t^i(u_t^i|o_t^i)}{\partial\theta_t^i}\right] \tag{15}$$

Il est à noter que la somme des gradients des politiques individuelles est un estimateur non-biaisé du gradient de la politique jointe. Néanmoins, la question reste ouverte de savoir comment tirer profit des méthodes acteur-critique (voir Section 3) pour réduire la variance de cet estimateur joint. De plus, l'algorithme `Reinforce` distribué est restreint à un apprentissage où l'échantillonnage des trajectoires est dicté par la politique apprise π (*on-policy*), quand

©AfIA 2018

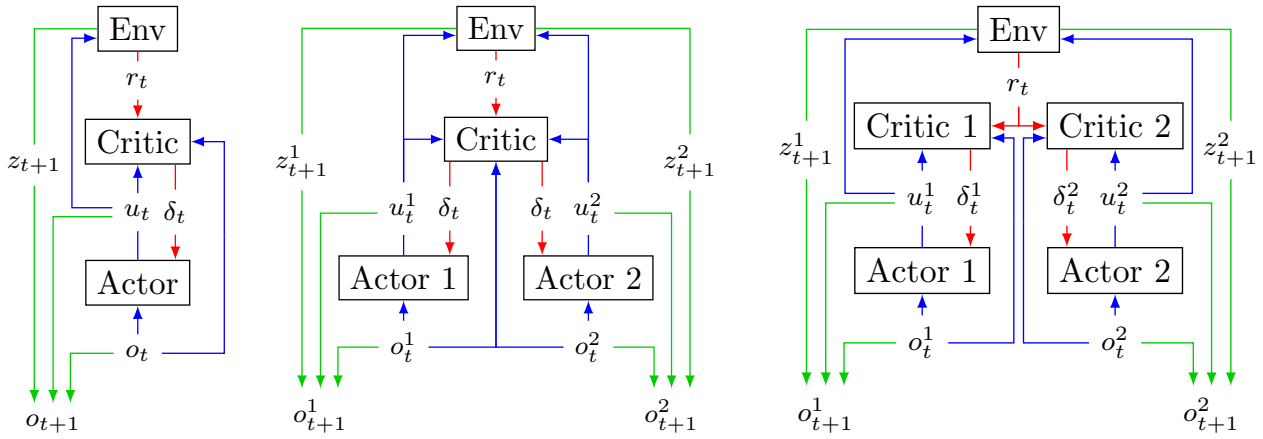


FIGURE 1 – Schéma algorithmique acteur-critique dans un modèle à deux agents pour les paradigmes : (*gauche*) CTCC, (*centre*) CTDC, et (*droite*) DTDC. Pour chaque figure, les flèches bleues représentent les commandes des agents sur l'environnement, les vertes montrent l'agrégation d'information pour le prochain pas de décision, et les rouges indiquent le signal de récompense propagé pour mettre à jour les différents paramètres.

des politiques d'exploration mieux choisies $\bar{\pi}$ peuvent fortement améliorer la qualité de l'optimum local atteint à la convergence [7].

4.3 CTDC

Le paradigme CTDC a été employé avec succès pour résoudre des problèmes de planification de Dec-POMDP [4, 9, 10, 11, 14, 17, 21, 27, 28]. Dans un tel paradigme, un coordinateur central apprend pour l'ensemble des agents pendant la phase d'entraînement puis assigne les politiques individuelles apprises aux agents correspondants avant le début de la phase d'exécution. Comme le montre la Figure 1 *(centre)*, les algorithmes acteur critique dans ce paradigme CTDC mettent à jour un critique central mais optimisent un acteur par agent. Des travaux récents en apprentissage par renforcement (profond) [12, 13, 18] exploitent ce paradigme, en particulier l'approche par gradient de la politique multi-agents contrefactuel, et l'algorithme COMA qui en découle. Malheureusement, cette algorithme se focalise sur un critique central apportant une solution au problème de l'assignation des crédits. Pour un Dec-POMDP M_n (faiblement) séparable [8], il est possible d'apprendre une contribution relative de chaque agent à la valeur de la politique jointe, apportant ainsi une réponse au problème d'assignation du crédit. En revanche, en général, la dynamique et les récompenses d'un Dec-POMDP peuvent être fortement corrélées, ce qui rend difficile voire impossible d'accorder du crédit à tel ou tel agent. Contrairement aux méthodes de planification, les méthodes d'ascension du gradient de la politique dans ce paradigme CTDC manquent de fondements théoriques, qui permettraient de définir la forme du critique centralisé et les règles de mise à jour préservant certaines garanties formelles.

5 Gradient de la politique pour les Dec-POMDP

Dans cette section, nous apportons une réponse aux limitations des paradigmes CTCC et DTDC, et étendons les schémas algorithmiques acteur-critique et acteur-critique naturel de M_1 vers M_n.

5.1 Théorème du Gradient de la Politique

Notre principal résultat est une extension du théorème du Gradient de la politique [26] de M_1 vers M_n. Avant de poursuivre, nous commençons par exposer quelques résultats préliminaires qui nous permettrons d'établir notre résultat principal de la section. Toutes les preuves sont disponibles dans le rapport de recherche [5].
Nous commençons par établir une expression des dérivées partielles des fonctions de valeurs $V^{\pi}_{0:T}$ par rapport aux vecteurs de paramètres $\theta^{1:n}_{0:T}$ dans le cas à horizon fini.

Lemme 2. *Soit un Dec-POMDP M_n, une politique jointe à évaluer $\pi \doteq (a_0, \ldots, a_T)$ et une politique d'exploration $\bar{\pi} \doteq (\bar{a}_0, \ldots, \bar{a}_T)$. Pour tout temps $t = 0, 1, \ldots, T$, agent $i \in \mathcal{I}_n$, état caché $x_t \in \mathcal{X}$, et historique joint $o_t \in \mathcal{O}_t$:*

$$\frac{\partial V^{\pi}_t(x_t, o_t)}{\partial \theta^i_t} = \mathbb{E}_{U_t \sim \bar{a}_t(\cdot|o_t)}\Big[\frac{a_t(U_t|o_t)}{\bar{a}_t(U_t|o_t)} Q^{\pi}_t(x_t, o_t, U_t) \frac{\partial \log a^i_t(U^i_t|o^i_t)}{\partial \theta^i_t}\Big]. \quad (16)$$

Voici à présent le résultat principal de la section.

Théorème 1. *Soit un Dec-POMDP M_n, une politique à évaluer $\pi \doteq (a_0, \ldots, a_T)$ et une politique d'exploration $\bar{\pi} \doteq (\bar{a}_0, \ldots, \bar{a}_T)$.*

1. *Dans le cas à horizon fini $T < \infty$, pour tout temps*

$t = 0, 1, \ldots, T$ et agent $i \in \mathcal{I}_n$:

$$\frac{\partial J(s_0; \theta^{1:n}_{0:T})}{\partial \theta^i_t} = \gamma^t \mathbb{E}_{(X_t, O_t, U_t) \sim \mathbb{P}(\cdot|\bar{a}_t, M_n)} \Big[\frac{a_t(U_t|O_t)}{\bar{a}_t(U_t|O_t)} Q^\pi_t(X_t, O_t, U_t) \frac{\partial \log a^i_t(U^i_t|O^i_t)}{\partial \theta^i_t} \Big]$$

2. *Dans le cas à horizon infini $T = \infty$, pour tout agent $i \in \mathcal{I}_n$:*

$$\frac{\partial J(s_0; \theta^{1:n})}{\partial \theta^i} = \mathbb{E}_{(X, \Sigma, U) \sim \mathbb{P}(\cdot|\bar{s}, \bar{a})} \Big[\frac{a(U|\Sigma)}{\bar{a}(U|\Sigma)} Q^\pi(X, \Sigma, U) \frac{\partial \log a^i(U^i|\Sigma^i)}{\partial \theta^i} \Big]$$

où

$$\bar{s}(x, \varsigma) \doteq \sum_{t=0}^{\infty} \gamma^t \mathbb{P}(X_t = x, \Sigma_t = \varsigma | M_n, \bar{a}, \psi, X_0, \varsigma_0)$$

Le théorème du gradient de la politique pour M_n (Théorème 1) est différent de celui pour M_1 [26]. Ce dernier part de l'hypothèse qu'un unique agent apprend à agir dans un (PO)MDP. À l'inverse, le Théorème 1 s'applique à un ensemble d'agents apprenant à contrôler l'évolution d'un POMDP de façon décentralisée. Les agents agissent indépendamment, mais leur estimation du gradient de la politique est guidée par une évaluation centralisée des fonctions $Q^\pi_{0:T}$. Pour utiliser cette propriété en pratique, il faut remplacer la véritable Q-valeur $Q^\pi_{0:T}$ par une approximation. Pour garantir que cette approximation est compatible – *c.à.d.* que le gradient calculé avec l'approximation est toujours dans la même direction que le véritable gradient – il faut lui donner une structure particulière que nous détaillons ci-dessous.

5.2 Approximations compatibles

Le résultat principal de cette section caractérise la forme des approximations compatibles $V^\sigma_{0:T}$ et $A^\nu_{0:T}$ respectivement pour les fonctions de valeur $V^\pi_{0:T}$ et les fonctions d'avantage $A^\pi_{0:T}$ pour un Dec-POMDP quelconque M_n. Ensemble, ces deux approximations permettent d'évaluer $Q^\pi_{0:T}(x_t, o_t, u_t) \doteq V^\pi_t(x_t, o_t) + A^\pi_t(x_t, o_t, u_t)$, pour tout temps $t = 0, 1, \ldots, T$, historique joint $o_t \in \mathcal{O}_t$ et commande jointe $u_t \in \mathcal{U}$.

Théorème 2. *Soit un Dec-POMDP M_n, des fonctions d'approximation $V^\sigma_{0:T}$ et $A^\nu_{0:T}$, de paramètres respectifs $\sigma^{1:n}_{0:T}$ et $\nu^{1:n}_{0:T}$. Ces approximations sont compatibles avec la politique jointe $\pi \doteq (a_0, \ldots, a_T)$ paramétrée par $\theta^{1:n}_{0:T}$ si l'une des deux conditions suivantes est vérifiée $\forall t = 0, 1, \ldots, T$:*

1. *$\forall i \in \mathcal{I}_n, \forall x_t \in \mathcal{X}, \forall o_t \in \mathcal{O}_t$,*

$$\frac{\partial V^\sigma_t(x_t, o_t)}{\partial \sigma^i_t} = \mathbb{E}_{U^i_t \sim a^i_t(\cdot|o^i_t)} \Big[\frac{\partial \log a^i_t(U^i_t|o^i_t)}{\partial \theta^i_t} \Big] \tag{17}$$

et σ minimise l'erreur quadratique moyenne (MSE) $\mathbb{E}[\epsilon_t(X_t, O_t, U_t)^2]$

2. *$\forall i \in \mathcal{I}_n, \forall x_t \in \mathcal{X}, \forall o_t \in \mathcal{O}_t, \forall u_t \in \mathcal{U}$,*

$$\frac{\partial A^\nu_t(x_t, o_t, u_t)}{\partial \nu^i_t} = \frac{\partial \log a^i_t(u^i_t|o^i_t)}{\partial \theta^i_t} \tag{18}$$

et ν minimise l'erreur quadratique moyenne (MSE) $\mathbb{E}[\epsilon_t(X_t, O_t, U_t)^2]$

où $\epsilon_t(x_t, o_t, u_t) \doteq Q^\pi_t(x_t, o_t, u_t) - V^\sigma_t(x_t, o_t) - A^\nu_t(x_t, o_t, u_t)$.

Quand l'une de ces conditions est vérifiée, on a bien, pour toute politique d'exploration $\bar{\pi} \doteq (\bar{a}_0, \ldots, \bar{a}_T)$:

$$\frac{\partial V^\pi_t(x_t, o_t)}{\partial \theta^i_t} = \mathbb{E}_{U_t \sim \bar{a}_t(\cdot|o_t)} \Big[\frac{a_t(U_t|o_t)}{\bar{a}_t(U_t|o_t)} \big(V^\sigma_t(x_t, o_t) + A^\nu_t(x_t, o_t, U_t) \big) \frac{\partial \log a^i_t(U^i_t|o^i_t)}{\partial \theta^i_t} \Big] \tag{19}$$

Le Théorème 2 est énoncé ci-dessus pour des politiques non-stationnaires et pour des problème à horizon fini $T < \infty$. Le résultat s'étend néanmoins très naturellement au cas à horizon infini et pour des politiques stationnaires où $a^i_t = a^i, \theta^i_t = \theta^i \ \forall t = 0, 1, \ldots, \infty, \forall i \in \mathcal{I}_n$. Le théorème montre comment les conditions de compatibilité d'une approximation de la Q-valeur pour M_1 se généralisent dans le cas d'un Dec-POMDP M_n. Parmi les propriétés notables de ces approximations centralisées compatibles, il faut relever leur séparabilité :

$$V^\sigma_t \colon (x_t, o_t) \mapsto \sum_{i \in I_n} \mathbb{E} \left[\frac{\partial \log a^i_t(U^i_t|o^i_t)}{\partial \theta^i_t} \right]^\top \sigma^i_t \tag{20}$$

dans le premier cas, ou alors

$$A^\nu_t \colon (x_t, o_t, u_t) \mapsto \sum_{i \in I_n} \left(\frac{\partial \log a^i_t(u^i_t|o^i_t)}{\partial \theta^i_t} \right)^\top \nu^i_t \tag{21}$$

dans le second. Quel que soit le cas, à l'instar des résultats existants pour M_1, l'approximation A^ν ou V^σ dont la structure n'est pas contrainte constitue un degré de liberté très important pour tenter de réduire la variance de l'estimation du gradient. Dans notre paradigme CTDC, elle peut entre autre accéder aux état cachés x_t, ou encore combiner de façon plus intriquée l'information jointe disponible. Par ailleurs, il est tout à fait envisageable d'exploiter la séparabilité des deux approximations compatibles :

$$Q^\pi_t(x_t, o_t, u_t) \approx \sum_{i \in I_n} \left(\frac{\partial \log a^i_t(u^i_t|o^i_t)}{\partial \theta^i_t} \right)^\top \nu^i_t + \sum_{i \in I_n} \mathbb{E} \left[\frac{\partial \log a^i_t(U^i_t|o^i_t)}{\partial \theta^i_t} \right]^\top \sigma^i_t + \tilde{\beta}_t(x_t, o_t, u_t) \tag{22}$$

avec $\tilde{\beta}_t$ une fonction de référence arbitraire, ne dépendant ni de ν, ni de σ, ni de θ et ne perturbant donc pas le gradient. Notons enfin que la séparabilité du critique centralisé ne nous permet pas de considérer les critiques individuels indépendamment l'un de l'autre, l'approximation du gradient est toujours jointe et guidée par cet unique critique centralisé.

©AfIA 2018

5.3 Algorithmes Acteur Critique pour le Contrôle Décentralisé

Dans cette section, nous utilisons les résultats du Théorème 2 pour donner naissance à un schéma algorithmique acteur-critique pour M_n adapté au paradigme CTDC, que nous appelons *acteur critique pour le contrôle décentralisé* (ACDC). Cet algorithme ne nécessite pas de connaissance à priori du modèle, et l'échantillonnage des trajectoires peut se servir d'une politique d'exploration différente de la politique à évaluer. Il est centralisé [1] et itératif. Chaque itération consiste en une étape d'évaluation de la politique apprise, suivi d'une étape d'amélioration de celle-ci. L'étape d'évaluation crée une base d'échantillons de trajectoires (mini-batch) $\mathcal{D}$ à partir de $\mathbb{P}(\Omega_{0:T}|\pi, M_n)$ et stocke en mémoire les erreurs TD correspondantes (*voir* lignes 6–11). L'étape d'amélioration met à jour l'ensemble des paramètres θ, ν, et σ dans la direction du gradient estimé à partir de la moyenne empirique sur la base $\mathcal{D}$ en exploitant la séparabilité des approximations compatibles (*voir* lignes 12–16).

Algorithm 1: Actor-Critic for Decentralized Control (ACDC).

1 `ACDC()`
2 Initialize $\theta_0, \nu_0, \sigma_0$ arbitrarily and $h \leftarrow 0$.
3 **while** θ_h *has not converged* **do**
4 `evaluation()` and `improvement()`
5 $h \leftarrow h + 1$

6 `evaluation()`
7 Initialize $\mathcal{D}^h_{0:T} \leftarrow \emptyset$
8 **for** $j = 1 \dots m$ *and* $t = 0 \dots T$ **do**
9 Create trajectories $(x_{t:t+1}, o_{t:t+1}, u_t) \sim p(\cdot|\theta^{1:n}_{0:t})$
10 Evaluate $\delta_t \leftarrow r_t + \gamma V^\sigma_{t+1}(x_{t+1}, o_{t+1}) - V^\sigma_t(x_t, o_t)$
11 Compose batch
$\mathcal{D}_{t,h} \leftarrow \{(o_t, u_t, \delta_t, a_t(u_t|o_t)/\bar{a}_t(u_t|o_t))\} \cup \mathcal{D}_{t,h}$

12 `improvement()`
13 **for** $i = 1 \dots n$ **do**
14 Baseline update
$\sigma^i_{t,h+1} \leftarrow \sigma^i_{t,h} + \alpha^\sigma_h \mathbb{E}_{\mathcal{D}_{t,h}}\{\delta_t \frac{a_t(u_t|o_t)}{\bar{a}_t(u_t|o_t)} \phi^i_t(o^i_t)\}$
15 Critic update
$\nu^i_{t,h+1} \leftarrow \nu^i_{t,h} + \alpha^\nu_h \mathbb{E}_{\mathcal{D}_{t,h}}\{\delta_t \frac{a_t(u_t|o_t)}{\bar{a}_t(u_t|o_t)} \phi^i_t(o^i_t, u^i_t)\}$
16 Actor update
$\theta^i_{t,h+1} \leftarrow \theta^i_{t,h} + \alpha^\theta_h \mathbb{E}_{\mathcal{D}_{t,h}}\{\frac{a_t(u_t|o_t)}{\bar{a}_t(u_t|o_t)} \phi^i_t(o^i_t, u^i_t)\nu^i_{t,h}\}$

Les taux d'apprentissage α^θ_h, α^ν_h et α^σ_h évoluent selon les conditions standard de Robbins et Monro pour les algorithmes d'approximation stochastiques [23], *c.à.d.* $\sum_{h=0}^\infty \alpha_h = \infty$, $\sum_{h=0}^\infty \alpha_h^2 < \infty$. De plus, si l'on suit les recommandations de [16], ces taux doivent être mis à jour à chaque itération tel que les paramètres des acteurs θ soient modifiés « plus lentement » que les paramètres ν and σ pour garantir la convergence. Pour faciliter la convergence d'une politique jointe pour une modification constante de ses paramètres, une méthode de choix est l'utilisation gradient naturel [1, 15]. L'algorithme ACDC naturel (NACDC) ne diffère de sa version initiale que par la formule de mise à jour des acteurs :

$$\theta^i_{t,h+1} \leftarrow \theta^i_{t,h} + \alpha^\theta_h \mathbb{E}_{\mathcal{D}_{t,h}}[\frac{a_t(u_t|o_t)}{\bar{a}_t(u_t|o_t)} \nu^i_{t,h}]$$

Pour conclure, cette section, quelques remarques sur les propriétés théoriques des algorithmes ACDC. D'une part, moyennant quelques conditions mineures, ils convergent avec probabilité 1 vers un optimum local puisque ce sont de véritables algorithmes d'ascension du gradient [7]. L'argument repose sur le fait qu'ils minimisent l'erreur de projection quadratique par descente de gradient stochastique, mais nous renvoyons le lecteur vers [7] pour plus de détails. D'autre part, ils terminent sur un optimum local qui est aussi un équilibre de Nash. En effet les dérivées partielles du critique centralisé par rapport à n'importe lequel de ses paramètres ne sont nulles qu'en un point d'équilibre, qui est aussi un optimum local.

6 Expérimentations

Dans cette section, nous montrons empiriquement les avantages du paradigme CTDC sur les paradigmes classiques CTCC et DTDC. Nos résultats semblent indiquer que les méthodes ACDC se comparent favorablement aux algorithmes existants sur de nombreux domaines multi-agents décentralisés de la littérature. Nous illustrons également les limitations de l'implémentation actuelle qui l'empêchent d'obtenir de meilleures performances.

6.1 Conditions expérimentales

Comme nous l'avons laisser transparaître tout au long de cet article, plusieurs facteurs clés peuvent affecter les performances des méthodes acteur-critique. Parmi ceux-ci, nous relèverons : le paradigme d'entraînement utilisé (CTCC *vs* DTDC *vs* CTDC) ; la représentation de la politique (stationnaire *vs* non-stationnaire) ; les structures d'approximation (linéaires *vs* réseaux de neurones récurrents (RNN)) ; la représentation des historiques (troncatures à mémoire finie *vs* états cachés d'un RNN). Nous implémentons trois variantes des méthodes acteur-critique qui combinent ces différents facteurs. Sauf mention contraire, nous ferons référence à chaque variante mise en œuvre par le nom du paradigme qu'elle implémente, par exemple CTDC pour notre algorithme ACDC, suivi de la structure de représentation interne utilisée quand cela est pertinent, par exemple : « *CTDC_trunc(K)* » pour un algorithme ACDC utilisant un historique tronqué des K dernières observations comme entrée d'une politique linéaire non-stationnaire, ou encore « *DTDC_rnn* » pour l'algorithme Reinforce distribué utilisant un réseau de neurone récurrent dans une politique stationnaire (voir Figure 2).

1. Il est néanmoins possible de le faire tourner de façon distribuée en donnant le moyen aux agents de collaborer pendant l'entraînement en se communiquant leur informations locales.

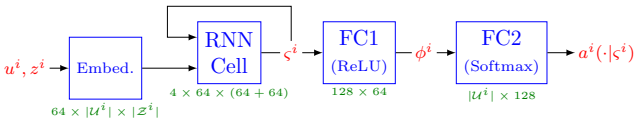


FIGURE 2 – Réseau de neurones récurrent utilisé comme structure « acteur » de chaque agent $i \in \mathcal{I}_n$. Une couche LSTM met à jour un état caché interne à partir du précédent et d'une encapsulation (embedding) d'une paire commande-observation. Elle est suivie d'un perceptron dont les activations sont rectifiées linéairement (ReLU) qui génère un vecteur de caractéristiques (features) ϕ^i, qui sont combinées linéairement par un second perceptron dont la sortie est normalisée par un softmax pour donner la règle de décision conditionnelle $a^i(\cdot|\varsigma^i)$.

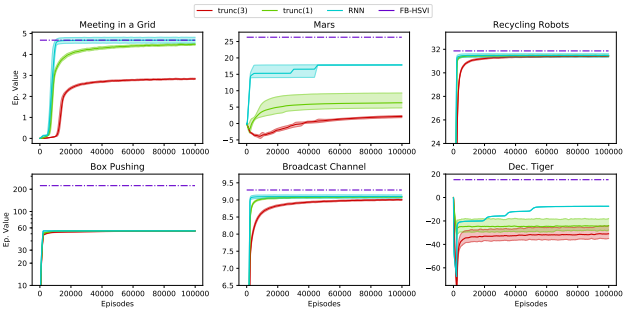


FIGURE 3 – Comparaison des différentes structures de représentation des historiques à $T = 10$.

Nous avons mené nos expérimentations sur une station de calcul *Dell Precision Tower 7910* équipé d'un CPU *Intel Xeon* à 16 cœurs cadencés à 3GHz, de 16Go de RAM et d'une carte graphique *nVIDIA Quadro K620* munie de 2Go de mémoire vidéo. Nous avons simulé plusieurs bancs de test standard de la littérature des Dec-POMDP, – *Dec. Tiger*, *Broadcast Channel*, *Mars*, *Box Pushing*, *Meeting in a Grid*, et *Recycling Robots* pour les citer – tel qu'ils sont défini sur `http://masplan.org`. Le détail des meta-paramètres utilisés est donné dans le rapport de recherche [5].

6.2 Importance de la représentation de l'état interne

Dans cette section, nous mettons en place des expériences visant à comprendre la façon dont la représentation des historiques affectent les performances des méthodes `ACDC`. La Figure 3 comparent les récompenses cumulées moyennes obtenues avec des historiques tronqués de taille 1 et 3, celles obtenues avec des RNN, et la performance ϵ-optimal donnée par l'algorithme de planification centralisé `FB-HSVI` [11]. Pour des horizons courts, ici $T = 10$, CTDC rnn converge rapidement vers de bonnes solutions par rapport à *CTDC_trunc(*1*)* et *CTDC_trunc(*3*)*. Ceci laisse penser que *CTDC_rnn* apprends des représentations des historiques plus concises et discriminantes que la représentation tronquée. Il semblerait néanmoins que cette différence de performances s'amenuise quand l'horizon de planification augmente (non illustré ici). On notera néanmoins une certaine instabilité de la récompense cumulée empirique pour les RNN sur les tâches plus complexe comme *Dec. Tiger*, un banc de test mettant en avant l'importance de la collecte d'information avant de prendre une action décisive, où les pénalités en cas d'erreur sont, en valeur absolues, bien plus élevées que les récompenses en cas de succès.

Dans l'ensemble, nos expériences montrent un fort potentiel des RNN pour représenter les historiques. Ils ont l'avantage par rapport aux historiques tronqués d'apprendre automatiquement des classes d'équivalence et une représentation compacte de l'état interne, en se basant seulement sur la retro-propagation du gradient issu du signal de récompense. Il est aussi important de noter ici l'importance du choix des méta-paramètres, en particulier le taux d'apprentissage, qui, s'il est trop élevé, encourage des convergences prématurées vers des optima locaux insatisfaisants, sans laisser le temps à la politique d'explorer le reste des trajectoires possibles. Attention également à certaines propriétés spécifiques des modèles considérés, pour lesquelles les méthodes par ascension de gradient peinent à échapper au piège d'optima locaux. Nous n'avons pas pu identifier avec certitude quelles caractéristiques particulières avaient le plus d'impact négatif sur les performances, et nous laissons à de futures études l'analyse et l'exploration de méthodes plus fiables pour entraîner ces architectures.

6.3 Comparaison des différents paradigmes

Dans cette section, nous comparons les paradigmes CTCC, DTDC et CTDC en utilisant indifféremment les représentations ayant données les meilleurs résultats dans chaque cas. Nous incluons également à la comparaison les résultats issus de deux autres algorithmes de la littérature Dec-POMDP : un algorithme de planification ϵ-optimal appelé `FB-HSVI` [11], et un algorithme de planification par échantillonnage appelé Espérance-Maximisation Monte-Carlo (`MCEM`) [31] ; un choix justifié par de nombreuses similarités avec les méthodes acteur critique. Nous soulignons le fait que nous ne cherchons pas à égaler les performances de `FB-HSVI` qui est un algorithme de planification centralisé nécessitant la connaissance à priori du modèle. Il nous fournit des performances de référence très proches de l'optimal global. Quant à `MCEM`, les résultats sont repris de [31] [2].

Pour presque tous les bancs de test, CTDC semble prendre le meilleur sur les deux autres paradigmes, que ce soit à horizon fini ou infini. CTCC pâtit probablement du fléau de la dimension dans l'espace des historiques joints, et ne parvient pas à l'explorer de façon suffisamment efficace avant que les taux d'apprentissage rendent les mises à jour des paramètres négligeables, ou que le nombre maximum

2. Sauf pour *Meeting in a Grid* et *Broadcast Channel* pour lesquels les valeurs indiquées dans [31] étaient bien au delà de l'optimal, laissant penser à une anomalie.

©AfIA 2018

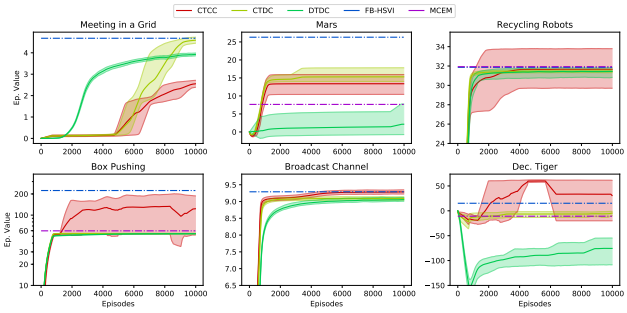


FIGURE 4 – Comparaison des paradigmes pour $T = 10$.

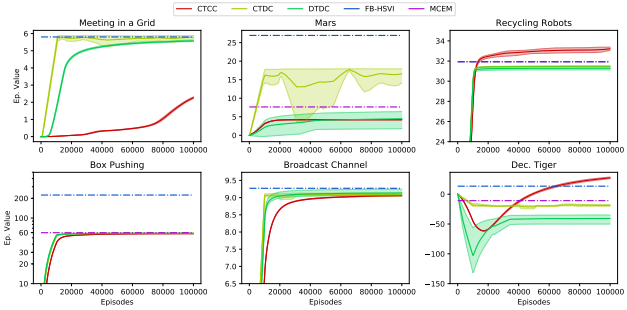


FIGURE 5 – Comparaison des paradigmes pour $T = \infty$.

d'itérations fixé au départ ne soit atteint. Le fait d'utiliser la politique apprise comme politique d'échantillonnage (on-policy) amplifie très certainement cet effet. Ayant un espace d'historiques individuels bien plus réduit à explorer, CTDC donne de meilleur résultats que CTCC dans ces expérimentations. Comparé à DTDC qui explore lui-aussi des espaces d'historiques de dimensions plus raisonnables, le paradigme CTDC a un net avantage, et l'utilisation d'un critique centralisé compatible donnent de meilleures performances sur la vitesse de convergence et la qualité de l'optimum local atteint. Bien que CTDC donnent des résultats favorables – ou au moins comparables – par rapport à l'algorithme de l'état de l'art `MCEM`, il y a encore une large marge d'amélioration pour atteindre les optima globaux donnés par `FB-HSVI` pour tous les bancs de test. Comme mentionné précédemment, c'est en partie dû à une compression encore imparfaite des historiques, mais aussi à des limitations intrinsèques des méthodes par ascension de gradient, qui ne peuvent garantir qu'un optimum local.

7 Conclusion

Cet article pose les fondations théoriques des méthodes acteur-critique pour les Dec-POMDP dans le paradigme CTDC. Dans ce paradigme, un algorithme acteur-critique centralisé apprend des politiques indépendantes, une par agent, guidé par un unique critique joint. Nous montrons qu'un critique centralisé compatible peut s'écrire comme la somme de critiques individuels, qui sont chacun des combinaisons linéaires des « caractéristiques » de la politique individuelle correspondante. Nos expérimentations montrent que nos méthodes acteur-critique appelées `ACDC` se démarquent favorablement des approches standard du RL pour un certain nombre de bancs de test de la littérature. L'implémentation actuelle de `ACDC` soulève deux problèmes bien connu du domaine : le compromis exploitation-exploration et la représentation des états internes individuels. En particulier pour ce dernier point, apprendre à projeter les historiques individuels vers des états internes proches des états d'occupation individuels est un défi de taille, auquel nous comptons continuer à contribuer dans le futur. Outre ce problème de représentation des historiques individuels, `ACDC` peut exploiter la séparabilité du critique joint compatible pour passer à l'échelle d'un assez grand nombre d'agents. Nous nous intéressons à présent à une application multi-agents décentralisée à large échelle, pour laquelle nous cherchons à exploiter cette propriété.

Références

[1] Shun-Ichi AMARI. "Natural Gradient Works Efficiently in Learning". In : *Neural Comput.* 10.2 (1998). ISSN : 0899-7667.

[2] Karl J ASTRÖM. "Optimal Control of Markov Decision Processes with Incomplete State Estimation". In : *Journal of Mathematical Analysis and Applications* 10 (1965).

[3] Richard E BELLMAN. "The Theory of Dynamic Programming". In : *Bulletin of the American Mathematical Society* 60.6 (1954).

[4] Daniel S BERNSTEIN et al. "The Complexity of Decentralized Control of Markov Decision Processes". In : *Mathematics of Operations Research* 27.4 (2002).

[5] Guillaume BONO et al. *On the Study of Cooperative Multi-Agent Policy Gradient*. Research Report RR-9188. INSA Lyon, INRIA, 2018.

[6] Craig BOUTILIER. "Planning, Learning and Coordination in Multiagent Decision Processes". In : *Proc. of the Sixth Conf. on Theoretical Aspects of Rationality and Knowledge*. 1996.

[7] Thomas DEGRIS, Martha WHITE et Richard S. SUTTON. "Linear Off-Policy Actor-Critic". In : *Proc. of the 29th Int. Conf. on ML, ICML 2012, Edinburgh, Scotland, UK, June 26 - July 1, 2012*. 2012.

[8] Jilles Steeve DIBANGOYE et al. "Exploiting Separability in Multi-Agent Planning with Continuous-State MDPs". In : *Proc. of the Thirteenth Int. Conf. on Autonomous Agents and Multiagent Systems*. 2014.

[9] Jilles Steeve DIBANGOYE et al. "Optimally Solving Dec-POMDPs As Continuous-state MDPs". In : *Proc. of the Twenty-Fourth Int. Joint Conf. on AI*. 2013.

[10] Jilles Steeve DIBANGOYE et al. *Optimally solving Dec-POMDPs as Continuous-State MDPs : Theory and Algorithms*. Research Report RR-8517. 2014.

[11] Jilles S DIBANGOYE et al. "Optimally Solving Dec-POMDPs as Continuous-State MDPs". In : *Journal of AI Research* 55 (2016).

[12] Jakob FOERSTER et al. "Counterfactual Multi-Agent Policy Gradients". In : (2018).

[13] Jayesh K. GUPTA, Maxim EGOROV et Mykel KOCHENDERFER. "Cooperative Multi-agent Control Using Deep Reinforcement Learning". In : *Autonomous Agents and Multiagent Systems*. 2017.

[14] Eric A HANSEN, Daniel S BERNSTEIN et Shlomo ZILBERSTEIN. "Dynamic Programming for Partially Observable Stochastic Games". In : *Proc. of the Nineteenth National Conf. on AI*. 2004.

[15] Sham KAKADE. "A Natural Policy Gradient". In : *Advances in Neural Information Processing Systems 14 (NIPS 2001)*. 2001.

[16] Vijay R. KONDA et John N. TSITSIKLIS. "Actor-Critic Algorithms". In : *Adv. in Neural Information Processing Systems 12*. 2000.

[17] Landon KRAEMER et Bikramjit BANERJEE. "Multi-agent reinforcement learning as a rehearsal for decentralized planning". In : *Neurocomputing* 190 (2016).

[18] Ryan LOWE et al. "Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments". In : *Adv. in Neural Information Processing Systems 30*. 2017.

[19] Volodymyr MNIH et al. "Human-level control through deep reinforcement learning". In : *Nature* 518.7540 (fév. 2015). ISSN : 0028-0836.

[20] Matej MORAVČÍK et al. "DeepStack : Expert-level artificial intelligence in heads-up no-limit poker". In : *Science* 356.6337 (2017).

[21] Frans A OLIEHOEK et al. "Incremental Clustering and Expansion for Faster Optimal Planning in Dec-POMDPs". In : *Journal of AI Research* 46 (2013).

[22] Leonid PESHKIN et al. "Learning to Cooperate via Policy Search". In : *Sixteenth Conf. on Uncertainty in Artificial Intelligence (UAI-2000)*. 2000.

[23] H ROBBINS et S MONRO. "A stochastic approximation method". In : *The annals of mathematical statistics* 22.3 (1951).

[24] Yoav SHOHAM et Kevin LEYTON-BROWN. *Multiagent Systems : Algorithmic, Game-Theoretic, and Logical Foundations*. New York, NY, USA, 2008. ISBN : 0521899435.

[25] Richard S SUTTON et Andrew G BARTO. *Introduction to Reinforcement Learning*. 2nd. Cambridge, MA, USA, 2016. ISBN : 0262193981.

[26] Richard S SUTTON et al. "Policy Gradient Methods for Reinforcement Learning with Function Approximation". In : *Proc. of the 12th Int. Conf. on Neural Information Processing Systems*. Cambridge, MA, USA, 1999.

[27] Daniel SZER et François CHARPILLET. "An Optimal Best-First Search Algorithm for Solving Infinite Horizon DEC-POMDPs". In : *Proc. of the Fifteenth European Conf. on ML*. 2005.

[28] Daniel SZER, François CHARPILLET et Shlomo ZILBERSTEIN. "MAA* : A Heuristic Search Algorithm for Solving Decentralized POMDPs". In : *Proc. of the Twenty-First Conf. on Uncertainty in AI*. 2005.

[29] Ming TAN. "Multi-agent Reinforcement Learning : Independent vs. Cooperative Agents". In : *Readings in Agents*. San Francisco, CA, USA, 1998.

[30] Ronald J WILLIAMS. "Simple statistical gradient-following algorithms for connectionist reinforcement learning". In : *ML* 8.3 (1992).

[31] Feng WU, Shlomo ZILBERSTEIN et Nicholas R JENNINGS. "Monte-Carlo Expectation Maximization for Decentralized POMDPs". In : *Proc. of the Twenty-Fourth Int. Joint Conf. on AI*. 2013.

[32] Xinhua ZHANG, Douglas ABERDEEN et S. V. N. VISHWANATHAN. "Conditional Random Fields for Multi-agent Reinforcement Learning". In : *Proc. of the 24th Int. Conf. on ML* (2007).

©AfIA 2018

Classification d'images en apprenant sur des échantillons positifs et non labélisés avec un réseau antagoniste génératif

F. Chiaroni[1,2] M-C. Rahal[1] F. Dufaux[2] N. Hueber[3]

[1] Institut VEDECOM, équipe Perception du véhicule à conduite déléguée
[2] L2S, CNRS, CentraleSupelec, Univ Paris-Sud, Univ Paris-Saclay
[3] Institut Saint-Louis Franco-Allemand (ISL), équipe ELSI

florent.chiaroni@l2s.centralesupelec.fr

Abstract

In this article, we suggest a novel approach for image classification task from a positive unlabeled datased. Its proper functioning is based on generative adversarial networks (GANs) abilities. These allow us to generate fake images whose distribution is close to the distribution of the negative samples included in the unlabelled dataset available, while remaining different from the distribution of positive samples that are not labeled. Then we train a CNN classifier with the positive samples and the fake samples generated, as it would have been done with a classical Positive Negative dataset. Tests performed on three different image classification datasets show that the system is stable in its behavior with a non negligible fraction of positive samples present in the unlabeled dataset. Although very different, this method outperforms the state of the art in PU learning on the RGB CIFAR-10 dataset.

Résumé

Dans ce document, nous proposons une nouvelle approche répondant à la tâche de classification d'images à partir d'un apprentissage sur données positives et non-labélisées. Son bon fonctionnement repose sur certaines particularités des réseaux antagonistes génératifs (GANs). Ces derniers nous permettent de générer des fausses images dont la distribution se rapproche de la distribution des échantillons négatifs inclus dans le jeu de données non labélisé disponible, tout en restant différente de la distribution des échantillons positifs non labélisés. Ensuite, nous entraînons un classifieur convolutif avec les échantillons positifs et les faux échantillons générés, tel que cela aurait été fait avec un jeu de données classique de type Positif Négatif. Les tests réalisés sur trois jeux de données différents de classification d'images montrent que le système est stable dans son comportement jusqu'à une fraction conséquente d'échantillons positifs présents dans le jeu de données non labélisé. Bien que très différente, cette méthode surpasse l'état de l'art PU learning sur le jeu de données RVB CIFAR-10.

Mots Clef

Apprentissage Positif Non Labélisé (PU learning), Classification d'Images, Apprentissage Profond, Apprentissage de Représentations, Modèles Génératifs.

1 Introduction

Les méthodes d'apprentissage utilisant des filtres à noyaux de convolution ont démontré de bonnes performances de prédiction dans le domaine du traitement d'image, et plus particulièrement pour la tâche de classification d'images. Pour réaliser de telles performances, de grands jeux de données entièrement labélisés sont requis. De nos jours, plusieurs jeux de données distincts peuvent être amenés à être fusionnés pour cette raison afin d'augmenter la capacité de généralisation d'un modèle d'apprentissage tel que cela est proposé dans YOLO9000 [18]. Par ailleurs, pour atténuer ce besoin de grands jeux de données labélisés, des méthodes d'apprentissage semi-supervisé existent [16]. Mais, si un objet n'appartenant à aucune classe labélisée du jeu de données d'entraînement doit être traité, il reste difficile de prédire le comportement du modèle entraîné à son égard.

Néanmoins, une idée pouvant répondre à ce problème consiste à se focaliser principalement sur les données qui nous intéressent. Cela est le cas pour les méthodes de One-Class Classification (OCC) [8], détection de nouveauté [14] où il est utilisé uniquement des échantillons de la classe d'intérêt ; la classe positive. Cependant, à notre connaissance, les méthodes OCC ont une performance limitée lorsqu'elles sont appliquées à des tenseurs de données de grande dimensionalité tels que des images. De plus, il est souvent facile d'acquérir des échantillons non labélisés susceptibles de contenir des informations pertinentes à propos des contre-exemples de la classe d'intérêt. De cette manière, nous abordons le problème d'apprentissage Positif Non labélisé (apprentissage PU). Il se trouve que les méthodes d'apprentissage Positif Non labélisé ont été appliquées récemment à des données de type images tel que la méthode Rank Pruning (RP) [13]. Cette méthode est la plus performante de l'état de l'art dans un contexte où l'on n'a pas de connais-

 ©AfIA 2018

sances à priori sur les fractions d'échantillons bruités. Elle est cependant coûteuse en calculs car elle consiste à réaliser plusieurs entraînements consécutifs du même classifieur de manière à éliminer les échantillons les moins pertinents pendant la phase d'entraînement. De plus, selon [12], ces méthodes deviennent compétitives lorsque le nombre d'échantillons non labélisés dans le jeu de données d'entraînement augmente considérablement. Cela est un avantage lorsque l'on peut obtenir facilement des données non labélisées.
Par ailleurs, les réseaux génératifs antagonistes (GANs) ont attiré notre attention en raison de leur capacité à générer de faux échantillons x_F qui ont une distribution $p_G(x_F)$ qui tend vers la distribution $p_{data}(x_R)$ des échantillons réels x_R utilisés pendant son entraînement. Le GAN originel [5] contient un modèle génératif G et un modèle discriminatif D. Ces deux modèles possèdent une structure de type perceptron multi-couche. Un vecteur de bruit z, composé de variables aléatoires continues, est placé en entrée de G. D est entraîné à distinguer les échantillons réels des faux échantillons générés par G, pendant que ce dernier est entraîné à produire des faux échantillons qui doivent sembler réels au possible. Cet entraînement adversaire consiste à utiliser la fonction d'évaluation minimax $V(G, D)$:

$$\min_G \max_D V(G, D) = \mathbb{E}_{x_R \sim p_{data}(x_R)} log D(x_R) + \mathbb{E}_{z \sim p_z(z)} log[1 - D(G(z))].$$

Lorsque D ne peut plus distinguer les vrais échantillons des faux, nous obtenons la propriété suivante, avec y_D le scalaire de sortie prédit :

$$p_G(x_F) \xrightarrow[y_D \to \frac{1}{2}]{} p_{data}(x_R).$$

D'autres variantes du GAN sont apparues telles que le DCGAN [15], qui adapte sa structure au traitement d'images en intégrant des couches convolutives. Le Wasserstein GAN (WGAN) [1] utilise d'une part la distance $Earth - Mover$ (EM) dans sa fonction de coût, et d'autre part limite les valeurs des poids de son modèle à un certain intervalle, afin de rectifier le problème d'instabilité des précédentes versions du GAN.
En raison de leur capacité à apprendre des représentations pertinentes d'un point de vue sémantique et de leur efficacité déjà démontrée en apprentissage semi-supervisé [20], nous avons décidé d'exploiter d'une certaine manière leurs avantages pour une application d'apprentissage PU. Parallèlement à notre étude, l'approche [7] est apparue pour répondre à la même problématique en utilisant un modèle d'apprentissage de type GAN. [7] requière deux modèles génératifs et trois discriminateurs pour l'étape générative contre un générateur et un discriminateur pour notre approche. Cela devrait nous confèrer un temps de calcul moindre et un apprentissage mieux maîtrisé. Mais leur étude s'est arrêtée à la description fonctionnelle de leur modèle [1]. Ici, l'approche proposée que l'on nomme Positive-GAN ("PGAN" par la suite) a été testée sur trois jeux de données différents et dont les résultats sont très prometteurs en termes robustesse et de performance de prédiction pour le traitement d'images complexes. Il surpasse l'état de l'art sur le jeu de données le plus difficile que nous avons testé.
Le document est organisé tel que ci-dessous. Dans la section suivante nous présentons la méthode. Les expérimentations et les résultats sont présentés dans la troisième section. Pour finir, une conclusion est faite sur notre approche et de futures directions de recherche sont suggérées.

2 Méthode d'apprentissage proposée : Le Positive-GAN

Dans cette section, nous décrivons notre système d'apprentissage PU de manière générique et focalisons la description sur la méthode d'entraînement. Notre méthode d'apprentissage Positive-GAN (PGAN) consiste à substituer l'absence d'échantillons négatifs labélisés x_N avec les faux échantillons x_F générés par notre GAN, dont la distribution est proche au possible de celle des x_N, tout en étant différente de celle des échantillons positifs x_P. La figure 1 illustre le fonctionnement du système.

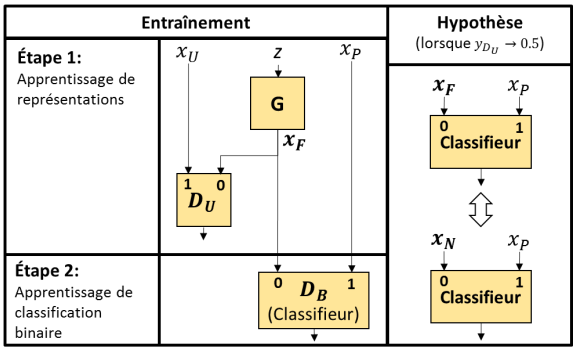


FIGURE 1 – Système d'apprentissage PU proposé : Positive-GAN.

Lors de l'étape 1, le GAN est entraîné avec les échantillons non labélisés x_U à partir du jeu de données d'entraînement

1. À noter qu'après la date de soumission de notre article à CNIA-RJCIA, une nouvelle version de la publication [7] est apparue, en version pre-print, où l'étude de leur méthode a été complétée : Leur nouvelle version datant du 4 avril 2018 inclue une partie théorique et quelques tests comparatifs sur les jeux de données USPS et MNIST auxquels il aurait été pertinent de se comparer, s'il n'y avait pas eu de conflits de dates. Leur nouvelle version nécessite la connaissance à priori de la fraction d'échantillons positifs inclus dans le jeu de données non labélisé pour fonctionner. Les tests ont été réalisés avec une faible proportion d'échantillons positifs labélisés, ce qui met en avant l'intérêt des méthodes génératives pour l'augmentation de dataset. Cependant, des détails sur l'intialisation des hyper-paramètres λ_P, λ_N et λ_U auraient été appréciés, ainsi que la réalisation de tests sur des bases de données plus complexes telles que CIFAR-10 avec des structures de type réseaux convolutifs. En effet, bien que leur méthode semble fonctionnelle avec une structure de type peceptron multicouche, les étapes de convolution rendent plus difficile l'effondrement d'un générateur et donc la divergence de leur générateur G_N pour l'apprentissage de la distribution des échantillons négatifs à partir d'échantillons non labélisés et d'échantillons positifs labélisés.

PU qui contient une fraction $\pi \in (0,1)$ d'échantillons positifs et une fraction $1-\pi$ d'échantillons négatifs x_N. Le système Positif-Non labélisé inclue trois modèles convolutifs avec différents rôles respectifs :

— Le modèle discriminateur D_U est entraîné à distinguer les vrais échantillons x_U des faux non labélisés générés x_F, avec $y_{D_U} \in (0,1)$ sa valeur de sortie prédite.

— Le modèle génératif G prend en entrée un vecteur de bruit z constitué de variables aléatoires continues, et fournit en sortie, dans le même format que x_U, les faux échantillons $x_F = G(z)$. G est entraîné de manière antagoniste à D_U afin de générer des faux échantillons tels que leur distribution $p(x_F)$ tend vers $p(x_U)$.

— Lors de la deuxième étape, une fois que l'entraînement du GAN est considéré comme terminé, le classifieur binaire convolutif D_B est entraîné à distinguer les échantillons réels positifs x_P des faux échantillons x_F.

Les explications présentées ci-dessous ont pour objectif de développer l'intuition derrière le système proposé.

Nous rappelons que le jeu de données non labélisé est composé d'une fraction π d'échantillons positifs x_P et d'une fraction $1-\pi$ d'échantillons négatifs x_N. Ainsi, si le GAN est correctement entraîné sur les échantillons non labélisés x_U, on peut en déduire que :

$$p(x_F) \xrightarrow[y_{D_U} \to \frac{1}{2}]{} p(x_U)$$

$$\Leftrightarrow p(x_F) \xrightarrow[y_{D_U} \to \frac{1}{2}]{} \pi\, p(x_P) + (1-\pi)\, p(x_N),$$

et l'on admet alors comme forte hypothèse pour la suite que $p(x_F) = \pi\, p(x_{FP}) + (1-\pi)\, p(x_{FN})$, avec :

$$\begin{cases} p(x_{FP}) \xrightarrow[y_{D_U} \to \frac{1}{2}]{} p(x_P) \\ p(x_{FN}) \xrightarrow[y_{D_U} \to \frac{1}{2}]{} p(x_N). \end{cases}$$

Lorsque $y_{D_U} \to \frac{1}{2}$, nous démarrons la deuxième étape du PGAN. Par ailleurs, un GAN n'est pas parfait dans son fonctionnement lorsqu'il se voit être appliqué à des tenseurs de grandes dimensions, ainsi :

$$p(x_{FP}) \neq p(x_P), \text{ et } p(x_{FN}) \neq p(x_N). \tag{1}$$

Il est donc alors possible d'estimer une distance d non nulle dans la fonction de coût du classifieur D_B, tel que :

$$d(p(x_P), p(x_F)) \;\Leftrightarrow \begin{cases} d(p(x_P), p(x_{FP})) \\ d(p(x_P), p(x_{FN})). \end{cases}$$

Mais, bien que calculée, la distance $d(p(x_P), p(x_{FP}))$ n'est pas exploitée dans l'application finale où nous traitons uniquement des échantillons réels avec le classifieur D_B. Ainsi, lorsque $p(x_{FN}) \xrightarrow[y_{D_U} \to \frac{1}{2}]{} p(x_N)$ et que D_B a été également correctement entraîné, nous obtenons l'équivalence :

$$d(p(x_P), p(x_{FN})) \Leftrightarrow d(p(x_P), p(x_N)). \tag{2}$$

Nous sommes donc capable de calculer la distance qui nous intéresse. En transférant ce raisonnement dans notre méthode PU, cela revient à affirmer les équivalences suivantes à la sortie de la fonction de coût L_{D_B} du classifieur D_B lorsque $y_{D_U} \to \frac{1}{2}$:

$$\begin{aligned} L_{D_B} &= \mathbb{E}_{x_P \sim p(x_P)} log D_B(x_P) \\ &+ \mathbb{E}_{z \sim p_z(z)} log[1 - D_B(G(z))] \end{aligned}$$

$$\begin{aligned} \Leftrightarrow L_{D_B} &= \mathbb{E}_{x_P \sim p(x_P)} log D_B(x_P) \\ &+ \mathbb{E}_{x_N \sim p(x_N)} log[1 - D_B(x_N)]. \end{aligned}$$

Ainsi, grâce à l'hypothèse de fonctionnement proposée ci-dessus on peut affirmer que la méthode PGAN devient similaire à un entraînement sur des échantillons positifs et négatifs respectivement labélisés, tout en s'éloignant d'un entraînement de type apprentissage PU, malgré le fait que le jeu de données que nous utilisons contienne uniquement des échantillons positifs labélisés et des échantillons non labélisés. De plus, intuitivement, calculer $d(p(x_P), p(x_{FP}))$ peut favoriser dans une certaine mesure, l'apprentissage des frontières de $p(x_P)$. Cependant, deux risques peuvent survenir avec cette méthode :

— Si les échantillons x_U contiennent majoritairement des échantillons x_P, alors il est possible que G ne soit plus apte à générer suffisamment de faux échantillons similaires aux échantillons x_N.

— Si G génère des faux échantillons ayant une distribution égale à celle des vrais échantillons, contrairement à l'inégalité 1, alors le PGAN deviendrait équivalent en termes de performances à un entraînement classique PU. Mais lorsque la dimensionalité des images à traiter devient large, alors ce risque disparaît. De plus, il peut être atténué en utilisant une structure pour le classifieur D_B dont les performances en prédiction sont meilleures que celles de la structure utilisée pour le discriminateur D_U.

Trouver une solution permettant d'éviter ces deux risques de se produire reste une question ouverte très intéressante pour améliorer la fiabilité de la méthode.

3 Expérimentations

3.1 Réglages des tests

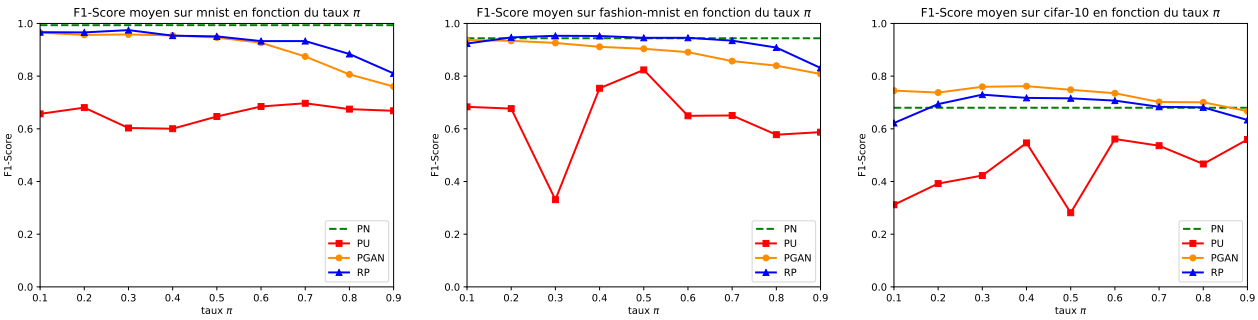
Les expériences ont été réalisées sur les trois jeux de données MNIST [10], Fashion-MNIST [21] et CIFAR-10 [9]. Nous avons comparé notre approche à RP [13], qui est à notre connaissance la meilleure méthode d'apprentissage bruité (noisy learning) et PU ne nécessitant pas de connaissances à priori de la fraction π. De plus, l'implémentation de l'auteur est disponible [2]. Nous indiquons aussi la performance du classifieur entraîné sur le jeu de données d'entraînement initial contenant des échantillons entièrement labélisés positifs et négatifs, et nous appelons évidemment cette méthode PN, que nous considérons comme la référence du cadre idéal. Nous comparons aussi le PGAN à un

2. *https : //github.com/cgnorthcutt/rankpruning*

©AfIA 2018

TABLE 1 – Résultats comparatifs en fonction des F1-Scores mesurés sur MNIST, Fashion-MNIST et CIFAR-10 avec le classifieur entraîné sur 20 époques.

	ref	$\rho = 0.5, \pi = 0.1$			$\rho = 0.5, \pi = 0.3$			$\rho = 0.5, \pi = 0.5$			$\rho = 0.5, \pi = 0.7$		
Jeux de données	PN	PU	PGAN	RP	PU	PGAN	RP	PU	PGAN	RP	PU	PGAN	RP
0	0.997	0.633	0.974	**0.992**	0.445	**0.973**	0.955	0.320	0.973	**0.991**	0.689	**0.902**	0.880
1	0.998	0.774	0.971	**0.995**	0.642	0.979	**0.996**	0.851	0.958	**0.994**	0.884	0.863	**0.993**
2	0.990	0.395	0.972	**0.975**	0.658	**0.959**	0.923	0.795	**0.947**	0.936	0.692	0.914	**0.987**
3	0.996	0.716	0.963	**0.991**	0.620	0.953	**0.991**	0.766	**0.934**	0.882	0.729	**0.885**	0.829
4	0.997	0.512	0.964	**0.972**	0.802	0.952	**0.995**	0.717	**0.945**	0.933	0.551	0.914	**0.977**
5	0.993	0.701	0.974	**0.985**	0.725	**0.950**	0.943	0.799	**0.949**	0.910	0.626	0.873	**0.973**
6	0.992	0.708	**0.962**	0.928	0.758	0.959	**0.992**	0.699	0.971	**0.993**	0.613	0.944	**0.990**
7	0.995	0.603	**0.962**	0.947	0.433	0.960	**0.991**	0.620	0.926	**0.988**	0.783	0.737	**0.979**
8	0.995	0.741	**0.949**	0.929	0.506	0.941	**0.982**	0.339	0.922	**0.941**	0.651	**0.849**	0.818
9	0.981	0.785	**0.959**	0.954	0.442	0.956	**0.979**	0.561	0.939	**0.941**	0.750	0.865	**0.904**
AVG_{MNIST}	0.993	0.657	0.965	**0.967**	0.603	0.958	**0.975**	0.647	0.946	**0.951**	0.697	0.875	**0.933**
T-shirt/top	0.908	0.724	**0.926**	0.899	0.206	0.91	**0.937**	0.821	0.873	**0.947**	0.695	0.802	**0.91**
Trouser	0.993	0.815	0.983	**0.993**	0.247	0.969	**0.989**	0.938	0.953	**0.99**	0.681	0.911	**0.984**
Pullover	0.932	0.635	**0.9**	0.887	0.29	0.885	**0.925**	0.695	0.865	**0.917**	0.657	0.842	**0.888**
Dress	0.952	0.601	0.941	**0.948**	0.312	0.925	**0.955**	0.852	0.893	**0.914**	0.631	0.853	**0.882**
Coat	0.882	0.614	**0.909**	0.847	0.252	0.889	**0.942**	0.788	0.845	**0.92**	0.686	0.83	**0.918**
Sandal	0.995	0.793	0.945	**0.977**	0.444	0.964	**0.98**	0.819	0.923	**0.985**	0.67	0.919	**0.981**
Shirt	0.818	0.446	**0.852**	0.758	0.398	0.846	**0.847**	0.797	0.819	**0.873**	0.554	0.792	**0.853**
Sneaker	0.983	0.73	**0.973**	**0.973**	0.271	0.952	**0.979**	0.865	0.943	**0.967**	0.64	0.922	**0.977**
Bag	0.989	0.772	**0.978**	0.976	0.536	0.947	**0.99**	0.837	0.96	**0.965**	0.685	0.757	**0.977**
Ankle boot	0.985	0.704	0.964	**0.979**	0.354	0.973	**0.986**	0.824	0.963	**0.976**	0.609	0.942	**0.976**
$AVG_{\text{F-MNIST}}$	0.944	0.683	**0.937**	0.924	0.331	0.926	**0.953**	0.824	0.904	**0.945**	0.651	0.857	**0.935**
Plane	0.727	0.341	**0.818**	0.669	0.557	0.784	**0.795**	0.295	**0.758**	0.743	0.621	**0.731**	0.718
Auto	0.78	0.506	**0.801**	0.695	0.492	0.737	**0.829**	0.414	0.789	**0.798**	0.521	0.734	**0.783**
Bird	0.447	0.175	**0.688**	0.56	0.439	**0.744**	0.68	0.184	**0.694**	0.644	0.359	**0.688**	0.542
Cat	0.5	0.125	**0.658**	0.384	0.272	**0.722**	0.651	0.249	**0.718**	0.67	0.446	0.69	**0.698**
Deer	0.698	0.272	**0.68**	0.605	0.232	**0.708**	**0.708**	0.3	**0.708**	0.64	0.43	**0.633**	0.602
Dog	0.567	0.2	**0.632**	0.539	0.37	**0.756**	0.648	0.258	**0.746**	0.733	0.514	0.678	**0.712**
Frog	0.691	0.35	**0.837**	0.666	0.418	0.793	**0.794**	0.256	**0.788**	0.769	0.693	**0.75**	0.749
Horse	0.786	0.373	**0.693**	0.653	0.515	**0.757**	0.723	0.26	0.751	**0.759**	0.611	0.675	**0.711**
Ship	0.832	0.313	**0.821**	0.764	0.565	0.809	**0.831**	0.324	0.775	**0.785**	0.623	0.716	**0.755**
Truck	0.771	0.462	**0.822**	0.685	0.367	**0.786**	0.637	0.272	**0.754**	0.617	0.539	**0.724**	0.564
$AVG_{\text{CIFAR-10}}$	0.680	0.312	**0.745**	0.622	0.423	**0.760**	0.730	0.281	**0.748**	0.716	0.536	**0.702**	0.684



(a) (b) (c)

FIGURE 2 – F1-Scores moyens après 20 époques d'entraînement pour le classifieur en fonction du taux π qui varie entre 0.1 et 0.9 avec un pas de 0.1, pour PN (vert), PU (rouge), RP (bleu) et PGAN (orange) sur MNIST (a), Fashion-MNIST (b) et CIFAR-10 (c).

entraînement que l'on nomme PU, qui est équivalent à PN, mais avec une substitution des échantillons négatifs par des échantillons non labélisés.
Pour ces expérimentations, les méthodes PN, PU, RP et PGAN sont testées avec exactement le même classifieur convolutif afin d'être impartial. Nous avons utilisé le modèle convolutif de classification d'imagettes proposé par tensorflow [3] pour rester générique. Il contient deux couches convolutives successives, respectivement suivies d'une étape de max-pooling, et se finit par deux couches entièrement connectées consécutives. La fonction d'activation en sortie de chaque couche est ReLU, mis à part pour la dernière où softmax est appliquée. Nous avons uniquement modifié la dimension de sortie de la dernière couche que l'on fait passer de 10 neurones à 2, afin d'être adapté à notre tâche de classification binaire. Le classifieur est entraîné sur 20 époques. Pour les images 32x32x3 de CIFAR-10, les largeurs et longueurs des tenseurs d'entrée et de sortie des deux couches convolutives sont adaptées, et la profondeur des filtres à noyaux de la première couche convolutive est établie à 3 afin de correspondre aux trois cannaux de ces images RVB. Mais le nombre de filtres et leurs largeur et longueur restent inchangés.
Pour l'étape générative du PGAN, nous avons associé la méthode d'entraînement du WGAN [1] à l'architecture du DCGAN [15] en raison de leurs performances. À noter qu'avec la distance EM, $p(x_F)$ tend vers $p(x_U)$ lorsque y_{D_U} tend vers 0. Bien que cela ne soit pas une nécessité, dans le cadre de ces expériences, le vecteur de bruit d'entrée z est constitué de variables aléatoires continues de distributions uniformes. La durée d'entraînement de notre modèle génératif dépend de la complexité du jeu de données à traiter : 10 époques pour MNIST, 20 pour Fashion-MNIST, et 100 pour CIFAR-10. Pour ce dernier, nous faisons les mêmes modifications dans la structure de D_U et G, tel que cela a été expliqué précédemment pour le classifieur.
Au sujet de la création de notre jeu de données d'entraînement, ρ correspond à la fraction d'échantillons positifs du jeu de données initial qui contient n_P échantillons positifs. Ces $\rho \times n_P$ échantillons collectés sont ensuite introduits dans notre jeu de données non-labellisé U_{train}, qui contient initiallement uniquement des échantillons négatifs N dont le nombre total est n_N. π est la fraction d'échantillons positifs P que l'on impose dans le jeu de données non labellisé d'entraînement U_{train}. Pour se faire, nous retirons de U_{train} un certain nombre d'échantillons négatifs que nous n'exploitons pas, de manière à respecter π. U_{train} contient alors à la fois des N et des P selon les paramètres ρ et π. Nous établissons qu'avec $\pi \in [\frac{1}{\frac{n_N}{\rho\ n_P}+1}, 1)$ et $\rho \in (0, 1)$, nous pouvons alors obtenir consécutivement, avec P_{train} l'ensemble d'échantillons positifs d'entraînement, les deux jeux de données d'entraînement suivants :

$$P_{train} = \{(1-\rho)\, n_P\ P\ ;\ 0\ N\},$$

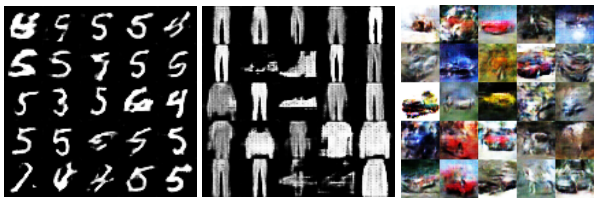
$$U_{train} = \{\rho\, n_P\ P\ ;\ \frac{1-\pi}{\pi}\ \rho\, n_P\ N\},$$

où les notations $a\ P$ et $b\ N$ désignent respectivement a éléments positifs et b éléments négatifs.
Pour trouver les équations définissant U_{train} selon les paramètres π et ρ, et l'intervalle des valeurs possibles pour π, nous utilisons n_U qui représente le nombre total d'échantillons non labélisés contenus dans l'ensemble U_{train}, tel que :

$$\begin{aligned} U_{train} &= \{\pi\, n_U\ P\ ;\ (1-\pi) n_U\ N\} \\ &= \{\rho\, n_P\ P\ ;\ (1-\pi)\frac{\rho\, n_P}{\pi}\ N\}, \end{aligned}$$

car nous imposons $\rho\, n_P = \pi\, n_U$.
Or, pour que cela soit réalisable, il faut que $(1-\pi)\frac{\rho\, n_P}{\pi}$ soit inférieur ou égal à n_N. Cela revient donc à dire que $\pi \in [\frac{1}{\frac{n_N}{\rho\ n_P}+1}, 1)$.
Les résultats présentés ci-dessous sont tous réalisés avec $\rho = 0.5$ et pour plusieurs valeurs de π.

3.2 Résultats



(a) (b) (c)

FIGURE 3 – Images générées par G avec $\rho = 0.5$ et $\pi = 0.5$ après 10 époques sur MNIST (a), 20 sur Fashion-MNIST (b), et 100 sur CIFAR-10 (c). Les classes positives respectives sont ici "5", "$trouser$" and "$automobile$".

Dans la figure 3, nous présentons quelques fausses images générées par G, respectivement pour MNIST, Fashion-MNIST et CIFAR-10. Nous pouvons remarquer que les images générées par G semblent visuellement acceptables, ce qui indique d'un point de vue qualitatif le bon fonctionnement du modèle génératif. Afin d'obtenir un tel résultat, plus les images sont grandes et complexes, et plus il faut entraîner le GAN sur un grand nombre d'époques.
Pour calculer le F1-Score, la fonction ArgMax est appliquée aux deux neurones de sortie du classifieur. Si l'indice du premier neurone est renvoyé par ArgMax, alors l'échantillon traité est classifié comme négatif. Sinon, il est considéré comme positif. De plus, étant donné que les jeux de données de test contiennent 9 fois plus d'échantillons négatifs que d'échantillons positifs, une fois toutes les prédictions de test réalisées, nous adaptons les proportions des échantillons négatifs à celle des positifs de manière à obtenir un F1-Score pertinent. Le tableau 1 montre une partie des F1-Scores comparatifs mesurés pour chaque classe pour chacun des trois jeux de données exploités, et respectivement pour chaque méthode testée. Sur Fig. 2, il peut être

3. $https://github.com/tensorflow/tensorflow/blob/master/tensorflow/examples/tutorials/mnist/mnist_softmax.py$

©AfIA 2018

observé que la méthode PN est une bonne référence sur MNIST et Fashion-MNIST. Nous trouvons que l'efficacité de la méthode d'apprentissage PGAN est équivalente à celle de la méthode RP jusqu'à $\pi = 0.5$ sur MNIST et $\pi = 0.3$ sur Fashion-MNIST. Son efficacité décline ensuite un peu plus vite que pour RP, mais tout en conservant un score acceptable. Sur CIFAR-10 le F1-Score moyen est systématiquement meilleur pour notre méthode PGAN. Aussi, notre méthode présente de meilleurs résultats que la référence PN jusqu'à $\pi = 0.8$, ce qui est très intéressant. Cela est probablement dû au fait que les échantillons générés représentent une plus grande variété de distributions d'échantillons négatifs que celles inclues dans le jeu de données initial. De plus, le F1-Score du PGAN est significativement et systématiquement meilleur que la méthode PU sur l'ensemble des trois jeux de données, même avec seulement 10% d'échantillons positifs parmis les échantillons non labélisés, autrement dit avec $\pi = 0.1$.

La figure 4 présente l'étude de la robustesse de l'approche PGAN. Les figures 4.a et 4.b montrent que la méthode PGAN a comparativement à RP une meilleure stabilité dans son fonctionnement de manière à permettre de prédire plus facilement l'évolution de son F1-Score en fonction de π pour chaque classe du jeu de données. La figure 4.c nous montre que le classifieur se stabilise et converge après 10 époques d'entraînement. Pour réaliser l'histogramme de la figure. 4.d, nous avons récupéré la valeur du deuxième neurone de sortie du classifieur qui correspond à la probabilité prédite pour une image d'appartenir à la classe positive. On peut observer que les distributions respectives des échantillons de test positifs et négatifs estimées par le PGAN sont de forme gaussienne, ce qui est une caractéristique intéressante pour des applications réelles.

TABLE 2 – Stabilité moyenne des performances (F1-Scores) des méthodes PGAN et RP en fonction de π sur les jeux de données MNIST, Fashion-MNIST et CIFAR-10

jeux de données	PGAN	RP	$\frac{E_{RP}}{E_{PGAN}}$
MNIST	**0.00039**	0.00172	4.410
Fashion-MNIST	**0.00016**	0.00025	1.563
CIFAR-10	**0.00044**	0.00182	4.136

En complément aux figures 4.a et 4.b, le tableau 2 présente la quantification des robustesses réalisée pour RP et PGAN en fonction de π en ce qui concerne leurs performances de prédiction, pour chacun des trois jeux de données de test. Pour ce faire, nous avons lissé respectivement la courbe $s(\pi)$ de chaque classe représentant l'évolution du F1-Score en fonction de π. Les courbes lissées $\tilde{s}(\pi)$ ont été obtenues en appliquant un filtre moyen avec un noyau de taille 3. Ensuite, l'erreur quadratique moyenne MSE pour chaque classe est calculée entre $\tilde{s}(\pi)$ et $s(\pi)$ tel que ci-dessous, avec k le nombre d'échantillons de $\tilde{s}(\pi)$:

$$MSE = \frac{1}{k}\sum_{i=1}^{k}(\tilde{s}^{(i)} - s^{(i+1)})^2. \quad (3)$$

Puis, nous calculons les erreurs moyennes E_{PGAN} et E_{RP} pour chaque jeu de données, ainsi que le ratio E_{RP} : E_{PGAN}. On peut de cette manière constater que notre méthode a systématiquement un comportement plus stable, qui est d'un facteur 4 sur MNIST et CIFAR-10.

4 Conclusion

Ainsi, nous avons démontré que l'approche d'apprentissage PU proposée surpasse l'état de l'art sur les images RVB complexes du jeu de données CIFAR-10, et a un comportement plus stable sur l'ensemble des jeux de données testés jusqu'à une fraction acceptable π d'échantillons positifs dans le jeu de données non labélisé d'entraînement. Ces résultats sont en cohérence avec le raisonnement formulé et permettent ainsi d'envisager des applications PU sur des données de plus grandes dimensions. Le PGAN ne nécessite pas de connaissances à priori sur la fraction d'échantillons positifs non labélisés. Cependant, il reste à étudier plus en profondeur les risques de fonctionnement indiqués dans la section 1, afin de garantir un fonctionnement idéal pour cette approche.

L'optimisation du système peut se prolonger en testant d'autres récentes variantes du GAN tels que le BEGAN [3], le WGAN-GP [6], ou bien d'autres modèles génératifs de type auto-encodeurs variationnels (VAEs) par exemple, afin de généraliser l'approche proposée aux réseaux génératifs. Une autre idée peut être d'exploiter le vecteur latent z du GAN pour réaliser des opérations arithmétiques linéaires, tel que dans [4], afin de générer des faux échantillons dont on pourrait peut-être ainsi mieux gérer la distribution. Dans cette même idée, trouver un moyen d'exploiter les données positive labélisées pour la phase d'entraînement du générateur est envisagé.

Étant données les performances prometteuses obtenues, une future orientation certaine est d'étendre cette méthode à l'analyse de plus grandes images et donc permettre la réalisation de tâches plus complexes telles que la détection d'objets [19], [11], [17] ou la segmentation sémantique [2].

Références

[1] M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein generative adversarial networks. In *International Conference on Machine Learning*, pages 214–223, 2017.

[2] V. Badrinarayanan, A. Kendall, and R. Cipolla. Segnet : A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12) :2481–2495, 2017.

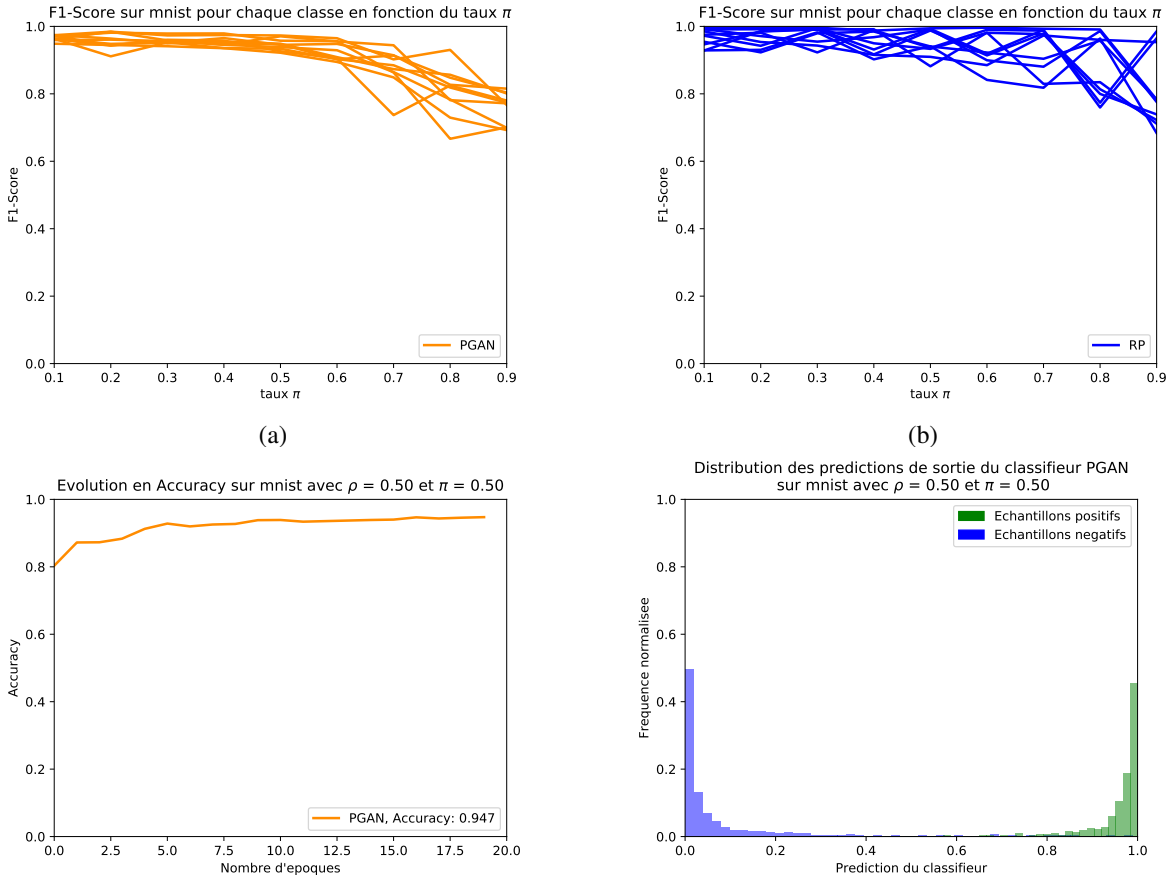


FIGURE 4 – Analyse de la robustesse sur MNIST. évolution du F1-Score pour chaque classe en fonction de π pour PGAN (a), et pour RP [13] (b). (c) montre l'évolution de l'Accuracy pendant l'entraînement du PGAN avec la classe positive "5" et $\pi = 0.5$. (d) est l'histogramme des distributions des valeurs de sortie du deuxième neurone du classifieur à son 20ème époque d'entraînement de (c) pour les échantillons de test positifs (vert) et négatifs (bleu).

[3] D. Berthelot, T. Schumm, and L. Metz. Began : Boundary equilibrium generative adversarial networks. *arXiv preprint arXiv :1703.10717*, 2017.

[4] P. Bojanowski, A. Joulin, D. Lopez-Paz, and A. Szlam. Optimizing the latent space of generative networks. *arXiv preprint arXiv :1707.05776*, 2017.

[5] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

[6] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville. Improved training of wasserstein gans. In *Advances in Neural Information Processing Systems*, pages 5769–5779, 2017.

[7] M. Hou, Q. Zhao, C. Li, and B. Chaib-draa. A generative adversarial framework for positive-unlabeled classification. *arXiv preprint arXiv :1711.08054*, 2017.

[8] S. S. Khan and M. G. Madden. One-class classification : taxonomy of study and review of techniques. *The Knowledge Engineering Review*, 29(3) :345–374, 2014.

[9] A. Krizhevsky and G. Hinton. Learning multiple layers of features from tiny images. 2009.

[10] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11) :2278–2324, 1998.

[11] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg. SSD : Single shot multibox detector. In *European Conference on Computer Vision*, pages 21–37. Springer, 2016.

[12] G. Niu, M. C. du Plessis, T. Sakai, Y. Ma, and M. Sugiyama. Theoretical comparisons of positive-unlabeled learning against positive-negative learning. In *Advances in Neural Information Processing Systems*, pages 1199–1207, 2016.

©AfIA 2018

[13] C. G. Northcutt, T. Wu, and I. L. Chuang. Learning with confident examples : Rank pruning for robust classification with noisy labels. In *Proceedings of the Thirty-Third Conference on Uncertainty in Artificial Intelligence*, UAI'17. AUAI Press, 2017.

[14] M. A. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko. A review of novelty detection. *Signal Processing*, 99 :215–249, 2014.

[15] A. Radford, L. Metz, and S. Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv :1511.06434*, 2015.

[16] A. Rasmus, M. Berglund, M. Honkala, H. Valpola, and T. Raiko. Semi-supervised learning with ladder networks. In *Advances in Neural Information Processing Systems*, pages 3546–3554, 2015.

[17] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once : Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.

[18] J. Redmon and A. Farhadi. YOLO9000 : Better, Faster, Stronger. *arXiv preprint arXiv :1612.08242*, 2016.

[19] S. Ren, K. He, R. Girshick, and J. Sun. Faster R-CNN : Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, pages 91–99, 2015.

[20] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen. Improved techniques for training gans. In *Advances in Neural Information Processing Systems*, pages 2234–2242, 2016.

[21] H. Xiao, K. Rasul, and R. Vollgraf. Fashion-mnist : a novel image dataset for benchmarking machine learning algorithms, 2017.

Ontologie pour l'intégration de données d'observation de la Terre et contextuelles basée sur les relations topologiques

C. Comparot[1] C. Trojahn[1] H. Arenas[2] N. Aussenac-Gilles[2]

[1] IRIT, CNRS et Université Toulouse 2, France
[2] IRIT, CNRS et Université de Toulouse, France

prenom.nom@irit.fr

Résumé

Nous proposons une ontologie pour l'intégration de données d'observation de la Terre et de données contextuelles, à l'aide de propriétés et de relations spatiales et temporelles. Cette ontologie spécialise les standards SOSA, GeoSPARQL et OWL-Time. La base de connaissances décrite par cette ontologie est alimentée grâce à un processus qui sélectionne, transforme et intègre de données géospatiales hétérogènes (méta-données d'image satellite, données météorologiques, unités administratives, couverture terrestre, etc.). Ce processus s'appuie sur le tuilage des images pour traiter les données ayant une composante spatiale fixe, alors que les relations temporelles sont calculées à la volée à partir d'une topologie temporelle. Nous validons cette approche par un cas d'étude exploitant des méta-données d'image satellite Sentinel.

Mots Clef

Ontologies, intégration de données, données spatiales et temporelles, SOSA.

Abstract

We propose an ontology to support the integration of Earth observation data with contextual data, using spatial and temporal properties and relations. This ontology specializes several standard models : SOSA, GeoSPARQL and OWL-Time. The process that feeds the Knowledge Base structured by this ontology selects, transforms and integrates heterogeneous data with spatial features i.e. metadata of satellite images, weather report data, administrative units and land cover. This process relies on image tiling to manage data with a fixed spatial property whereas temporal relations are evaluated when needed. We validate our approach on a case-study using the metadata of Sentinel satellite images.

Keywords

Ontologies, data integration, spatial and temporal data, SOSA.

1 Introduction

L'observation de la Terre offre une valeur ajoutée à une grande diversité de domaines. Récemment l'Agence Spatiale Européenne (ESA) a lancé le programme Sentinel avec deux types de satellites, Sentinel-1 and Sentinel-2, qui transmettent des images de haute qualité (entre 8 à 10 To de données quotidiennement). Ces images de la Terre sont captées selon différentes technologies et libres d'accès. Cette disponibilité des données ouvre de nombreuses perspectives économiques grâce à de nouvelles applications dans des domaines aussi variés que l'agriculture, l'environnement, l'urbanisme, l'océanographie ou encore la climatologie. Ces applications métier ont néanmoins un besoin de coupler les images avec des données sur les zones observées. Ces données sont accessibles à partir de différentes sources dans des formats hétérogènes et des temporalités différentes : elles peuvent être statiques, comme les données sur la couverture terrestre, ou dynamiques, comme les observations météo. Elles peuvent être utiles par exemple pour indiquer qu'une image contient une région touchée par un phénomène tel qu'un tremblement de terre ou une canicule, et ensuite pour décider des actions à mener dans cette zone ou conduire à des analyses à plus long terme. Plus encore, en exploitant les caractéristiques spatio-temporelles d'un phénomène (son empreinte spatiale et sa date), il devient possible de savoir si une entité localisée dans l'empreinte de l'image (e.g. une ville) a subi le même phénomène.

Les images satellites étant décrites par des méta-données, le problème revient à intégrer à ces méta-données des données provenant de sources variées et très hétérogènes (format, représentation). L'apport des technologies sémantiques pour faciliter cette tâche a été démontré dans des travaux antérieurs [25, 26], en particulier grâce à l'utilisation d'ontologies comme représentations formelles des connaissances d'un domaine donné. Ainsi, dans la continuité des travaux sur l'accès et l'intégration de données via les ontologies (i.e.OBDA pour "Ontology-based Data Access" et OBDI pour "Ontology-based Data Integration") [21, 20, 8], nous avons conçu un vocabulaire sémantique permettant de représenter les données des diffé-

 ©AfIA 2018

rentes sources envisagées et d'y accéder de façon homogène. Notre approche a consisté à construire une ontologie modulaire permettant de répondre aux besoins d'exploitation propres à chaque source de données. Cette approche permet de réduire partiellement la complexité et l'hétérogénéité des données, et ainsi de faciliter le peuplement de l'ontologie. Une caractéristique commune aux observations de la Terre est qu'elles peuvent être liées via leurs propriétés topologiques spatiales et temporelles. Un préalable au peuplement de l'ontologie (i.e. la production d'entrepôts RDF), a été de convertir les différents ensembles de données géo-spatiales fournis dans des formats hétérogènes (shapefile, KML, CSV, GeoJSON, TIFF), en un format commun, JSON, voire d'assurer, le cas échéant, la compatibilité de leurs propriétés et relations spatiales et temporelles, puis de les stocker dans une base de données NoSQL MongoDB.

Nous présentons dans cet article le vocabulaire que nous avons défini pour assurer la description sémantique homogène des différents types de données sous forme d'entités ayant des propriétés spatiales et temporelles. Une partie des données géospatiales à intégrer aux méta-données d'image sont des données contextuelles mesurées à la surface de la Terre, que nous traitons comme des données de capteurs. Ce vocabulaire spécialise ainsi des vocabulaires connus du web des données (LOD pour Linked Open Data), dont SOSA [1] pour les données de capteurs, GeoSPARQL [17] pour gérer les relations et coordonnées spatiales, et OWL Time [2] pour traiter la dimension temporelle. De fait, nous nous sommes appuyés sur une topologie des entités, en distinguant notamment les données dynamiques, dont la validité est fournie par la composante temporelle, et les données statiques, pour lesquelles la composante spatiale joue un rôle plus primordial.

Nous montrons, à travers un cas d'étude, comment cette ontologie permet de prendre en compte les spécificités de différentes sources en termes de vocabulaire, périodicité et volume, et comment elle peut améliorer l'accès aux données. A titre d'exemple, nous avons généré des entrepôts RDF qui mettent en relation des informations sur la couverture terrestre (pourcentage de forêts, d'eau, etc.) d'une part, et des relevés de stations météorologiques (température, humidité, etc.) d'autre part, avec les images satellites. Ce travail est mené dans le cadre du projet SparkinData visant à construire une plate-forme cloud de données d'observations de la Terre. Le cas d'étude s'appuie sur l'annotation sémantique de méta-données d'images brutes fournies par le CNES (Centre National d'Etudes Spatiales).

Le reste de l'article est organisé ainsi : la partie 2 expose des travaux liés ; en partie 3, nous présentons le modèle d'intégration ; son exploitation pour représenter en RDF les diverses sources d'informations envisagées est l'objet de la partie 4 ; la partie 5 décrit différentes stratégies mises en oeuvre pour intégrer les données des entrepôts RDF via les relations topologiques. Enfin, nous concluons et présentons des perspectives à ce travail en partie 6.

2 Travaux liés

2.1 Modèle sémantique pour l'imagerie satellitaire

Le projet européen TELIOS (Virtual Observatory Infrastructure for Earth Observation Data) [3] a été pionnier dans l'utilisation de représentations sémantiques pour décrire des images de satellites et faciliter l'accès à des données d'observation de la Terre par des agences telles que la NASA ou l'ESA. Le cadre applicatif de ce projet est la gestion des incendies en Grèce [29]. Les images sont classifiées et représentées sous forme vectorielle, et à chaque vecteur est associée une estimation des risques d'incendie. Les vecteurs sont liés à des sources de données externes par le biais d'ontologies [28]. Ces ontologies exploitent notamment le standard stRDF qui étend RDF pour les données spatiales et temporelles ; seule la composante spatiale est exploitée dans les analyses réalisées.

Le travail de [10] et [11] a été influencé par TELEIOS. Leur système permet d'ajouter des annotations sémantiques à des images produites par des radars à synthèse d'ouverture (SAR) selon le processus suivant : 1) extraction des informations pertinentes à partir des fichiers de métadonnées des images (fichiers XML) ; 2) division de l'image en *tuiles* de tailles différentes et organisation de ces tuiles sous une forme pyramidale ; 3) extraction puis transformation sous forme de vecteurs des caractéristiques de chaque tuile. Ce système est capable de trouver de l'information en utilisant des requêtes portant sur les méta-données, la sémantique et le contenu spatial des images. Cette approche associe aux images d'observation de la Terre, des données provenant des sources externes comme Corine Land Cover, Urban Atlas, Geonames et LinkedGeoData. Elle s'appuie sur un modèle ontologique, SAR, qui réutilise la taxonomie TerraSARX et étend celle de [28].

L'approche de [16] utilise à la fois des sources de données externes et des relations temporelles. Elle s'appuie sur le concept de *prisme spatio-temporel* ("space-time prism") qui décrit les caractéristiques de trajectoires. Un outil très connu pour l'analyse spatiale est le *buffer* ; il s'agit d'une zone située autour d'une région d'intérêt, mesurée dans une certaine unité spatiale. Le *prisme spatio-temporel* est similaire à un *buffer* auquel on a ajouté une dimension temporelle. Une fois la zone autour de la caractéristique d'intérêt définie, il est possible d'identifier d'autres propriétés contenues dans le *prisme spatio-temporel*. Ainsi, dans les trajectoires étudiées par [16], certains points sont caractérisés par des valeurs temporelles de type `xsd:dateTime`. Les valeurs extraites déterminent si un élément est contenu ou non dans un *prisme spatio-temporel* donné.

1. `https://www.w3.org/2015/spatial/wiki/SOSA\_Ontology`
2. `https://www.w3.org/TR/owl-time/(10/2017)`
3. http ://www.earthobservatory.eu (08/2018)

Compte-tenu du volume d'information à gérer, la recommandation RDF DataCube du W3C [7] suggère de lier les images à des tuiles de telle sorte à faire des assertions sur les tuiles. Dans ce cadre, les tuiles sont des zones carrées géo-localisées définies suite à une décomposition de la surface de la terre à partir d'une grille. Chaque image fournie par Sentinel-2 Single Tile (S2ST) a déjà une tuile. Nous nous appuyons également sur la notion de tuile dans notre processus d'intégration.

2.2 Données géographiques et LOD

La collecte et l'intégration de données géographiques produites par une diversité de disciplines est au coeur du projet Digital Earth [13]. Rendre les données géographiques disponibles puis interopérables à un niveau sémantique est une préoccupation qui conduit à appliquer les principes du Linked Data [6], et ainsi exposer, partager, et intégrer les données sur le Web via des URI déréférençables [15]. Les recommandations du W3C pour publier des données spatiales sous forme de données liées du LOD attirent l'attention sur la représentation des relations spatiales et des systèmes de référence (CRS pour "Coordinate Reference Systems") [27]. Ainsi, de nombreuses ontologies et vocabulaires sont proposés pour représenter des données géospatiales volumineuses au format *raster*. Le W3C suggère l'ontologie RDF Data Cube (QB), combinée à d'autres ontologies standards du W3C et de l'OGC dont SSN (Semantic Sensor Network) [4], OWL-Time [5], SKOS [6], PROV-O [7] et la récente extension de DataCube pour les entités spatio-temporelles, QB4ST [8].

L'OGC a introduit la notion de *données géo-liées* ("geolinked data") pour faire référence aux données liées géographiquement. Dans les premiers travaux, la géométrie était stockée dans un ensemble de données géospatiales séparé, et non directement comme valeur d'attributs. Cette option étant trop contraignante lorsqu'il faut comparer la géométrie de chaque entité, les entrepôts actuels mémorisent ensemble, une représentation RDF de la géométrie et des entités spatiales. [3] a identifié divers types de géométries (point, ligne ou polygone) et divers outils pour construire une représentation RDF de la géométrie (comme Geometry2RDF [9] ou TripleGeo [10]). Il modélise aussi quatre vocabulaires pour représenter les CRS, les entités topographiques et leurs géométries. Ces ontologies étendent des vocabulaires existant et offrent deux avantages supplémentaires : l'utilisation explicite du CRS identifié par des URI pour la géométrie, et la possibilité de décrire des géométries structurées en RDF. Ces données sont publiées comme la base de données française GEOFLA.

Une autre considération à intégrer pour représenter des données d'observations de la Terre est la différence de validité temporelle des données. Certaines données, comme la position des stations météorologiques, des villes et de la plupart des lieux administratifs, et même la couverture terrestre, sont valides pour une très longue période, plus longue que celle de l'application, et peuvent être considérées stables ou statiques. Les flux de données, au contraire, fournissent en continu de nouvelles données à intervalles de temps réguliers. Par exemple, les mesures de température sont données toutes les 3 heures par les bulletins Météo France, et des dizaines d'images satellites et leurs métadonnées sont disponibles sur le serveur PEPS chaque jour.

3 Le modèle d'intégration

3.1 Rôle de la géolocalisation et de la date

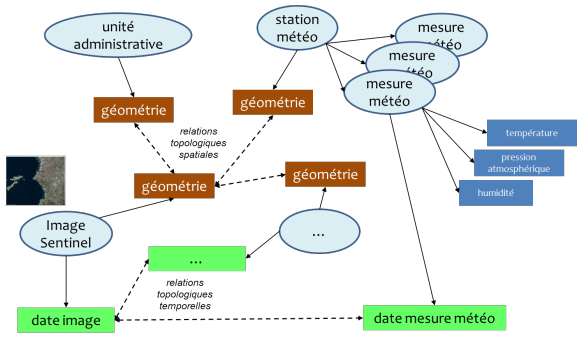


FIGURE *1 – Illustration du rôle pivot de la géométrie et de la date pour associer les données par des relations spatiales et temporelles.*

Le modèle retenu pour intégrer des données aux métadonnées des images repose sur l'hypothèse que ces données sont géolocalisées et datées. Pour chacune, on doit connaître a minima un point déterminé par sa latitude / longitude, ou une zone géolocalisée (appelée « géometrie » dans la Figure 1). C'est par le biais de la géométrie que des données peuvent être associées à une même image ou à une zone dans une image. Par exemple sur la Figure 1, les mesures météorologiques (humidité, température, pression) ainsi que les unités administratives (villes, régions, etc.) sont géolocalisées et couvrent une superficie plus ou moins grande. Pour comparer les géométries des données à lier, il faut les ramener à une même unité, ce qui permet de savoir quelle zone de l'image est concernée ou décrite par ces données. Les images étant datées, dès lors que les autres données sont aussi datées, comme les mesures météorologiques, il est possible de les lier par des relations temporelles, et de savoir par exemple quelles sont les températures relevées sur la zone couverte par l'image durant les 3 jours qui ont précédé la prise d'image.

Nous détaillons à présent le modèle conçu comme support à cette approche. Pour permettre l'intégration de divers types de données, ce modèle, présenté sur la Figure 2 comporte une partie générique (les trois cadres *time*, *sosa*

4. http://purl.oclc.org/NET/ssnx/ssn
5. https://www.w3.org/TR/owl-time
6. http://www.w3.org/2004/02/skos/core
7. https://www.w3.org/TR/prov-o
8. https://www.w3.org/TR/qb4st/
9. https://github.com/boricles/geometry2rdf
10. https://github.com/GeoKnow/TripleGeo

©AfIA 2018

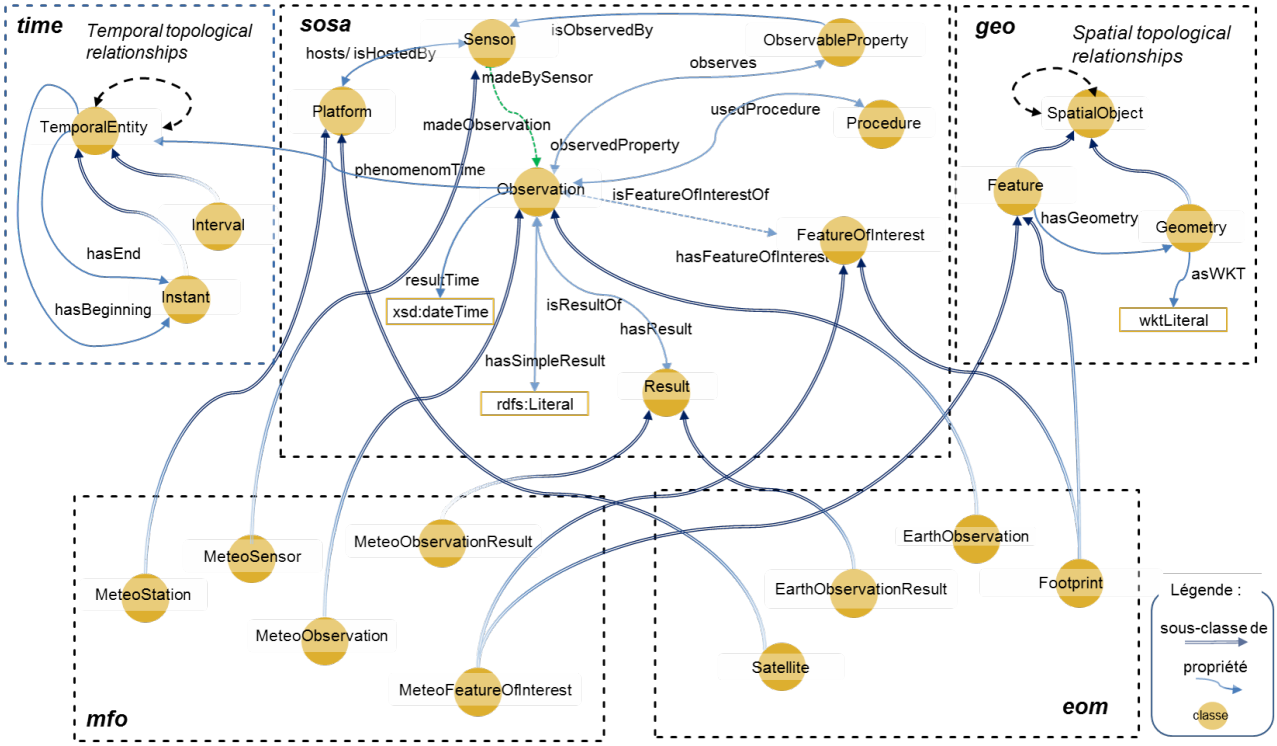


FIGURE 2 – *Le modèle d'intégration de données.*

et *geo* de la partie haute de la figure), composée de classes et de propriétés tirés de vocabulaires existants (respectivement OWL-Time, SOSA, GeoSPARQL), et une partie spécifique aux données à intégrer (cadres de la partie basse de la Figure 2). La partie spécifique comprend a minima un vocabulaire pour décrire les métadonnées d'images (cadre *eom* pour "Earth Observation Model") et autant de vocabulaires que de types de données à intégrer (un seul vocabulaire est mentionné ici, *mfo* "pour MétéoFrance Observation", représentant les données météorologiques).

3.2 Les vocabulaires réutilisés

Pour faciliter le partage de connaissances, une bonne pratique dans la construction d'ontologies consiste à identifier des ontologies existantes à réutiliser, à les connecter ou les étendre. Nous commençons donc par décrire les vocabulaires sur lesquels s'appuie la partie générique de notre modèle : OWL-Time, SOSA et GeoSPARQL, désignés par la suite respectivement par les préfixes `time`, `sosa` et `geo`. D'autres vocabulaires peuvent être réutilisés en fonction des données à intégrer pour la partie spécifique.

Vocabulaire pour représenter méta-données d'images et données de capteurs. Un grand nombre de modèles ont été développés pour représenter des méta-données d'observations par les communautés des sciences de la Terre et de l'environnement. O&M (Observations and Measurements), développé par le groupe SWE de l'OGC, fournit une vision utilisateur des observations. L'OGC Sensor Observation Service and Web Feature Service en a fourni une implémentation XML (spécification XSD). Plus général, DCAT est un vocabulaire recommandé par le W3C pour publier des métadonnées de catalogues sur le Web et disponible au format RDF [22].

Pour faciliter l'intégration de données provenant de différents types de capteurs, l'ontologie SSN [11] décrit des réseaux de capteurs numériques ainsi que les fonctionnalités, propriétés et mesures de ces capteurs. Construite comme une sous-couche de l'ontologie formelle DOLCE-UltraLite (DUL), SSN est générique au sens où elle ne définit pas toutes les propriétés spécifiques à des capteurs et des observations particuliers. En revanche, on peut la lier à des ontologies ou des vocabulaires propres à un domaine tels que l'ontologie SWEET (Semantic Web for Earth and Environmental Terminology) qui décrit des données environnementales [24]. Pour faciliter sa réutilisation, et éviter les risques d'incompatibilité notamment avec les déclarations DUL, une version modulaire de SSN a été publiée. Elle est composée de plusieurs ontologies couvrant le même domaine, mais avec une portée sémantique différente pour s'adapter à des contextes d'utilisation différents. SOSA (Sensor, Observation, Sample, and Actuator) est le module ontologique noyau de la nouvelle version de SSN ; à ce titre, il regroupe les termes centraux de l'ontologie. [9] proposent des alignements entre différents modèles de données d'observations, dont SOSA, à l'aide de l'ontologie PROV-O, pour faciliter l'intégration des données produites

11. `https://www.w3.org/TR/vocab-ssn/`

avec chacun d'eux. Les axiomes d'alignement sont fournis en RDF dans un module de la spécification SOSA/SSN. Ils s'appuient notamment sur le fait qu'une observation dans le vocabulaire SOSA (`sosa:Observation`) est considérée comme une activité (`prov:Activity`) de capteur, i.e. tout processus (calcul, simulation, interprétation, etc.) appliquant une procédure pour obtenir une estimation d'une valeur de propriété. La classe `sosa:Observation` y est ainsi définie comme une sous-classe de `prov:Activity`.

Nous avions dans un premier temps fait le choix de DCAT et SSN pour représenter respectivement les catalogues de méta-données d'images et les données issues de stations météorologiques [1, 2]. SOSA étant pertinent pour une vaste gamme d'applications, dont l'imagerie satellite, nous l'avons finalement retenu pour représenter à la fois les méta-données d'image et les données météorologiques.

Vocabulaire pour représenter la composante spatiale des données. Pour représenter l'information géospatiale, une des normes les plus connues est GeoSPARQL, un standard de l'OGC qui définit une petite ontologie pour la représentation de caractéristiques, de relations et de fonctions spatiales [17, 4, 23]. Le cadre *geo* de la Figure 2 présente les principales classes de GeoSPARQL. La classe `geo:Feature` représente toute entité du monde réel ayant une empreinte spatiale. Cette empreinte est décrite par une "géométrie" (point, polygone, etc.), instance de la classe `geo:Geometry`. Une entité est liée à sa géométrie via la propriété `geo:hasGeometry`. Les coordonnées d'une géométrie sont décrites via la propriété `geo:asWKT`. Il est possible de lier GeoSPARQL à une ontologie de n'importe quel domaine en spécialisant la classe `geo:Feature` par une classe de l'ontologie de domaine considérée.

Parmi les alternatives à GeoSPARQL, GeoRDF sert à représenter des propriétés géographiques de points telles que la latitude, la longitude, l'altitude (en utilisant WGS84 comme référentiel) ; GeoOWL permet de représenter l'empreinte d'objets spatiaux plus complexes (lignes, rectangles, polygones) ; stSPARQL est une autre extension de SPARQL conçue pour interroger des graphes RDF spatio-temporels [18, 5]. Nous avons retenu GeoSPARQL car il offre la possibilité de raisonner sur des entités (`geo:Feature`) ayant une forme géométrique (`geo:Geometry`), à savoir proposer des relations entre des entités sur la base des relations entre leurs géométries (inclusion, recouvrement, etc). GeoSPARQL assure l'expression de relations spatiales binaires sous forme de propriétés des `geometries` et des `features`, et de fonctions servant à les comparer sur la base des propriétés. Parmi les fonctions spatiales de GeoSPARQL, on peut citer `equals`, `disjoint`, `intersects`, `touches`, `within`, *contains*, `overlaps`, et `crosses`. Grâce à des requêtes utilisant ces fonctions, il est possible de comparer, à la volée, deux `geometries` ou deux `features`, ou de comparer une ressource géolocalisée à une position fournie explicitement. La requête GeoSPARQL ci-après recherche les "objets" situés dans (fonction GeoSPARQL `sfWithin`) un polygone (un triangle) défini par les coordonnées de ses trois sommets :

```
select ?obj_loc
WHERE {
 ?obj_loc a geo:Feature .
 ?obj_loc geo:hasGeometry ?geo_obj .
 FILTER ( geof:sfWithin( ?geo_obs ,
   "POLYGON ((-3.562096 42.073807,
     -0.44234999999998 42.476536,
     -0.78979299999997 43.973148)
     )"^^geo:wktLiteral) )
}
```

Vocabulaire pour représenter la composante temporelle des données. Parmi les modèles de données et langages de requête du web sémantique gérant la composante temporelle de données, on trouve OWL-Time [12], le standard du W3C, et SWRL Temporal Ontology. Le cadre *time* de la Figure 2 présente les principales classes de OWL Time. La classe `time:TemporalEntity` représente toute entité ayant une temporalité, i.e. une date de début (propriété `time:hasBeginning`) et une date de fin (propriété `time:hasEnd`), et donc une durée (propriété `time:hasDuration`). Une entité dont la durée n'est pas nulle ("début" et "fin" sont différents) est un *intervalle*, sinon il s'agit d'un *instant* (`time:Interval` et `time:Instant` étant des spécialisations de `time:TemporalEntity`). Les entités temporelles peuvent être liées par des relations binaires comme `meets`, `overlaps`, `during` issues de l'algèbre des intervalles d'Allen et servant de base pour les raisonnements spatio-temporels.

Le temps a été associé aux triplets RDF à diverses fins. Ainsi, pour donner à un triplet RDF une dimension temporelle représentant sa période de validité (`valid-time`), notion issue des Bases de Données temporelles, qui s'oppose à celle de "date de transaction" (`transaction-time`) indiquant quand un élément a été enregistré dans la base, [18, 5] proposent le langage stRDF. Un fait y est représenté par un quadruplet *<s,p,o> :t* où *t* est une variable de type `xsd:dateTime` (un instant) ou `strdf:period` (un intervalle de temps). stSPARQL, le langage de requête associé à stRDF, propose des fonctions pour établir des relations temporelles et ainsi rechercher des faits qui se sont produits entre deux dates. Des représentations plus complexes permettent de représenter les évolutions d'objets au cours du temps. Ainsi le modèle Continuum s'appuie sur une ontologie des *fluents* et enrichit GeoSPARQL pour modéliser les changements qui s'opèrent sur des entités spatiales [14]. Pour représenter la date des observations, SOSA intègre une dimension temporelle basée sur OWL-Time. Nous utilisons donc aussi ce vocabulaire.

3.3 Les vocabulaires spécifiques du modèle

Nous avons donc adopté SOSA pour décrire les données d'observations (méta-données d'image, observations météorologiques, etc.). La propriété

12. `https://www.w3.org/TR/owl-time/` (10/2017)

©AfIA 2018

```
g-mfo:obs_07005_20180204150000_tminsol a mfo:Observation .
g-mfo:obs_07005_20180204150000_tminsol sosa:phenomenonTime
                                        g-mfo:interval_1517713200-1517756400 .
g-mfo:interval_1517713200-1517756400 a time:Interval .
g-mfo:interval_1517713200-1517756400 time:hasBeginning g-mfo:instant_1517713200 .
g-mfo:interval_1517713200-1517756400 time:hasEnd g-mfo:instant_1517756400 .
g-mfo:instant_1517713200 a time:Instant .
g-mfo:instant_1517713200 time:inXSDDateTime "2018-02-04T03:00:00"^^xsd:dateTime .
g-mfo:instant_1517713200 time:inXSDDateTimeStamp "1517713200"^^xsd:dateTimeStamp .
g-mfo:obs_07005_20180204150000_tminsol sosa:hasResult
                                        g-mfo:obs_07005_20180204150000_tminsol_result .
g-mfo:obs_07005_20180204150000_tminsol_result a mfo:Result .
g-mfo:obs_07005_20180204150000_tminsol_result a qudt-1-1:QuantitativeValue .
g-mfo:obs_07005_20180204150000_tminsol_result qudt-1-1:unit qudt-unit-1-1:Kelvin .
g-mfo:obs_07005_20180204150000_tminsol_result qudt-1-1:numericValue "274.75"^^xsd:double .
g-mfo:obs_07005_20180204150000_tminsol sosa:hasFeatureOfInterest
                                        <http://melodi.irit.fr/lod/mfo/foi_07005> .
g-mfo:obs_07005_20180204150000_tminsol sosa:observedProperty g-mfo:Temperature .
g-mfo:obs_07005_20180204150000_tminsol sosa:madeBySensor g-mfo:station_07005_Thermometer .
g-mfo:obs_07005_20180204150000_tminsol sosa:usedProcedure g-mfo:procedure_tminsol .
```

TABLE *1 – Extrait de la représentation RDF d'un relevé de température réalisé le 04/02/2018, entre 3h et 15h.*

`sosa:phenomenomTime` ayant `sosa:Observation` pour domaine et `time:TemporalEntity` pour co-domaine sert à représenter la période de temps durant laquelle le résultat d'une observation a été obtenu. Pour représenter les relations spatiales, nous complétons notre modèle avec GeoSPARQL. Nous partons du principe que pour toute observation géolocalisée, son "*point d'intérêt*" (`FeatureOfInterest`) correspond à la zone observée ; cette zone étant caractérisée par une géométrie, elle est représentée comme un `geo:Feature`. Ainsi, pour chaque type d'observations issues de capteurs, notre approche consiste à définir un nouveau vocabulaire, avec un espace de nom spécifique, approprié à la source des données. Ce vocabulaire comporte au moins une classe qui spécialise à la fois `sosa:FeatureOfInterest` et `geo:Feature`.
Les deux cadres en bas de la Figure 2 correspondent à deux vocabulaires développés selon cette approche pour représenter les méta-données d'images satellites pour l'un (*eom*), les mesures de stations météorologiques pour l'autre (*mfo*). Ces vocabulaires sont présentés dans la partie 4.1. Dès lors qu'on est capable de représenter la composante spatiale d'une image comme instance de `geo:Feature`, on peut la lier à d'autres informations ayant une composante spatiale également définie comme une `geo:Feature`. De même, on peut lier par une relation temporelle un enregistrement de méta-données d'image avec des mesures observées par ailleurs, telles que les données météo, ou avec des périodes d'intérêt (par exemple "une semaine après la prise de l'image").

4 Entrepôts RDF pour les données d'observations de la Terre

Nous précisons ici comment nous avons adapté les ontologies réutilisées pour concevoir un vocabulaire propre à chaque source de données. Parmi les données géolocalisées, nous avons distingué les *données dynamiques*, dont la validité est fournie par la composante temporelle, et les *données statiques*, qui ne sont pas datées a priori ou auxquelles sont associés de très longs intervalles de temps. Nous présentons à présent les deux (sous-)modèles résultant et leur instanciation pour enrichir les méta-données d'images. Il est à noter que la propriété `time:hasTime` de OWL-Time est définie comme "un prédicat générique pour associer une entité temporelle à n'importe quoi". Donc même si notre modèle ne le mentionne pas, toute entité décrite par une URI peut être estampillée par une composante temporelle (une date ou un intervalle de temps).

4.1 Les observations (données dynamiques)

Le modèle est spécialisé dans 2 modules dédiés à chacune des sources d'observations issues de capteurs décrites ici, à savoir les images satellites et les observations météo.

Les méta-données d'images. Dans le projet SparkInData, nous utilisons des enregistrements de méta-données d'images Sentinel [13]. La périodicité de Sentinel-1 est de douze jours, celle de Sentinel-2 étant de cinq jours. Les enregistrements de méta-données sont obtenus au format GeoJSON à partir de RESTO, un service géré par le CNES [12]. L'API RESTO permet de spécifier les paramètres à retrouver, à savoir des métadonnées d'enregistrements comme la couverture nuageuse, l'intervalle de temps, la zone géographique d'intérêt, etc. Nous collectons ces informations toutes les nuits. L'URL suivante fait appel à RESTO pour retourner tous les enregistrements de métadonnées de la collection S2ST pour la France, captés entre le 19/09/2017 `23:00` et le 25/09/2017 `00:00` : `https://peps.cnes.fr/resto/api/collections/S2ST/search.json?q=France\&startDate=17-09-19T23:00:00\&completionDate=2017-09-25T00:00:00.`

13. `https://sentinel.esa.int/web/sentinel/missions/` (07/2018)

Les méta-données d'images sont obtenues sous forme de fichiers GeoJSON ; elles sont converties en RDF et représentées à l'aide du vocabulaire *eom* comme des "observations de la Terre", i.e. des instances de `eom:EarthObservation` (cadre *eom* de la Figure 2). Cette classe étant une sous-classe de `sosa:Observation`, la composante spatiale des observations est fournie via la classe *eom :Footprint*, spécialisation de `geo:Feature` et `sosa:FeatureOfInterest` : un `footprint` est décrit par un polygone fermé (une géométrie) délimitant la zone terrestre couverte par l'image. La classe `eom:Footprint` étant une sous-classe de `geo:Feature`, connaissant l'empreinte des images (leur `footprint`), il est facile d'identifier les caractéristiques d'un autre type de données qui recouvrent la position des images. La *dimension temporelle* d'un enregistrement de méta-données d'image indique le moment où l'image a été prise. Elle est fournie via la propriété `sosa:phenomenomTime` de l'observation correspondante.
Le module est complété par ailleurs pour fournir des données propres aux images satellites telles que le domaine spectral de l'image (infrarouge, UV, etc.), la direction orbitale du satellite (ascendante, descendante) ou encore la couverture nuageuse de l'image.

Les observations météorologiques. Comme données contextuelles dynamiques, nous utilisons les informations météo de `SYNOP Meteo France` [14]. Ces observations sont réalisées toutes les trois heures par chacune des 62 stations françaises. Un fichier séparé contient la liste des stations avec leur position respective (un point fixe repéré par ses coordonnées géographiques). Selon une approche très similaire à celle de [19], nous utilisons des classes tirées de SOSA que nous spécialisons pour définir des classes propres au domaine de la météorologie. Dans un premier temps, nous sommes restés au plus près du schéma de la source MétéoFrance utilisée. Nous envisageons de faire évoluer ce modèle en intégrant le vocabulaire SWEET [24], devenu un standard pour les données météorologiques.
Nous représentons une station météo comme une instance de la classe `mfo:MeteoStation`, une sous-classe de `sosa:Platform`. Les capteurs fonctionnant sur une station météo sont représentés comme instances de `mfo:MeteoSensor`, sous-classe de `sosa:Sensor`. La position géographique où s'effectuent les mesures est représentée comme instance de la classe `mfo:MeteoFeatureOfInterest`, une sous-classe de `sosa:FeatureOfInterest`.
Comme pour *eom*, le module *mfo* a été complété pour représenter les variables fournies avec les relevés météo ; il y a 28 variables par relevé, mais beaucoup d'entre elles ne sont pas renseignées. Nous avons ainsi instancié la classe `sosa:ObservableProperty` avec notamment des individus représentant la température (`mfo:Temperature`), l'humidité (`mfo:HUMIDITY`) ou la vitesse du vent (`mfo:WIND_SPEED`), variables dont une valeur est relevée toutes les trois heures (cf. Table 1). Le module est enrichi régulièrement par de nouvelles instances (tous les 5 ou 12 jours pour les images satellites, quotidiennement pour les relevés météo).

4.2 Les données supports (statiques)

Le modèle peut être enrichi pour représenter des données dites statiques, i.e. qui ont une composante spatiale mais dont la représentation de la composante temporelle n'est pas nécessaire car leur valeur ne varie pas dans l'échelle de temps considérée. Nous avons créé deux modules de ce type, *admin* et *grid*, présentés sur la Figure 3. Pour mettre en relation ces données avec les données dynamiques via des relations spatiales, un concept de chacun de ces modules spécialise la classe `geo:Feature`. Les classes de ces modules sont instanciées une fois pour toutes.

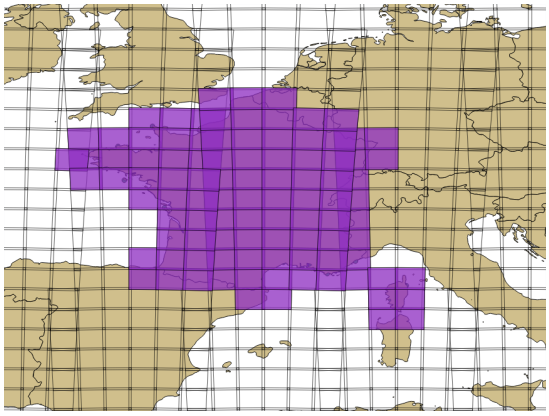


FIGURE *4 – Les tuiles de l'ESA définies pour la France métropolitaine*

Les tuiles d'images. Les images Sentinel ont des caractéristiques différentes en fonction du capteur qui les a prises. En septembre 2016, l'ESA a commencé à diffuser des images Sentinel 2 sous forme de paquets de tuiles simples (images S2ST), chaque tuile représentant une zone fixe de la surface du globe de taille approximative de 100 x 100 km. Une image S2ST correspond donc à un fragment d'image S2. L'intérêt par rapport à une image S2 est que l'utilisateur peut mieux sélectionner la surface qui l'intéresse et ne télécharger que l'information souhaitée. Un fichier S2ST est aussi moins volumineux : il peut faire environ 500 Mo alors que celui d'une image S2 avant tuilage peut faire plus de 3 Go.
La grille de l'ESA décrivant le tuilage de la surface terrestre (Cf. Figure 4), i.e l'empreinte de chaque tuile, est fournie dans un fichier Grid au format KML [15]. Nous avons converti ce fichier en RDF à l'aide du vocabulaire *grid* dé-

14. `https://donneespubliques.meteofrance.fr/` (07/2018)

15. `https://sentinel.esa.int/web/sentinel/missions/sentinel-2/news/-/article/sentinel-2-tiling-grid-updated`

©AfIA 2018

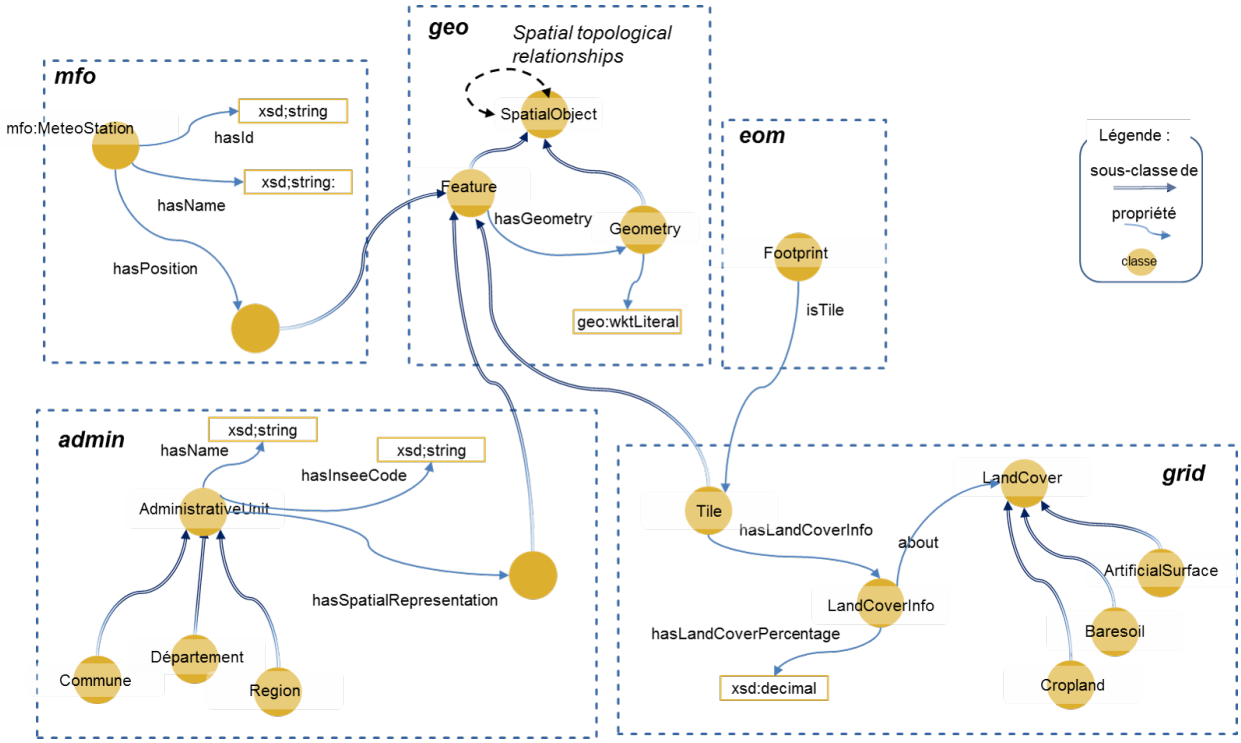


FIGURE *3 – Les vocabulaires* admin *et*grid *intégrant des données "statiques".*

crit en bas de la Figure 3 : chaque tuile est représentée comme une instance de la classe `Tile`, spécialisation de `geo:Feature`, dont la propriété `geo:hasGeometry` correspond à son empreinte. Le module *eom* (cf. Section 4.1) a été enrichi avec la propriété `eom:isTile` pour associer une tuile au `footprint` d'une image. Ainsi en reliant des données à une tuile, on associe indirectement ces données à toutes les images associées à cette tuile.

La couverture terrestre. Pour associer des données aux tuiles, nous avons utilisé une autre source de données statiques, le GLC-SHARE (Global Land Cover SHARE) produit par le FAO, qui donne des informations sur la couverture terrestre. Cette dernière est définie à partir d'une nomenclature permettant de classer les zones en fonction du type d'occupation des sols ou du type de surface (surface artificielle, terre cultivée, zone forestière, etc.). Les données du GLC-SHARE sont fournies sous forme d'une image au format TIFF dont chaque pixel correspond à une échelle spatiale approximativement de 1 km^2. La valeur d'un pixel est un entier indiquant la classe la plus fréquente pour la zone couverte par le pixel. Nous avons donc calculé la composition de la couverture terrestre de chaque tuile de l'ESA de la France. Avec le vocabulaire *grid*, une tuile est liée à un ensemble de graphes RDF (via la propriété `hasLandCoverPercentage`) décrivant chacun le pourcentage d'une classe GLC-SHARE (`CropLand`, `Baresoil`, etc.) sur la surface couverte par la tuile.

Les stations météorologiques. Les informations sur les stations météo, tirées du fichier fourni par *SYNOP Meteo France*, fournissent entre autre la position géographique des mesures météo. Le modèle *mfo* (cf. Section 4.1) a ainsi été complété notamment avec la propriété `hasPosition` dont le codomaine est une spécialisation de `geo:Feature`. Le code suivant est un extrait de la représentation RDF de la station météorologique d'Embrun :

```
g-mfo:MS_07591 rdf:type    mfo:MeteoStation .
g-mfo:MS_07591 sosa:hosts g-mfo:Sensor_07591_humidity .
g-mfo:MS_07591 sosa:hosts g-mfo:Sensor_07591_temperature .
g-mfo:MS_07591 mfo:hasId "07591"^^xsd:String .
g-mfo:MS_07591 mfo:hasName "EMBRUN"^^xsd:String .
g-mfo:MS_07591 mfo:hasPosition g-mfo:MS_pos_07591 .
g-mfo:MS_pos_07591 rdf:type  geo:Feature .
g-mfo:MS_pos_07591 geo:hasGeometry g-mfo:MS_geo_07591 .
g-mfo:MS_geo_07591 rdf:type    geo:Geometry .
g-mfo:MS_geo_07591 geo:asWKT
       "POINT ((6.502333 44.565667))^^geo:wktLiteral" .
```

La Figure 5 est un graphe d'instances illustrant (partiellement) la représentation d'observations météo avec le vocabulaire *mfo* : 2 observations mesurant respectivement les précipitations et la pression atmosphérique, ont été réalisées par la même la station météo. La figure met en évidence qu'une seule URI (`g-mfo:MeteoStation_pos_07005` sur l'exemple) est utilisée pour représenter à la fois la position de la station (propriété `mfo:hasPosition`) et le lieu de chacune des observations (propriété `sosa:hasFeatureOfInterest`).

Les unités administratives. Afin de lier les observations de la Terre à des entités administratives (villes, départe-

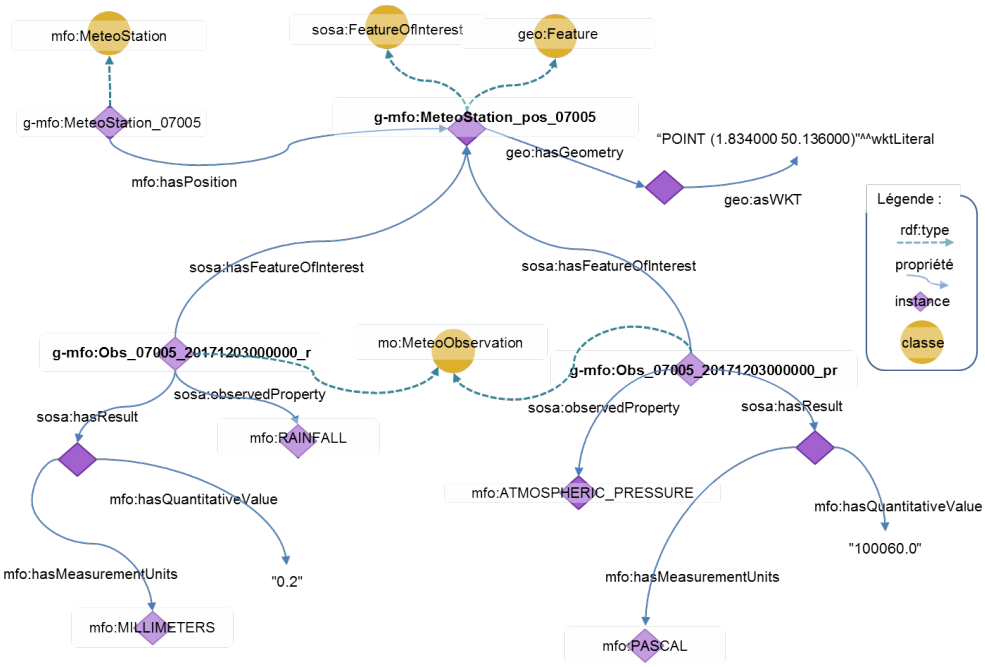


FIGURE 5 – *Instanciation du sous-modèle "mfo" : le "feature of interest" des observations d'une station météo est le lieu où se trouve la station.*

ments, régions, etc.) françaises à partir de leur position géographique (point ou polygone), nous avons enrichi le modèle avec le module *admin* (Cf. Figure 3). La classe `AdministrativeUnit` est caractérisée par une propriété `hasSpatialRepresentation` dont le co-domaine est une sous-classe de `geo:Feature`.

Le modèle a été instancié avec des données provenant de la plate-forme ouverte des données publiques françaises `data.gouv.fr` [16]. Ces données sont fournies au format shapefile. Les entités administratives de notre triplestore (instances de `AdministrativeUnit`) sont alignées via la propriété `owl:sameAs` avec les territoires français de l'ontologie géographique de l'INSEE [17], qui ne peut pas être utilisée directement pour établir des relations spatiales car elle ne contient pas d'information géométrique. Ci-après un exemple de graphe RDF, celui de la région 54 :

```
admin:region_54 a admin:AdministrativeUnit .
admin:region_54 a admin:Region .
admin:region_54 owl:sameAs <http://id.insee.fr/geo/region/54> .
admin:region_54 admin:hasInseeCode "54"^^xsd:String .
admin:region_54 admin:hasName "Poitou-Charentes"^^xsd:String .
admin:region_54 a geo:Feature .
admin:region_54 geo:hasGeometry l_admin:region_54_geo .
admin:region_54_geo geo:asWKT "MULTIPOLYGON(((
-1.0988062299633785 45.64032288975508, ...
-1.0988062299633785 45.64032288975508))...)"^^wkt:Literal .
```

16. https://www.data.gouv.fr
17. http://rdf.insee.fr/def/index.html

5 Intégration de données RDF via leurs relations spatiales

Chacun des modules des Figures 2 et 3 est décrit dans un fichier OWL spécifique. Le modèle global est ainsi composé de 4 ontologies [18]. Le calcul des relations spatiales et temporelles permet de relier les données de chacun de ces entrepôts, et exploite le tuilage. En liant des éléments spatiaux aux tuiles de l'ESA, les connaissances s'y rapportant sont extrapolées au niveau des images. Par exemple, sachant que les images $[img1, img2, img3]$ partagent la tuile $tile_1$, et que cette tuile recouvre $adminUnit_i$, il est possible d'inférer que $[img1, img2, img3]$ recouvrent aussi $adminUnit_i$.
Pour calculer une relation spatiale R (appartenance ou non-intersection, par exemple) entre deux ensembles de données *S* et *T* d'une base de connaissances, nous comparons chaque entité *s* de *S* à chaque entité *t* de *T*, pour vérifier l'existence de *R(s,t)*. Ceci est réalisé à l'aide d'une requête comme celle ci-dessous qui recherche l'ensemble des triplets `< s, geo:sfIntersect, t >` tels que les empreintes géographiques de *s* et *t* ont une intersection non vide :

```
CONSTRUCT {?s geo:sfIntersects ?t .}
WHERE {
 GRAPH <S> {?s geo:hasGeometry/geo:asWKT ?s_geometry.}
 GRAPH <T> {?t geo:hasGeometry/geo:asWKT ?t_geometry.}
 FILTER( geof:sfIntersects(?s_geometry , ?t_geometry) )
 }
```

18. http://melodi.irit.fr/ontologies/eom.owl; http://melodi.irit.fr/ontologies/mfo.owl; http://melodi.irit.fr/ontologies/grid.owl; http://melodi.irit.fr/ontologies/administrativeUnits.owl

©AfIA 2018

Cette approche, consistant à faire le produit cartésien de deux ensembles, est de complexité quadratique si bien que, dès que les jeux de données à combiner sont volumineux, calculer ces relations au moment de l'interrogation peut être extrêmement coûteux en temps et inacceptable pour les applications temps-réel. C'est pourquoi il est préférable de les pré-calculer. Nous avons distingué le cas des données à composante spatiale seule (statiques) de celles ayant en plus une composante temporelle (dynamiques).

5.1 Intégration des données statiques

Pour les données statiques, il est raisonnable d'envisager d'enregistrer dans l'entrepôt RDF les relations spatiales entre les données RDF correspondantes, si ces jeux de données ont une taille raisonnable. Si le volume est trop conséquent, il est nécessaire d'avoir recours à des techniques d'optimisation. Ainsi, en nous appuyant sur l'indexation spatiale fournie par le tuilage des images S2ST, nous avons calculé les relations entre les données de chacun des jeux que nous venons de présenter : stations météo de *mfo*, unités administratives de *admin* et tuiles de *grid*. Ainsi, il est possible de lier les stations météorologiques aux images en calculant uniquement les relations spatiales entre les stations météo et les tuiles de l'ESA. La requête SPARQL ci-dessous interroge les jeux de données instanciant les modèles *eom*, *grid* et *admin*. Elle retrouve les images S2ST avec une couverture nuageuse inférieure à 10%, collectées à un moment donné, et dont l'empreinte géographique couvre (propriété `geo:sfContains`) une zone géographique particulière et ayant un type d'occupation du sol particulier. L'utilisateur a défini une date qui, comme la couverture nuageuse et le type d'occupation du sol, est utilisée pour filtrer les données pertinentes (cf. les deux dernières clauses FILTER).

```
select distinct ?s2st_result ?tileId
{
?s2st a md:EarthObservation .
?s2st md:featureOfInterest ?fi .
?s2st md:result ?s2st_result .
?s2st_result md:product ?s2st_product.
?s2st_product md:cloudCoverPercentage ?s2st_cloudCover .
?s2st md:phenomenomTime ?s2st_period .
?s2st_period time:hasBeginning ?s2st_time .
?fi md:isTile ?tile .
?tile geo:hasGeometry ?tile_geo .
?tile_geo geo:sfContains ?admin_geo .
?adminUnit admin:hasSpatialRepresentation ?admin_sr .
?admin_sr geo:hasGeometry ?admin_geo .
?adminUnit admin:hasName ?adminUnitName .
?adminUnitName bif:contains 'Alpes' .
?tile grd:hasLandCoverInfo ?tile_lc .
?tile_lc grd:hasLandCoverPercentage ?lc_perc .
?lc_perc grd:about ?lc_class .
?lc_perc grd:percentage_value ?lc_value .
?tile grd:tileId ?tileId .
FILTER (?s2st_time = ?user_time)
FILTER (?lc_value>1 and ?lc_class=grd:ArtificialSurface
            and ?s2st_cloudCover<10)
}
```

Nous avons testé deux méthodes pour calculer les relations spatiales et créer les triplets RDF les représentant. Nous venons d'exposer la première qui met en oeuvre des requêtes SPARQL incluant des fonctions GeoSPARQL. La seconde a consisté à développer un programme en Python, à l'aide du module Shapely qui permet de faire des opérations spatiales sur des données au format WKT. Avec la solution Python, il faut 0,50s (sans optimisation particulière) pour calculer toutes les relations spatiales entre 1 élément d'un jeu de données *S* et 50 éléments d'un jeu de données *T*, alors qu'il faut 11s avec GeoSPARQL pour calculer une seule relation entre 1 élément de *S* et 50 éléments de *T*.

5.2 Intégration des données ayant une composante temporelle

Il est possible d'établir des relations temporelles entre un enregistrement de méta-données d'image et des données ayant une indication temporelle comme les relevés météo, ou des périodes d'intérêt définies par l'utilisateur. A défaut de disposer d'entités qui servent de référentiel temporel (le pendant des tuiles de l'ESA pour la composante spatiale), nous exploitons l'intervalle de temps défini par un utilisateur lors de la recherche d'images, comme un buffer temporel fournissant un contexte aux enregistrements de méta-données.

La Figure 6 illustre ce principe. Une interface permet à l'utilisateur de dessiner un rectangle pour spécifier des contraintes spatiales et de saisir une période de temps pour définir un empan temporel. A partir de ces informations, le système génère une requête GeoSPARQL qui permet de retrouver les observations météo, les métadonnées et les images dont la composante temporelle recouvre l'empan temporel et dont la localisation vérifie les contraintes spatiales (plus précisément, l'empreinte de ces images a une intersection non nulle avec le rectangle choisi par l'utilisateur). Les images recherchées sont des images Sentinel 2 "classiques", qui ne sont pas liées à une unique tuile comme les S2ST, et donc pour lesquelles les relations spatiales n'ont pas été pré-calculées. La requête SPARQL ci-dessous sélectionne des enregistrements de méta-données d'images dont les mesures météorologiques associées respectent deux contraintes :

1. elles proviennent de stations situées dans l'empreinte de l'image : `?img_geo` étant la géométrie d'une image et `?ws_geo` celle d'une station météo, la géométrie de l'image doit contenir celle de la station météo, ce qu'indique la fonction `geof:sfContains` dans le filtre `FILTER(geof:sfContains(?img_geo,?ws_geo))`;
2. elles ont été collectées durant une période donnée ; l'utilisateur a défini une période d'intérêt d'une semaine après la prise d'image (24x7=168 heures, 'PT168H'), qui est utilisée au sein des deux dernières clauses `FILTER`.

```
SELECT  ?img ?obs_time ?humidityDataVal ?humidityUnits
WHERE{
?img a eom:EarthObservation .
?img eom:featureOfInterest ?img_foi .
?img_foi geo:hasGeometry ?img_geo .
?img eom:phenomenonTime ?img_temp .
?img_temp time:hasEnd/time:inXSDDateTime ?img_end .
?ws a mfo:MeteoStation .
?ws mfo:hasPosition ?ws_pos .
?ws_pos geo:hasGeometry ?ws_geo .
?ws sosa:hosts ?sensor .
?sensor sosa:observes mfo:HUMIDITY .
?obs sosa:madeBySensor ?sensor .
?obs sosa:hasResult ?humidity .
?humidity mfo:hasQuantitativeValue ?humidityDataVal .
?humidityVal mfo:isClassifiedBy ?humidityUnits .
?obs sosa:phenomenonTime ?obs_time .
?obs_time time:hasEnd/time:inXSDDateTime ?obs_end .
BIND (?img_temp-?obs_time as ?diffDateTime)
FILTER (geof:sfContains(?img_geo,?ws_geo))
FILTER (?diffDateTime<'PT168H'^^xsd:dayTimeDuration)
FILTER (?diffDateTime>'PT0H'^^xsd:dayTimeDuration)
}
```

Faute de place, nous n'avons pas indiqué les éléments de la requête filtrant les images présentes dans la zone spécifiée par l'utilisateur.

La Figure 6 montre l'information météorologique associée à une image faisant partie de la liste retournée par une requête de ce type. L'empreinte de l'image est représentée par un polygone

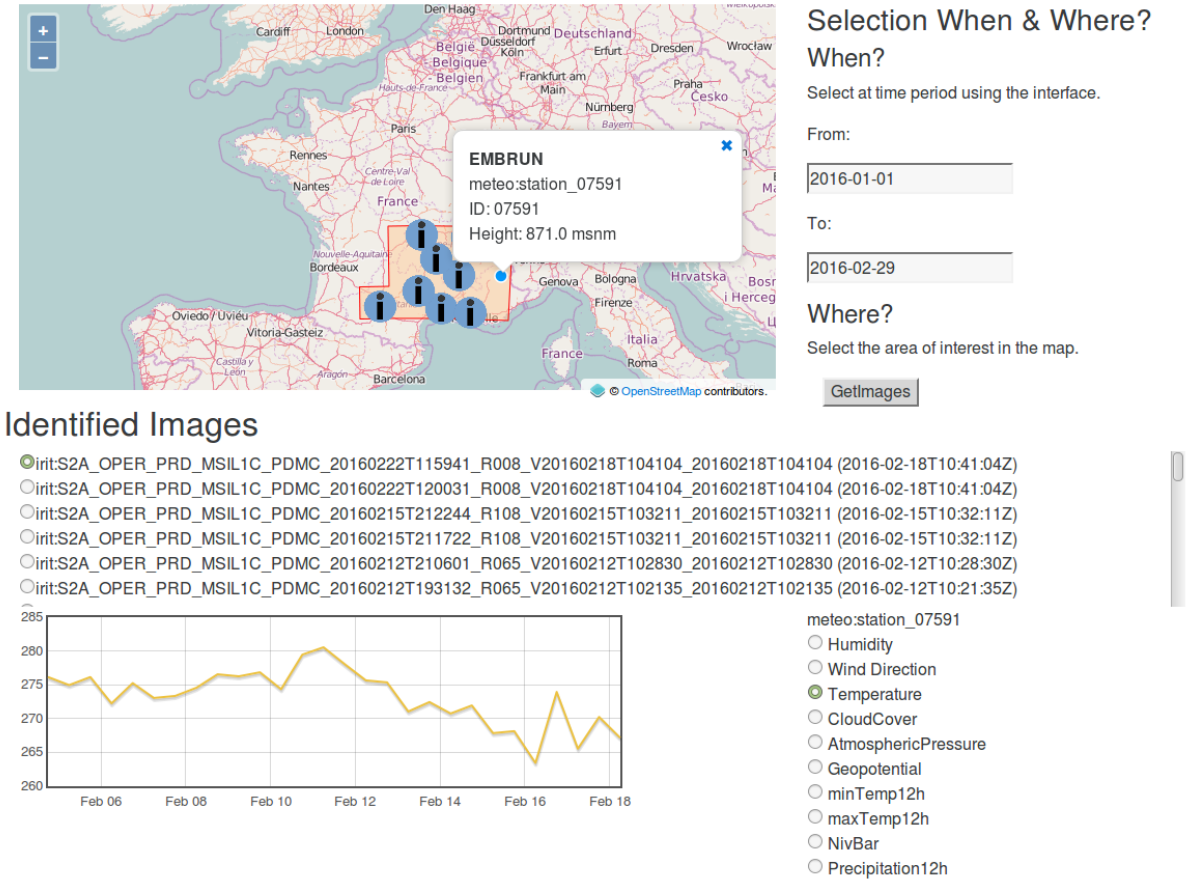


FIGURE 6 – *Affichage des informations météorologiques associées à une image satellite en réponse à la requête.*

orange. Chaque cercle bleu correspond à une station météorologique. En bas de la carte, figurent des séries temporelles qui représentent l'évolution de la variable *Temperature* pendant la période choisie.

6 Conclusion

Nous avons proposé une ontologie pour l'intégration de données d'observation de la Terre et donnés contextuelles, basée sur des relations topologiques spatiales et temporelles. Cette ontologie spécialise des standards, notamment SOSA, GeoSPARQL et OWL-Time. Nous avons défini également un processus d'intégration qui sélectionne, transforme et intègre de données géospatiales hétérogènes (méta-données d'image satellite, données météorologiques, unités administratives, couverture terrestre, etc.). Ce processus s'appuie sur le tuilage des images pour traiter les données ayant une composante spatiale fixe, les relations temporelles, quant à elle, sont calculées à la volée à partir d'une topologie temporelle. Nous avons présenté un cas d'étude basé exploitant des méta-données d'image satellite Sentinel.

Dans la continuité de ces travaux, nous envisageons de considérer des sources propres à un domaine métier pour traiter un cas d'usage (l'agriculture et des rapports bulletins agricoles) et fournir des règles et des fonctionnalités de raisonnement pour faciliter les analyses. Il serait intéressant aussi d'identifier des patrons qui augmenteraient la pertinence de l'image, en s'appuyant sur des règles et du raisonnement. On pourrait par exemple analyser les séries temporelles de température et détecter une canicule. Nous envisageons également d'intégrer des alignements automatiques pour mettre en correspondance les métadonnées d'images et les vocabulaires, mais également les données annotées et d'autres sources du LOD.

Références

[1] H. Arenas, N. Aussenac-Gilles, C. Comparot, and C. Trojahn. Semantic integration of geospatial data from earth observations. In *Knowledge Engineering and Knowledge Management - EKAW 2016 Satellite Events*, pages 97–100, 2016.

[2] H. Arenas, N. Aussenac-Gilles, C. Comparot, and C. Trojahn. Un modèle pour l'intégration sémantique de données géolocalisées liées à l'observation de la terre. In *Spatial Analysis and Geomatics - Atelier EXtraction de Connaissances à partir de donnéEs Spatialisées (SAGEO-EXCES), Nice*, 2016.

[3] G. A. Atemezing. *Publishing and consuming geo-spatial and government data on the semantic web*. PhD thesis, Thesis, 04 2015.

[4] R. Battle and D. Kolas. Enabling the Geospatial Semantic Web with Parliament and GeoSPARQL. *Semantic Web*, 3(October 2012) :355–370, 2012.

©AfIA 2018

[5] K. Bereta, P. Smeros, and M. Koubarakis. Representation and querying of valid time of triples in linked geospatial data. In *The Semantic Web : Semantics and Big Data*, pages 259–274, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[6] L. M. V. Blázquez, B. Villazón-Terrazas, Ó. Corcho, and A. Gómez-Pérez. Integrating geographical information in the linked digital earth. *Int. Journal of Digital Earth*, 7(7) :554–575, 2014.

[7] D. Brizhinev, S. Toyer, K. Taylor, and Z. Zhang. Publishing and using earth observation data with the rdf data cube and the discrete global grid system. Technical report, W3C and OGC, 2017.

[8] M. Console and M. Lenzerini. Reducing global consistency to local consistency in ontology-based data access - extended abstract. In *Informal Proceedings of the 27th International Workshop on Description Logics, Vienna, Austria, July 17-20, 2014.*, pages 496–499, 2014.

[9] S. Cox. Prov ontology supports alignment of observational data (models). In *Modeling and Simulation Society of Australia and New Zealand*, pages 403–409, 2017.

[10] D. Espinoza-Molina and M. Datcu. Earth-observation image retrieval based on content, semantics, and metadata. *IEEE Trans. on Geoscience and Remote Sensing*, 51(11) :5145–5159, 2013.

[11] D. Espinoza-Molina, C. Nikolaou, C. O. Dumitru, K. Bereta, M. Koubarakis, G. Schwarz, and M. Datcu. Very-High-Resolution SAR Images and Linked Open Data Analytics Based on Ontologies. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(4) :1696 – 1708, 2015.

[12] J. Gasperi. Semantic Search Within Earth Observation Products Database Based on Automatic Tagging of Image Content. In *Proc. of the Conf. on Big Data from Space*, pages 4–6, 2014.

[13] A. Gore. The digital earth. *Australian Surveyor*, 43(2) :89–91, 1998.

[14] B. Harbelot, H. Arenas, and C. Cruz. un modèle sémantique spatio-temporel pour capturer la dynamique des environnements. In *14 ème conférence Extraction et Gestion des Connaissances*, page 39 à 54, Rennes, France, 2014.

[15] T. Heath and C. Bizer. *Linked Data : Evolving the Web into a Global Data Space ; Lectures on the Semantic Web : Theory and Technology*. Morgan & Claypool, 2011.

[16] C. Keßler and C. J. Q. Farmer. Querying and integrating spatial-temporal information on the Web of Data via time geography. *Journal of Web Semantics*, 35 :25–34, 2015.

[17] D. Kolas, M. Perry, and J. Herring. Getting started with GeoSPARQL. Technical report, OGC, 2013.

[18] M. Koubarakis and K. Kyzirakos. Modeling and querying metadata in the semantic sensor web : The model strdf and the query language stsparql. In *The Semantic Web : Research and Applications*, pages 425–439, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[19] L. Lefort, J. Bobruk, A. Haller, K. Taylor, and A. Woolf. A linked sensor data cube for a 100 year homogenised daily temperature dataset. In *Proc. of the 5th International Conference on Semantic Sensor Networks - Volume 904*, SSN'12, pages 1–16. CEUR-WS.org, 2012.

[20] M. Lefrançois, A. Zimmermann, and N. Bakerally. A SPARQL extension for generating RDF from heterogeneous formats. In *Proc. Extended Semantic Web Conference (ESWC'17)*, Portoroz, Slovenia, May 2017.

[21] M. Lenzerini. Ontology-based data management. In *Proc. of the 20th ACM Int. Conference on Information and Knowledge Management*, CIKM '11, pages 5–6, New York, USA, 2011. ACM.

[22] F. Maali and J. Erickson. Data Catalog Vocabulary (DCAT), 2014.

[23] M. Perry and J. Herring. OGC GeoSPARQL-A geographic query language for RDF data. Technical report, Open Geospatial Consortium, 2012.

[24] R. Raskin. Guide to SWEET Ontologies. Technical Report 9, NASA/Jet Propulsion Lab, Pasadena, CA, USA, 2006.

[25] F. Reitsma and J. Albrecht. Modeling with the semantic web in the geosciences. *IEEE Intelligent Systems*, 20(2) :86–88, 2005.

[26] D. Sukhobok, H. Sánchez, J. Estrada, and D. Roman. Linked data for common agriculture policy : Enabling semantic querying over sentinel-2 and lidar data. In *Proceedings of the ISWC 2017 Posters & Demonstrations and Industry Tracks co-located with 16th Int. Semantic Web Conference, Vienna, Austria, Oct. 23-25, 2017*, 2017.

[27] J. Tandy, L. van den Brink, and P. Barnaghi. Spatial data on the web best practices, w3c working group note. Technical report, W3C and OGC, 2017.

[28] TELEIOS. Virtual Observatory Infrastructure for Earth Observation Data. `http://www.earthobservatory.eu/search/node/ontologies`, 2016. Accessed : 2018-03.

[29] The TELEIOS Team. Building Remote Sensing Applications Using Scientific Database And Semantic Web Technologies. In *Image Information Mining Conference : Knowledge Discovery from Earth Observation Data*, pages 2–5, 2012.

Choisir un encodage CNF de contraintes de cardinalité performant pour SAT

T. Delacroix
IMT Atlantique - Dépt. LUSSI, Brest, France
thomas.delacroix@imt-atlantique.fr

Résumé

Cet article répond à une double problématique : (1) comment choisir un encodage CNF pour des contraintes de cardinalité de type $\#k(x_1, ..., x_n)$ où $\#$ peut être l'un des symboles $\leq, =, \geq$; (2) déterminer un encodage CNF performant pour les contraintes de cardinalité plus générales $\in \mathcal{K}(x_1, ..., x_n)$ où $\mathcal{K} \subset [\![0, n]\!]$. Pour ce faire, on introduit d'abord un nouvel encodage séquentiel bidirectionnel. On décrit alors un processus pour choisir l'encodage le plus performant pour une contrainte de cardinalité donnée s'appuyant sur une comparaison de différents encodages pour tous les cas possibles de valeurs n et k. Enfin, on montre que l'encodage séquentiel bidirectionnel permet de répondre à la problématique (2).

Mots Clef

CNF, SAT, encodage, contraintes de cardinalité.

Abstract

This article has a double aim : (1) define a process for choosing the most efficient CNF encoding for cardinality constraints of type $\#k(x_1, ..., x_n)$ where $\#$ is one of the following symbols $\leq, =, \geq$; (2) determine an efficient CNF encoding for the more general cardinality constraints of type $\in \mathcal{K}(x_1, ..., x_n)$ where $\mathcal{K} \subset [\![0, n]\!]$. In order to do this, we introduce a new sequential bidirectional encoding. We then describe a process for choosing the most efficient encoding for a given cardinality constraint based on a comparison of different encodings for all possible values of n and k. Finally, we show that the sequential bidirectional encoding can be used to reach our second aim.

Keywords

CNF, SAT, encoding, cardinality constraints.

1 Introduction

Parmi les solveurs modernes en programmation sous contrainte les plus performants, on trouve aujourd'hui un certain nombre de solveurs CNF-SAT. Les palmarès récents du MiniZinc Challenge en témoignent [13, 11].
Un solveur CNF-SAT permet d'obtenir une valuation pour laquelle une expression logique sous forme conjonctive normale (CNF) donnée est satisfaite lorsqu'il en existe une. Pour une contrainte particulière, la performance du solveur dépend donc à la fois de l'algorithme de résolution du solveur et à la fois de la façon dont la contrainte est exprimée sous forme CNF.

De nombreux travaux se sont donc penchés sur la question de savoir comment exprimer des contraintes classiques des CSP sous forme CNF de manière performante pour les solveurs CNF-SAT [1, 2, 3, 4, 5, 7, 10, 14, 15]. C'est le cas notamment pour les contraintes de cardinalité [1] de type $\#k(x_1, ..., x_n)$ où $\#$ peut être l'un des symboles $\leq, =, \geq$. En effet, un encodage naïf de ces contraintes contient, dès que n augmente, beaucoup trop de clauses pour pouvoir être utilisé de manière raisonnable en pratique. De nombreux encodages CNF ont donc été proposés, fonctionnant tous sur le même principe général : des variables supplémentaires sont introduites de manière à réduire drastiquement le nombre de clauses.

Parmi les travaux existants, on trouve des comparaisons des différents encodages proposés [2, 9, 12]. Toutefois, ces comparaisons ne sont pas exhaustives. En effet, l'accent y est généralement mis sur le comportement des encodages lorsque n tend vers l'infini. Or, en pratique, on peut également être amené à considérer de très nombreuses contraintes de cardinalité de faible dimension. Une comparaison exhaustive des encodages existants s'impose donc afin d'essayer d'optimiser au maximum l'étape du choix de l'encodage dans la résolution SAT et cela constitue un des éléments central de cet article. On s'aperçoit alors qu'il n'y a pas un encodage plus performant que tous les autres mais de nombreux encodages performants selon les paramètres du problème. On montre également qu'il est possible de combiner des encodages pour obtenir de meilleures performances.

Par ailleurs, on introduit dans cet article un nouvel encodage : l'encodage séquentiel bidirectionnel. Cet encodage, dont la définition est assez naturelle, est particulièrement adapté pour considérer des contraintes de cardinalité plus complexes, notamment les contraintes de cardinalité correspondant à un intervalle. On montre par ailleurs qu'il permet de considérer des contraintes de cardinalité de type $\in \mathcal{K}(x_1, ..., x_n)$ où $\mathcal{K} \subset [\![0, n]\!]$ de manière performante ce qui représente une nouveauté.

1. Le terme contrainte de cardinalité est utilisé ici selon la nomenclature standard dans le contexte SAT et ne doit pas être confondu avec d'autres notions telles que celle de *global cardinality constraint* utilisée dans le contexte de la programmation sous contrainte.

©AfIA 2018

2 Encodage séquentiel bidirectionnel

Dans la suite, on considère des entiers n et k tels que $n \geq 2$ et $k \in [\![1, n-1]\!]$, les autres cas étant évidemment triviaux. La démarche qui mène à définir l'encodage proposé dans cette section s'apparente à la démarche qui mène à l'encodage séquentiel proposé par Carsten Sinz dans [12]. En effet, dans l'article précité, l'auteur définit les sommes partielles $s_i = \sum_{m=1}^{i} x_m$ et considère le j-ième bit $s_{i,j}$ de la représentation unaire de s_i. Il transpose alors ces bits en variable booléenne dans un encodage CNF pour aboutir à l'encodage ci-dessous pour la contrainte de cardinalité $\leq k(x_1, ..., x_n)$. Par la suite, on le désigne par le nom d'encodage séquentiel unidirectionnel et on le note $SeqU^n_{\leq k}$.

$$\begin{array}{l}
(\neg x_1 \vee s_{1,1}) \\
(\neg s_{1,j}) \qquad \forall j \in]\!]1, k]\!] \\
\left.\begin{array}{l}
(\neg x_i \vee s_{i,1}) \\
(\neg s_{i-1,1} \vee s_{i,1}) \\
\left.\begin{array}{l}
(\neg x_i \vee \neg s_{i-1,j-1} \vee s_{i,j}) \\
(\neg s_{i-1,j} \vee s_{i,j})
\end{array}\right\} \forall j \in]\!]1, k]\!] \\
(\neg x_i \vee \neg s_{i-1,k}) \\
(\neg x_n \vee \neg s_{n-1,k})
\end{array}\right\} \forall i \in]\!]1, n[\!]
\end{array} \qquad (SeqU^n_{\leq k})$$

Cependant, cette transposition contient une réduction qui aboutit à une perte d'information entre la variable $s_{i,j}$ tel qu'elle est encodée dans $SeqU^n_{\leq k}$ par rapport au bit $s_{i,j}$ décrit précédemment. En effet, le bit $s_{i,j}$ est équivalent à $s_i \geq j$. Or l'encodage $SeqU^n_{\leq k}$ donne $(s_i \geq j) \implies s_{i,j}$ mais pas l'implication réciproque. Cette perte d'information est volontaire car elle entraine un encodage plus restreint de la contrainte $\leq k(x_1, ..., x_n)$. Toutefois, ce choix n'est pas forcément judicieux lorsque l'on considère une contrainte $= k(x_1, ..., x_n)$ ou deux contraintes $\geq k_1(x_1, ..., x_n)$ et $\leq k_2(x_1, ..., x_n)$ définissant un intervalle.

L'encodage CNF qui suit permet d'encoder exactement l'ensemble des équivalences $(s_i \geq j) \iff s_{i,j}$ pour tout $i \in [\![1, n]\!]$ et pour tout $j \in [\![1, k+1]\!]$. On appelle encodage séquentiel bidirectionnel cet encodage et on le note $SeqB^n_{\#k}$.

$$\begin{array}{l}
(x_1 \vee \neg s_{1,1}) \\
(\neg x_i \vee s_{i,1}) \forall i \in [\![1, n]\!] \\
(\neg s_{j-1,j}) \quad \forall j \in]\!]1, k+1]\!] \\
\left.\begin{array}{l}
\left.\begin{array}{l}
(\neg s_{i-1,j} \vee s_{i,j}) \\
(x_i \vee s_{i-1,j} \vee \neg s_{i,j})
\end{array}\right\} \forall j \in [\![1, k+1]\!] \\
\left.\begin{array}{l}
(s_{i-1,j-1} \vee \neg s_{i,j}) \\
(\neg x_i \vee \neg s_{i-1,j-1} \vee s_{i,j})
\end{array}\right\} \forall j \in]\!]1, k+1]\!]
\end{array}\right\} \forall i \in]\!]1, n]\!]
\end{array} \qquad (SeqB^n_{\#k})$$

À partir de cet encodage, il est très facile d'obtenir la contrainte de cardinalité $\leq k(x_1, ..., x_n)$. En effet, il suffit de rajouter la clause $\neg s_{n,k+1}$. De même, la contrainte $\geq k(x_1, ..., x_n)$ s'obtient simplement par le rajout de la clause $s_{n,k}$. Enfin, la contrainte $= k(x_1, ..., x_n)$ s'obtient par le rajout de ces deux clauses. On note $SeqB^n_{\leq k}$, $SeqB^n_{\geq k}$ et $SeqB^n_{=k}$ les encodages respectifs correspondants.

3 Choisir son encodage

3.1 Comparaisons des encodages

En plus des encodages décrits précédemment, on va également considérer l'encodage naïf ($N^n_{\leq k}$) défini par :

$$\bigwedge_{\mathbf{i} \in \mathcal{C}_n^{k+1}} \bigvee_{j=1}^{k+1} \neg x_{i_j} \qquad (N^n_{\leq k})$$

où $\mathbf{i} = (i_1, ..., i_{k+1})$ est une combinaison appartenant à l'ensemble $\mathcal{C}_n^{k+1}$ des combinaisons de $k+1$ éléments de $[\![1, n]\!]$. On rajoute l'encodage proposé par Bailleux & Boufkhad dans [3] ($BB^n_{\geq k_1, \leq k_2}$). On note que les auteurs précités ne donnent pas d'expression explicite de leur encodage mais décrivent plutôt un algorithme permettant de le construire.

D'autres encodages de la littérature [15] ne sont pas considérés car ils ne satisfont pas la condition de performance relative à la propagation unitaire décrite initalement dans [3] (i.e. il ne permettent pas de vérifier la contrainte sur une valuation partielle des variables x_i).

Par ailleurs, par manque de temps, nous n'avons pas intégré ici d'encodage à base de réseaux [1, 2, 7]. En effet, nous souhaitons d'abord vérifier que la génération de tels encodages est bien linéaire en leur nombre total de clauses (ou de littéraux), de manière à ce que la comparaison soit valable. Ce travail reste donc à compléter sur ce point. En effet, de tels encodages peuvent comporter un nombre total de clauses inférieur à ceux des encodages considérés ici pour un certain nombre de valeurs de n et k. Ceci est notamment le cas de l'encodage de la contrainte $\leq k(x_1, ..., x_n)$ proposé par Asín et al. dans [2] dont le nombre total de clauses est égal à :

$$-3m + 6mK + \frac{3}{4}mK\log_2(K) + \frac{3}{4}mK\log_2(K) - 3K - \frac{3}{2}K\log_2(K)$$

avec $K = 2^{\lceil \log_2(k) \rceil}$ et $m = \lceil \frac{n}{K} \rceil$.

Le tableau 1 donne le nombres de clauses (avec le détail en fonction de la taille en nombre de littéraux de ces clauses) ainsi que le nombre de variables auxilaires pour chacun de ces encodages. Les valeurs figurant dans ce tableau ont été recalculées à partir des descriptions de ces encodages dans les articles précités [12, 3] ainsi que le présent article.

Dans la suite de cet article, on utilise les informations de ce tableau pour permettre de choisir l'encodage le mieux adapté aux différents cas étudiés. On fera également usage de la règle suivante pour obtenir un encodage d'une contrainte $\geq k(x_1, ..., x_n)$ en considérant l'encodage pour la contrainte $\leq k(x_1, ..., x_n)$ via l'utilisation de la règle suivante :

$$\geq k(x_1, ..., x_n) \iff \leq (n-k)(\neg x_1, ..., \neg x_n) \qquad (1)$$

De même, on pourra obtenir un encodage de la contrainte $= k(x_1, ..., x_n)$ en combinant différents encodages via l'utilisation de la constation suivante :

$$= k(x_1, ..., x_n) \iff (\leq k(x_1, ..., x_n) \wedge \geq k(x_1, ..., x_n)) \quad (2)$$

Encodage	Nombre de clauses composées de m littéraux		Nombre de variables auxiliaires
$SeqU^n_{\leq k}$	$m = 1$	$k - 1$	$nk - k$
	$m = 2$	$nk + 2n - 2k - 2$	
	$m = 3$	$nk - n - 2k + 2$	
	Total	$2nk + n - 3k - 1$	
$SeqB^n_{\#k}$	$m = 1$	k	$nk + n$
	$m = 2$	$2nk + 2n - 2k$	
	$m = 3$	$2nk + n - 2k - 1$	
	Total	$4nk + 3n - 3k - 1$	
$N^n_{\leq k}$	$m = k + 1$	$\binom{n}{k+1}$	0
$BB^n_{\geq k_1, \leq k_2}$ [2]	$m = 1$	$n - k_2 + k_1$	$n \log_2(n)$
	$m = 3$	$n^2 + 2n \log_2(n) + n - 2$	
	Total	$n^2 + 2n \log_2(n) + 2n - k_2 + k_1 - 2$	

TABLE 1 – Nombre et tailles de clauses pour chaque encodage considéré

3.2 Contrainte $\leq k(x_1, ..., x_n)$

Dans cette secion, on compare différents encodages de $\leq k(x_1, ..., x_n)$.

Conditions sur n et k par encodage			
$SeqU^n_{\leq k}$	$BB^n_{\geq 0, \leq k}$	$SeqB^{\neg n}_{\geq n-k}$	$N^n_{\leq k}$
$n \leq 5$ et $k \in$			
$\emptyset$	$\emptyset$	$\emptyset$	$[\![1, n[\![$
$6 \leq n \leq 8$ et $k \in$			
$[\![1, k_1[\![$	$\emptyset$	$\emptyset$	$[\![k_1, n[\![$
où $k_1 = n - 4$.			
$9 \leq n \leq 13$ et $k \in$			
$[\![1, k_1[\![$	$\emptyset$	$\emptyset$	$[\![k_1, n[\![$
où $k_1 = n - 3$.			
$14 \leq n \leq 30$ et $k \in$			
$[\![1, k_2[\![$	$\emptyset$	$[\![k_2, k_1[\![$	$[\![k_1, n[\![$
où $k_1 = n - 3$ et $k_2 = \lceil \frac{2}{3}(n+1) \rceil$.			
$31 \leq n \leq 36$ et $k \in$			
$[\![1, k_2[\![$	$\emptyset$	$[\![k_2, k_1[\![$	$[\![k_1, n[\![$
où $k_1 = n - 2$ et $k_2 = \lceil \frac{2}{3}(n+1) \rceil$.			
$37 \leq n$ et $k \in$			
$[\![1, k_3[\![$	$[\![k_3, k_2[\![$	$[\![k_2, k_1[\![$	$[\![k_1, n[\![$
où $k_1 = n - 2,\ k_2 = \lceil \frac{3n^2 - 2n \log_2(n) - 2n + 2}{4(n-1)} \rceil$ et $k_3 = \lceil \frac{n^2 + 2n \log_2(n) + n - 1}{2(n-1)} \rceil$.			

TABLE 2 – Encodage de $\leq k(x_1, ..., x_n)$ donnant le nombre total minimal de clauses en fonction de n et k

2. Les valeurs données ici sont exactes si n est une puissance de 2.

On remarque d'abord que pour une telle contrainte, l'encodage $SeqU^n_{\leq k}$ est clairement toujours plus performant que l'encodage $SeqB^n_{\leq k}$ donc on peut exclure ce dernier de notre comparaison. Par contre, en utilisant (1), on peut considérer l'encodage $SeqB^n_{\geq n-k}$ appliqué à $(\neg x_1, ..., \neg x_n)$. On note $SeqB^{\neg n}_{\geq n-k}$ cet encodage.
On cherche donc à comparer les encodages $N^n_{\leq k}$, $SeqU^n_{\leq k}$, $SeqB^{\neg n}_{\geq n-k}$, $BB^n_{\geq 0, \leq k}$. Une analyse complète des différents nombres totaux de clauses pour chacun de ces encodages permet de déterminer l'encodage offrant le plus petit nombre de clauses en fonction de n et k. On a réalisé cette analyse ici et regroupé les résultats dans le tableau 2. Ce tableau décrit, en fin de compte, une partition de l'ensemble des valeurs potentielles de n et k en 4 parties, chacune correspondant aux valeurs de n et k pour lesquelles l'encodage en colonne est optimal (pour le critère du nombre de clauses considéré ici).

Conditions sur n et k par encodage			
$SeqU^n_{\leq k}$	$BB^n_{\geq 0, \leq k}$	$SeqB^{\neg n}_{\geq n-k}$	$N^n_{\leq k}$
$n \leq 5$ et $k \in$			
$\emptyset$	$\emptyset$	$\emptyset$	$[\![1, n[\![$
$6 \leq n \leq 7$ et $k \in$			
$[\![1, k_1[\![$	$\emptyset$	$\emptyset$	$[\![k_1, n[\![$
où $k_1 = n - 3$.			
$8 \leq n \leq 10$ et $k \in$			
$[\![1, k_1[\![$	$\emptyset$	$\emptyset$	$[\![k_1, n[\![$
où $k_1 = n - 2$.			
$11 \leq n \leq 27$ et $k \in$			
$[\![1, k_2[\![$	$\emptyset$	$[\![k_2, k_1[\![$	$[\![k_1, n[\![$
où $k_1 = n - 2$ et $k_2 = \lceil \frac{10n^2 - 3n - 3}{3(5n-6)} \rceil$.			
$28 \leq n \leq 148$ et $k \in$			
$[\![1, k_2[\![$	$\emptyset$	$[\![k_2, k_1[\![$	$[\![k_1, n[\![$
où $k_1 = n - 1$ et $k_2 = \lceil \frac{10n^2 - 3n - 3}{3(5n-6)} \rceil$.			
$149 \leq n$ et $k \in$			
$[\![1, k_3[\![$	$[\![k_3, k_2[\![$	$[\![k_2, k_1[\![$	$[\![k_1, n[\![$
où $k_1 = n - 1,\ k_2 = \lceil \frac{7n^2 - 6n \log_2(n) - 6n + 4}{10(n-1)} \rceil$ et $k_3 = \lceil \frac{3n^2 + 6n \log_2(n) + 3n - 7}{5n - 8} \rceil$.			

TABLE 3 – Encodage de $\leq k(x_1, ..., x_n)$ donnant le nombre total minimal de littéraux en fonction de n et k

Le tableau 2 permet de choisir un encodage en fonction du nombre de clauses de cet encodage mais il ne prend pas en compte la taille de ces clauses (i.e. le nombre de littéraux par clause). Or la taille des clauses dans une contrainte CNF peut avoir une influence importante sur la rapidité d'un solveur SAT sur cette contrainte. Comme l'évaluation d'une valuation d'une contrainte CNF est linéaire par rapport à son nombre total de littéraux, on pourrait également considérer le nombre total de littéraux de chacun de encodages plutôt que leur nombre de clauses comme critère pour choisir un encodage. Même si le nombre total de clauses est généralement utilisé comme critère de comparaison des encodages dans l'état de l'art précité, on penche plutôt pour l'utilisation du nombre total de littéraux. En

©AfIA 2018

tout état de cause, on présentera systèmatiquement par la suite les valeurs obtenues pour chacun de ces deux critères. Le tableau 3 permet de déterminer l'encodage avec le plus petit nombre total de littéraux.

3.3 Contrainte $\geq k(x_1, ..., x_n)$

Conditions sur n et k par encodage			
$N^{\neg n}_{\leq n-k}$	$SeqB^n_{\geq k}$	$BB^n_{\geq k,\leq n}$	$SeqU^{\neg n}_{\leq n-k}$
$n \leq 5$ et $k \in$			
$[\![1, n[\![$	$\emptyset$	$\emptyset$	$\emptyset$
$6 \leq n \leq 8$ et $k \in$			
$[\![1, 5[\![$	$\emptyset$	$\emptyset$	$[\![5, n[\![$
$9 \leq n \leq 13$ et $k \in$			
$[\![1, 4[\![$	$\emptyset$	$\emptyset$	$[\![4, n[\![$
$14 \leq n \leq 30$ et $k \in$			
$[\![1, 4[\![$	$[\![4, k_1[\![$	$\emptyset$	$[\![k_1, n[\![$
où $k_1 = \lceil \frac{n-1}{3} \rceil$.			
$31 \leq n \leq 36$ et $k \in$			
$[\![1, 3[\![$	$[\![3, k_1[\![$	$\emptyset$	$[\![k_1, n[\![$
où $k_1 = \lceil \frac{n-1}{3} \rceil$.			
$37 \leq n$ et $k \in$			
$[\![1, 3[\![$	$[\![3, k_2[\![$	$[\![k_2, k_1[\![$	$[\![k_1, n[\![$
où $k_1 = \left\lceil \frac{n^2-2n\log_2(n)-n+1}{2(n-1)} \right\rceil$ et $k_2 = \left\lceil \frac{n^2+2n\log_2(n)-2n-2}{4(n-1)} \right\rceil$.			

TABLE 4 – Encodage de $\geq k(x_1, ..., x_n)$ donnant le nombre total minimal de clauses en fonction de n et k

Conditions sur n et k par encodage			
$N^{\neg n}_{\leq n-k}$	$SeqB^n_{\geq k}$	$BB^n_{\geq k,\leq n}$	$SeqU^{\neg n}_{\leq n-k}$
$n \leq 5$ et $k \in$			
$[\![1, n[\![$	$\emptyset$	$\emptyset$	$\emptyset$
$n = 6$ et $k \in$			
$\{1, 2, 3, 5\}$	$\emptyset$	$\emptyset$	$k = 4$
$n = 7$ et $k \in$			
$[\![1, 4[\![$	$\emptyset$	$\emptyset$	$[\![4, n[\![$
$8 \leq n \leq 10$ et $k \in$			
$[\![1, 3[\![$	$\emptyset$	$\emptyset$	$[\![3, n[\![$
$11 \leq n \leq 27$ et $k \in$			
$[\![1, 3[\![$	$[\![3, k_1[\![$	$\emptyset$	$[\![k_1, n[\![$
où $k_1 = \left\lceil \frac{5n^2-15n+3}{3(5n-6)} \right\rceil$.			
$28 \leq n \leq 148$ et $k \in$			
$[\![1, 2[\![$	$[\![2, k_1[\![$	$\emptyset$	$[\![k_1, n[\![$
où $k_1 = \left\lceil \frac{5n^2-15n+3}{3(5n-6)} \right\rceil$.			
$149 \leq n$ et $k \in$			
$[\![1, 2[\![$	$[\![2, k_2[\![$	$[\![k_2, k_1[\![$	$[\![k_1, n[\![$
où $k_1 = \left\lceil \frac{2n^2-6n\log_2(n)-11n+7}{5n-8} \right\rceil$ et $k_2 = \left\lceil \frac{3n^2+6n\log_2(n)-4n-4}{10(n-1)} \right\rceil$.			

TABLE 5 – Encodage de $\geq k(x_1, ..., x_n)$ donnant le nombre total minimal de littéraux en fonction de n et k

Dans cette section, on compare 5 encodages pour la contrainte $\geq k(x_1, ..., x_n)$. Par l'équivalence (1), on obtient bien la contrainte souhaitée en appliquant les encodages $N^n_{\leq n-k}$, $SeqU^n_{\leq n-k}$ et $SeqB^n_{\leq n-k}$ à $(\neg x_1, ..., \neg x_n)$. On note respectivement $N^{\neg n}_{\leq n-k}$, $SeqU^{\neg n}_{\leq n-k}$ et $SeqB^{\neg n}_{\leq n-k}$ ces trois encodages. On considère par ailleurs les deux encodages $SeqB^n_{\geq k}$ et $BB^n_{\geq k,\leq n}$.
Comme précédemment, on détermine les encodages donnant le nombre total minimal de clauses (tableau 4) ainsi que les encodages donnant le nombre total minimal de littéraux (5) en fonction de n et k. On remarque que, dans ce cas, c'est l'encodage $SeqB^{\neg n}_{\leq n-k}$ qui est écarté systématiquement.

3.4 Contrainte $= k(x_1, ..., x_n)$

Dans cette section, on compare 7 encodages différents de $= k(x_1, ..., x_n)$ parmi lesquels 2 sont des encodages mixtes entre deux encodages différents. Il s'agit des encodages :

1. $BB^n_{=k}$ (i.e. $BB^n_{\geq k,\leq k}$) ;
2. $SeqB^n_{=k}$;
3. $SeqB^{\neg n}_{=n-k}$;
4. $SeqU^n_{=k}$ (i.e. $SeqU^n_{\leq k}$ et $SeqU^{\neg n}_{\leq n-k}$) ;
5. $N^n_{=k}$ (i.e. $N^n_{\leq k}$ et $N^{\neg n}_{\leq n-k}$) ;
6. $NS^n_{=k}$ (i.e. $N^n_{\leq k}$ et $SeqU^{\neg n}_{\leq n-k}$) ;
7. $SN^n_{=k}$ (i.e. $SeqU^n_{\leq k}$ et $N^{\neg n}_{\leq n-k}$).

Comme dans les deux sections précédentes, on a determiné les encodages donnant le nombre total de clauses minimal en fonction de n et k ainsi que celui donnant le nombre total de littéraux minimal en fonction des mêmes paramètres. Chacun de ces résultats étant difficilement synthétisable en un seul tableau, on renvoie aux annexes pour le détail pour tous les entiers $n \in [\![6, 17]\!]$.

Encodage	Condition sur k	Proportion en $n \to +\infty$
$SN^n_{=k}$	$1 \leq k < 3$	0%
$SeqB^n_{=k}$	$3 \leq k < k_2$	25%
$BB^n_{=k}$	$k_2 \leq k < k_1$	50%
$SeqB^{\neg n}_{=n-k}$	$k_1 \leq k < n-2$	25%
$NS^n_{=k}$	$n-2 \leq k < n$	0%
où $k_1 = \left\lceil \frac{3n^2-2n\log_2(n)-2n+3}{4n-3} \right\rceil$ et $k_2 = \left\lceil \frac{n^2+2n\log_2(n)-n-3}{4n-3} \right\rceil$.		

TABLE 6 – Encodage de $= k(x_1, ..., x_n)$ donnant le nombre total minimal de clauses en fonction de k pour $n \geq 18$

Pour $n \leq 5$, c'est l'encodage naïf $N^n_{=k}$ qui donne les plus petits nombres de clauses et de littéraux quel que soit k. Les deux tableaux 6 et 7 donnent les résultats pour $n \geq 18$. On remarquera que les encodages $SeqU^n_{=k}$ et $N^n_{=k}$ en sont absents. Par ailleurs, on a indiqué, pour chaque encodage, la limite de la proportion à n donné de valeurs différentes de k pour laquelle cet encodage est optimal lorsque n tend vers l'infini.

Encodage	Condition sur k	Proportion en $n \to +\infty$
$SN^n_{=k}$	$k = 1$	0%
$SeqB^n_{=k}$	$2 \leq k < k_2$	30%
$BB^n_{=k}$	$k_2 \leq k < k_1$	40%
$SeqB^{\neg n}_{=n-k}$	$k_1 \leq k < n-1$	30%
$NS^n_{=k}$	$k = n-1$	0%
où $k_1 = \left\lceil \frac{7n^2-6n\log_2(n)-6n+1}{10n-9} \right\rceil$ et $k_2 = \left\lceil \frac{3n^2+6n\log_2(n)-3n-1}{10n-9} \right\rceil$.		

TABLE 7 – Encodage de $= k(x_1, ..., x_n)$ donnant le nombre total minimal de littéraux en fonction de k pour $n \geq 18$

3.5 Contrainte $\geq k_1, \leq k_2(x_1, ..., x_n)$

On étudie ici le cas d'une contrainte donnant un intervalle. On compare 7 encodages de cette contrainte, correspondant aux cas généraux des encodages de $= k(x_1, ..., x_n)$ de la section précédente. Il s'agit des encodages :

1. $BB^n_{\geq k_1, \leq k_2}$;
2. $SeqB^n_{\geq k_1, \leq k_2}$;
3. $SeqB^{\neg n}_{\geq n-k_2, \leq n-k_1}$;
4. $SeqU^n_{\geq k_1, \leq k_2}$ (i.e. $SeqU^n_{\leq k_2}$ et $SeqU^{\neg n}_{\leq n-k_1}$) ;
5. $N^n_{\geq k_1, \leq k_2}$ (i.e. $N^n_{\leq k_2}$ et $N^{\neg n}_{\leq n-k_1}$) ;
6. $SN^n_{\geq k_1, \leq k_2}$ (i.e. $SeqU^n_{\leq k_2}$ et $N^{\neg n}_{\leq n-k_1}$) ;
7. $NS^n_{\geq k_1, \leq k_2}$ (i.e. $N^n_{\leq k_2}$ et $SeqU^{\neg n}_{\leq n-k_1}$).

On ne fait pas ici une comparaison exhaustive de toutes les valeurs possibles de k et n comme on a pu le faire pour les contraintes précédentes car cela nous semble trop fastidieux et difficilement exposable de manière concise. On donne simplement les expressions des nombres totaux de clauses et nombres totaux de littéraux pour chacun de ces encodages. Ces expressions peuvent être évaluées au cas par cas pour déterminer une valeur minimale et choisir l'encodage associé lorsque n est faible.

Encodage	Nombre total de clauses
$BB^n_{\geq k_1, \leq k_2}$	$n^2 + 2n\log_2(n) + 2n - k_2 + k_1 - 2$
$SeqB^n_{\geq k_1, \leq k_2}$	$4nk_2 + 3n - 3k_2 + 1$
$SeqB^{\neg n}_{\geq n-k_2, \leq n-k_1}$	$4n^2 - 4nk_1 + 3k_1 + 1$
$SeqU^n_{\geq k_1, \leq k_2}$	$2n^2 + 2n(k_2 - k_1) - n - 3(k_2 - k_1) - 2$
$N^n_{\geq k_1, \leq k_2}$	$\binom{n}{k_2+1} + \binom{n}{k_1-1}$
$SN^n_{\geq k_1, \leq k_2}$	$\binom{n}{k_1-1} + 2nk_2 + n - 3k_2 - 1$
$NS^n_{\geq k_1, \leq k_2}$	$\binom{n}{k_2+1} + 2n^2 - 2nk_1 - 2n + 3k_1 - 1$

TABLE 8 – Nombre total de clauses pour les encodages de la contrainte $\geq k_1, \leq k_2(x_1, ..., x_n)$

Sinon, en dehors de certains effets de bord quand $k_1 \leq 2$ ou bien $k_2 \geq n-2$, on peut voir que, pour n suffisamment grand, ce sont les trois premiers encodages qui sont les plus performants. On peut alors utiliser le protocole suivant pour choisir un encodage :

Encodage	Nombre total de littéraux
$BB^n_{\geq k_1, \leq k_2}$	$3n^2 + 6n\log_2(n) + 4n - k_2 + k_1 - 6$
$SeqB^n_{\geq k_1, \leq k_2}$	$10nk_2 + 7n - 9k_2 - 1$
$SeqB^{\neg n}_{\geq n-k_2, \leq n-k_1}$	$10n^2 - 10nk_1 - 2n + 9k_1 - 1$
$SeqU^n_{\geq k_1, \leq k_2}$	$5n^2 + 5n(k_2 - k_1) - 7n - 9(k_2 - k_1) + 2$
$N^n_{\geq k_1, \leq k_2}$	$(n - k_2)\binom{n}{k_2} + k_1\binom{n}{k_1}$
$SN^n_{\geq k_1, \leq k_2}$	$k_1\binom{n}{k_1} + 5nk_2 + n - 9k_2 + 1$
$NS^n_{\geq k_1, \leq k_2}$	$(n - k_2)\binom{n}{k_2} + 5n^2 - 5nk_1 - 8n + 9k_1 + 1$

TABLE 9 – Nombre total de littéraux pour les encodages de la contrainte $\geq k_1, \leq k_2(x_1, ..., x_n)$

$$\begin{array}{lll} \textbf{Si} & \min(k_2, n-k_1) \geq \frac{3}{10}n & \textbf{choisir } BB^n_{\geq k_1, \leq k_2} \\ \textbf{Sinon si} & k_2 \leq \min\left(\frac{3}{10}n, n-k_1\right) & \textbf{choisir } SeqB^n_{\geq k_1, \leq k_2} \\ \textbf{Sinon} & & \textbf{choisir } SeqB^{\neg n}_{\geq n-k_2, \leq n-k_1} \end{array} \tag{3}$$

Le protocole ci-dessus utilise le critère du nombre total de littéraux minimal. Pour utiliser le critère du nombre total de clauses, il suffit de remplacer $\frac{3}{10}$ par $\frac{1}{4}$.

3.6 Contrainte $\in \mathcal{K}(x_1, ..., x_n)$

Dans cette section, on considère un sous-ensemble $\mathcal{K} = \{k_1, ..., k_m\} \subset [\![0, n]\!]$ avec $m \geq 2$ et $k_1 < k_2 < ... < k_m$. Le cas général d'une contrainte d'appartenance du cardinal à un sous-ensemble $\mathcal{K}$ se distingue entièrement des cas précédents. En effet, la proposition $k \in \mathcal{K}$ est équivalente à la disjonction $\bigvee_{i=1}^{m} [k = k_i]$. Ainsi, à part le cas particulier d'un intervalle traité précédemment, on ne peut pas se ramener directement à une conjonction des encodages CNF décrits précédemment. Or obtenir une contrainte CNF à partir d'une disjonction de contraintes CNF nécessite une opération supplémentaire. Effectuer cette opération est prohibitif si l'on considère directement des disjonctions des encodages précédents. Toutefois, on peut également considérer ces disjonctions au sein d'un encodage séquentiel bidirectionnel unique ce qui permet de rendre cette opération tout à fait envisageable.

En effet, la contrainte suivante est bien un encodage de la contrainte $\in \mathcal{K}(x_1, ..., x_n)$.

$$SeqB^n_{\#} \wedge \left(\bigvee_{i=1}^{m} (s_{n,k_i} \wedge \neg s_{n,k_i+1}) \right) \tag{4}$$

Cette encodage n'est pas une contrainte CNF mais on peut se ramener à une contrainte CNF équisatisfaisable facilement en introduisant m variables supplémentaires y_0, ..., y_{m-1} [8]. On définit ainsi l'encodage $SeqB^n_{\in\mathcal{K}}$ de la contrainte $\in \mathcal{K}(x_1, ..., x_n)$ comme étant la contrainte CNF ci-dessous :

©AfIA 2018

$$
\begin{aligned}
SeqB^n_{\#} \quad & \wedge \quad \bigwedge_{i=1}^{m} \left(s_{n,k_i} \vee \neg y_{i-1} \vee \left(\bigvee_{j=i}^{m-1} y_j \right) \right) \\
& \wedge \quad \bigwedge_{i=1}^{m} \left(\neg s_{n,k_i+1} \vee \neg y_{i-1} \vee \left(\bigvee_{j=i}^{m-1} y_j \right) \right) \\
& \wedge \quad y_0
\end{aligned}
\qquad (SeqB^n_{\in \mathcal{K}})
$$

Cette contrainte a un nombre total de clauses égal à :

$$4nk_m + 3n - 3k_m + 2m \tag{5}$$

et un nombre total de littéraux égal à :

$$10nk_m + 7n - 9k_m + m^2 + 3m - 2 \tag{6}$$

Enfin, on peut remarquer que l'encodage $SeqB^n_{\in \overline{\mathcal{K}}}$ (où $\overline{\mathcal{K}}$ est le complèmentaire de $\mathcal{K}$) appliqué à $(\neg x_1, ..., \neg x_n)$ donne également la contrainte $\in \mathcal{K}(x_1, ..., x_n)$. On note $SeqB^{\neg n}_{\in \overline{\mathcal{K}}}$ cet encodage. Il peut éventuellement être plus performant que l'encodage précédent selon $\mathcal{K}$. Son nombre total de clauses et son nombre total de littéraux s'obtiennent en remplaçant k_m par $\max(\overline{\mathcal{K}})$ et m par $n - m$ dans (5) et (6).

Le pire des cas est atteint pour une contrainte de type k pair. On a alors un nombre de clauses de l'ordre de $4n^2$ et un nombre de littéraux de l'ordre de $10.5n^2$ ce qui est tout à fait raisonnable. Par ailleurs, à notre connaissance, il s'agit du seul encodage CNF de la littérature de la contrainte $\in \mathcal{K}(x_1, ..., x_n)$.

4 Conclusion

Lorsque l'on recherche un algorithme optimisé pour un problème donné, il est rare de découvrir un seul algorithme qui soit le plus optimisé pour chacune des occurrences de ce problème. Pour le problème d'une résolution CNF-SAT d'une contrainte de cardinalité, on peut remarquer que chacun des encodages d'une contrainte de cardinalité considérés dans cet article est préférable aux autres pour au moins quelques valeurs différentes de k et n. Par ailleurs, les comportements des différentes solutions à l'infini ne préjugent en rien quant à leurs comportements en faible dimension. Par exemple dans le cas des contraintes $\leq k(x_1, ..., x_n)$ et $\geq k(x_1, ..., x_n)$, l'encodage proposé par Bailleux et Boufkhad est préférable (selon le critère du nombre total de littéraux) dans environ 10% des cas lorsque n est très grand mais n'est préférable en aucun cas lorsque $n < 149$. Afin de pouvoir déterminer l'encodage réellement le mieux adapté à une contrainte et des paramètres donnés, il est nécessaire de passer par une étude exhaustive des cas comme cela a été réalisé dans cet article.

On pourrait objecter que, si n est petit alors la taille de l'encodage le sera également et qu'en conséquence, même si on reste loin de l'optimum possible, on pourra négliger l'accroissement. Cet argument ne tient toutefois pas longtemps car un gain fixe sur une opération peut être non négligeable, d'autant plus si cette opération est répétée de nombreuses fois au cours d'un même processus.

Par ailleurs, le nouvel encodage séquentiel bidirectionnel qui est proposé dans cet article offre une performance nettement améliorée dans certains cas ainsi que des perspectives nouvelles.

En effet, dans le cas d'une contrainte de type $= k(x_1, ..., x_n)$, les nombres totaux de clauses et de littéraux sont au mieux quadratiques en n pour les autres encodages et ceci quel que soit k. Toutefois, dans un certain nombre de problèmes, k est fixé indépendamment de n. Les nombres totaux de clauses et de littéraux de l'encodage séquentiel bidirectionnel sont alors linéaires en n. Cet encodage a ainsi pu être mis à profit dans le cas de la résolution SAT d'un problème d'emploi du temps en BTS [6] qui comporte un certain nombre de contraintes de type $= 3(x_1, ..., x_n)$.

Cet encodage offre également de nouvelles perspectives via son application aux contraintes de type $\in \mathcal{K}(x_1, ..., x_n)$ pour lesquelles il n'existait pas avant, à notre connaissance, d'encodage performant.

Enfin, on peut noter que sa présentation explicite (semblable à celle de l'encodage séquentiel dans [12] et qui a sûrement contribué à sa popularité) en permet une implémentation directe.

Le travail de recherche présenté dans cet article est d'ordre théorique. Si les critères du nombre total de clauses ou du nombre total de littéraux sont généralement pertinents, ils ne suffisent pas à déterminer l'encodage le mieux adapté qui pourra dépendre non seulement du problème considéré mais également du solveur utilisé. Il appelle donc d'autres travaux qui permettront de mettre en pratique (tels que [6]) et de valider ces résultats. Il sera également prolongé de manière à intégrer au processus de choix des encodages à base de réseaux dont l'encodage de $\leq k(x_1, ..., x_n)$ défini dans [2].

5 Annexe

k \ n	6	7	8	9	10	11	12	13	14	15	16	17
1	6	6	6	6	6	6	6	6	6	6	6	6
2	5	6	6	6	6	6	6	6	6	6	6	6
3	5	5	6	6	6	6	6	6	6	6	6	6
4	5	5	5	4	2	2	2	2	2	2	2	2
5	7	7	7	4	1	1	1	2	2	2	2	2
6		7	7	7	3	1	1	1	1	1	1	2
7			7	7	7	3	1	1	1	1	1	1
8				7	7	7	3	3	1	1	1	1
9					7	7	7	3	3	1	1	1
10						7	7	7	3	3	1	1
11							7	7	7	3	3	3
12								7	7	7	3	3
13									7	7	7	3
14										7	7	7
15											7	7
16												7

TABLE 10 – Encodage de la contrainte $= k(x_1, ..., x_n)$ donnant le nombre total minimal de clauses pour $n \in [\![6, 17]\!]$

Les tableaux 10 et 11 donnent, en fonction de k et n, les encodages de $= k(x_1, ..., x_n)$ pour $n \in [\![6, 17]\!]$ ayant un

k \ n	6	7	8	9	10	11	12	13	14	15	16	17
1	6	6	6	6	6	6	6	6	6	6	6	6
2	6	6	6	6	6	6	6	6	6	6	6	2
3	5	6	4	2	2	2	2	2	2	2	2	2
4	7	7	4	4	4	2	2	2	2	2	2	2
5	5	7	4	4	4	4	2	2	2	2	2	2
6		7	7	3	4	4	4	4	2	2	2	2
7			7	7	3	3	3	4	4	4	2	2
8				7	7	3	3	3	3	4	4	4
9					7	7	3	3	3	3	3	4
10						7	7	3	3	3	3	3
11							7	7	3	3	3	3
12								7	7	3	3	3
13									7	7	3	3
14										7	3	3
15											7	3
16												7

TABLE 11 – Encodage de la contrainte $= k(x_1, ..., x_n)$ donnant le nombre total minimal de littéraux pour $n \in [\![6, 17]\!]$

nombre total minimal de clauses et de littéraux respectivement. Les encodages sont indiqués dans le tableau par un numéro qui correspond à leur numérotation dans la section 3.4.

Références

[1] R. Asín, R. Nieuwenhuis, A. Oliveras, and E. Rodríguez-Carbonell. Cardinality networks and their applications. In *International Conference on Theory and Applications of Satisfiability Testing*, pages 167–180. Springer, 2009.

[2] R. Asín, R. Nieuwenhuis, A. Oliveras, and E. Rodríguez-Carbonell. Cardinality networks : a theoretical and empirical study. *Constraints*, 16(2) :195–221, 2011.

[3] O. Bailleux and Y. Boufkhad. Efficient cnf encoding of boolean cardinality constraints. In *International conference on principles and practice of constraint programming*, pages 108–122. Springer, 2003.

[4] O. Bailleux, Y. Boufkhad, and O. Roussel. A translation of pseudo-boolean constraints to sat. *Journal on Satisfiability, Boolean Modeling and Computation*, 2 :191–200, 2006.

[5] O. Bailleux, Y. Boufkhad, and O. Roussel. New encodings of pseudo-boolean constraints into cnf. In *International Conference on Theory and Applications of Satisfiability Testing*, pages 181–194. Springer, 2009.

[6] T. Delacroix. Planifier l'épreuve e5 à l'aide d'un solveur sat. In *APIA, Conférence Nationale sur les Applications Pratiques de l'Intelligence Artificielle*, 2018.

[7] N. Eén and N. Sorensson. Translating pseudo-boolean constraints into sat. *Journal on Satisfiability, Boolean Modeling and Computation*, 2 :1–26, 2006.

[8] ENS. Concours d'admission - composition d'informatique - a -, 2016.

[9] A. M. Frisch and P. A. Giannaros. Sat encodings of the at-most-k constraint. some old, some new, some fast, some slow. In *Proc. of the Tenth Int. Workshop of Constraint Modelling and Reformulation*, 2010.

[10] J. Marques-Silva and I. Lynce. Towards robust cnf encodings of cardinality constraints. *Principles and Practice of Constraint Programming–CP 2007*, pages 483–497, 2007.

[11] MiniZinc. Minizinc challenge 2017 results, 2017.

[12] C. Sinz. Towards an optimal cnf encoding of boolean cardinality constraints. *CP*, 3709 :827–831, 2005.

[13] P. J. Stuckey, T. Feydy, A. Schutt, G. Tack, and J. Fischer. The minizinc challenge 2008–2013. *AI Magazine*, 35(2) :55–60, 2014.

[14] N. Tamura, M. Banbara, and T. Soh. Compiling pseudo-boolean constraints to sat with order encoding. In *Tools with Artificial Intelligence (ICTAI), 2013 IEEE 25th International Conference on*, pages 1020–1027. IEEE, 2013.

[15] J. P. Warners. A linear-time transformation of linear inequalities into conjunctive normal form. *Information Processing Letters*, 68(2) :63–69, 1998.

©AfIA 2018

Inférence bayésienne pour l'estimation de déformations larges par champs gaussien: application au recalage d'images multi-modales

Thomas Deregnaucourt[1] Chafik Samir[1] Anne-Françoise Yao[2]
[1]LIMOS, CNRS UMR 6620, Université Clermont Auvergne, France
[2]LMBP, CNRS UMR 6158, Université Clermont Auvergne, France

thomas.deregnaucourt@uca.fr - chafik.samir@uca.fr - anne.yao@uca.fr

Résumé

Le problème de recalage d'images consiste à estimer la déformation globale entre une image source I_1 et une image cible I_2. Dans ce contexte, nous nous intéressons à l'estimation d'un champ de déformation U sur le domaine $\Omega = [0,1]^2$ de I_1 sachant U sur un ensemble fini de courbes $\beta \in \Omega$. Pour ce faire, nous proposons une nouvelle méthode basée sur des modèles gaussiens pour recaler des images multimodales. La méthode proposée commence par résoudre le problème de correspondance entre les courbes βs puis estime le déplacement sur tout Ω. La solution optimale est calculée à l'aide du maximum de vraisemblance et de l'inférence bayésienne. D'après les résultats obtenus sur des données réelles et simulées, la déformation résultante a l'avantage d'être exacte sur les observations et d'être lisse sur Ω.

Mots Clef

Recalage d'images, Statistique spatiale, Processus gaussiens, Inférence bayésienne.

Abstract

Image registration aims to estimate the global deformation between a target image I_1 and a reference image I_2. In this context, we will focus on estimating a random field U on the I_1 domain $\Omega = [0,1]^2$ based on observations of U on a finite set of curves $\beta \in \Omega$. Indeed, we present a new multimodal image registration method based on Gaussian random fields. The proposed method first find the optimal correspondences between curves βs then estimate the deformation vector field on Ω. The optimal solution is computed using Maximum Likelihood and Bayesian inference. Based on results using both real and simulated data, the resulting deformation has the advantage of being exact on the observations as being sufficiently smooth over the whole Ω.

Keywords

Image registration, Spatial statistics, Gaussian process, Bayesian inference.

1 Introduction

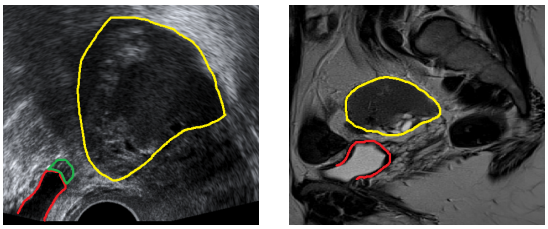
Le recalage d'images est une méthode qui vise à estimer la transformation, soumise à certaines contraintes, d'une image source I_1 vers une image cible I_2, afin de fusionner leurs informations complémentaires. Cette méthode est utilisée dans de nombreux domaines d'applications [11, 15, 2]. En imagerie médicale on utilise le recalage d'images pour détecter des maladies, valider un traitement, comparer les données du patient avec des atlas anatomiques, etc. [12]. L'estimation de cette déformation est basée soit sur les intensités, soit sur des caractéristiques géométriques, soit sur les deux [12, 10]. Dans le premier cas, on cherche une transformation concervant la correspondance entre les niveaux de gris, et dans le second la correspondance entre des points, des courbes, etc.

Pour cet article, nous nous sommes intéressés au problème de l'endométriose. Cette maladie est provoquée par l'apparition de muqueuse utérine, aussi appelé endomètre, en dehors de la cavité utérine. Elle touche approximativement 10% des femmes en âge de procréer, et peut provoquer divers symptômes tels que des douleurs pelviennes chroniques, une dysenterie sévère, une infertilité, etc. [3]. Il est alors nécessaire de pouvoir détecter si une patiente à l'endométriose afin de la traiter efficacement, que ce soit par des antalgiques, des traitements hormonaux, ou par chirurgie dans les cas les plus sévères. L'endomètre pouvant pénétrer d'autres tissus et organes, les méthodes de détection de l'endométriose se basent sur plusieurs modalités d'images, donnant des informations complémentaires. Plus précisément, l'échographie permet d'avoir une estimation de l'infiltration de l'endomètre dans d'autres tissus, et l'imagerie par résonance magnétique (IRM) une position précise des kystes [3]. La fusion des données IRM/échographie permet alors d'avoir un diagnostic précis, mais nécessite un recalage entre ces deux modalités.

Dans notre cas, les deux images I_1 et I_2 représentent respectivement l'échographie et l'IRM d'un même organe. Cependant, comme le montre la Figure 1, les modalités d'images ont des distributions d'intensités différentes, ce qui rend inefficace les méthodes de recalage basées sur les intensités. D'autre part, un spécialiste peut extraire le con-

 ©AfIA 2018

tour des organes présents dans les deux images. Nous allons alors utiliser ces dernières pour effectuer le recalage. Pour ce faire, nous estimerons le champ de déformation entre les deux images, à l'aide des champs gaussiens.



(a) Echographie (b) IRM

Figure 1: Exemple de courbes extraites manuellement issues d'images multimodales représentant les mêmes organes (en rouge et en jaune respectivement) : échographie à gauche et IRM à droite.

Le reste du papier est organisé de la manière suivante. Dans la section 2 nous formaliserons notre problème, et expliquerons en détail notre méthode de résolution. Puis nous présenterons nos résultats, sur des données synthétiques et réelles, dans la section 3. Enfin, la section 4 conclura ce papier.

2 Méthodologie

2.1 Formulation du problème

Soient Ω un domaine fermé de $\mathbb{R}^2$, $\beta_1 \subset I_1$ les courbes de l'échographie et $\beta_2 \subset I_2$ celles de l'IRM. On cherche à estimer un champ de déformation Ψ sur Ω transformant I_1 en I_2. Ce champ doit déformer β_1 en β_2, tout en étant lisse. Pour résoudre ce problème, il nous faut:

1. Trouver une correspondance optimale entre β_1 et β_2.
2. Estimer une déformation Ψ induite par un champ de déplacement U, c'est-à-dire telle que:

$$\begin{aligned} \Psi : \Omega &\rightarrow \Omega \\ X &\mapsto \Psi(X) = X + U(X) \end{aligned}$$

et vérifiant la contrainte $\Psi(\beta_1) = \beta_2$

Nous allons tout d'abord nous intéresser au premier problème.

2.2 Correspondance optimale entre courbes

Afin de trouver une correspondance optimale entre les courbes, nous adaptons les travaux de Srivastava et al. [13]. Dans ce papier, les auteurs s'intéressaient à l'analyse des formes, et cherchaient une invariance aux transformations préservant la forme, c'est-à-dire à la translation, la rotation, la mise à l'échelle et la re-paramétrisation. Dans notre cas, la translation et la rotation sont déjà fixées pour toute l'image, et la mise à l'échelle n'est pas une nuisance. Ainsi, nous cherchons seulement l'invariance à la re-paramétrisation. Par brièveté, nous ne décrivons le processus que pour les courbes ouvertes, mais celui-ci peut être étendu simplement à des courbes fermés [7].
Soit $\beta : [0,1] \rightarrow \mathbb{R}^2$ une courbe ouverte paramétrisée. On utilise par la suite la représentation square-root velocity function (SRVF) q de β, défini par:

$$\begin{aligned} q : [0,1] &\rightarrow \mathbb{R}^2 \\ t &\mapsto \frac{\dot{\beta}(t)}{\sqrt{||\dot{\beta}(t)||_2}} \end{aligned}$$

Le mapping $\beta \Longleftrightarrow (\beta(0), q)$ étant une bijection, il est possible de revenir aux courbes originales en stockant le premier point de ces dernières. On note $\mathcal{C}$ l'espace de SRVFs:

$$\mathcal{C} = \left\{ q \in \mathbb{L}^2([0,1], \mathbb{R}^2) \mid \int_0^1 ||q(t)||_2^2 dt = 1 \right\}$$

Comme on recherche une représentation des courbes invariantes aux re-paramétrisations, nous allons utiliser des classes d'équivalence. Nous définissons d'abord le groupe des re-paramétrisations Γ:

$$\Gamma = \{\gamma : [0,1] \rightarrow [0,1] \mid \gamma(0) = 0, \gamma(1) = 1, 0 < \dot{\gamma} < \infty\}$$

La re-paramétrisation d'une courbe β par $\gamma \in \Gamma$ est donnée par $\beta \circ \gamma$, et la SRVF de cette courbe re-paramétrisée est alors $(q \circ \gamma)\sqrt{\dot{\gamma}}$. Ainsi, pour unifier tous les éléments de $\mathcal{C}$ représentant la même courbe, on défini nos classes d'équivalence par $[q] = \{(q \circ \gamma)\sqrt{\dot{\gamma}} \mid \gamma \in \Gamma\}$. On note $\mathcal{S} = \mathcal{C}/\Gamma = \{[q], q \in \mathcal{C}\}$ l'ensemble des classes d'équivalence. Afin de comparer deux courbes, on impose la métrique $\mathbb{L}^2$ à $\mathcal{S}$. Sous la représentation SRVF, la métrique $\mathbb{L}^2$ correspond à une métrique élastique sur l'espace original des courbes [8], ce qui permet de déformer les courbes pour avoir la correspondance optimale. Cette déformation optimale entre deux points sur $\mathcal{S}$ est obtenue par le chemin géodésique, et la distance entre elles est définie par la longueur du chemin.
Afin de voir comment ceci peut résoudre notre problème, on note q_1 et q_2 la représentation SRVF respective des deux courbes β_1 et β_2. Afin de calculer la géodésique entre leurs classes d'équivalence $[q_1]$ et $[q_2]$, on fixe q_1, et on cherche la re-paramétrisation optimale de q_2 en résolvant $\hat{\gamma} = \arg\inf_{\gamma \in \Gamma} ||q_1 - (q_2 \circ \gamma)\sqrt{\dot{\gamma}}||_2^2$. La re-paramétrisation $\hat{\gamma}$ donnera alors la correspondance optimale entre les courbes.
Par la suite, on note $\{X_i, i = 1, \cdots, N\}$ l'ensemble de N points représentant la discrétisation de β_1 et $\{U_i = U(X_i), i = 1, \cdots, N\}$ les déplacements correspondants donnés par $\beta_2 \circ \hat{\gamma}$. Afin de résoudre le second problème, nous supposons que U est un champ gaussien.

2.3 Champ de déformation gaussien

On rappelle qu'un champ gaussien U est défini par:

$$U(X) \sim \mathcal{N}(\mu(X), C(X))$$

où $\mu(X)$ et $C(X)$ sont respectivement la moyenne et la variance de $U(X)$. Une loi gaussienne étant entièrement décrite par sa moyenne et sa variance, il nous suffit de trouver ces derniers pour définir notre champ. Pour ce faire, nous supposons tout d'abord que U est un champ stationnaire, c'est-à-dire que $\mu(X) = \mu$, $\forall X \in \Omega$. On a alors:

$$\mathcal{N}(\mu(X), C(X)) = \mu + \mathcal{N}(0, C(X))$$

ce qui implique que μ est une translation sur l'image I_1. Nous pouvons alors supposer que $\mu = 0$. Nous estimons ensuite $C(X)$ par une méthode paramétrique. Il existe un grand choix de fonctions de covariance candidates dans la littérature [14, 1]. Dans ce travail, nous avons choisi C comme fonction de covariance de Matérn :

$$C(h) = \tau \frac{2^{1-\nu}}{\Gamma(\nu)} (h\alpha)^{\nu} K_{\nu}(h\alpha) \tag{1}$$

où h est la corrélation spatiale et K la fonction de Bessel modifiée de seconde espèce. Généralement, $\tau > 0$ est appelé le paramètre de variance (marginale), $\alpha > 0$ le paramètre d'échelle et $\nu > 0$ le paramètre de lissage. Si $\nu = \frac{1}{2} + k, k \in \mathbb{N}$, l'équation 1 se réduit au produit d'une exponentielle et d'un polynôme[5]:

$$C(h) = \tau e^{-h\alpha} \sum_{l=0}^{k} \frac{(k+l)!}{(2k)!} \binom{k}{l} (2h\alpha)^{k-l}$$

Lorsque $\nu = \frac{1}{2}$, C est la fonction de covariance exponentielle, et elle devient gaussienne pour $\nu = +\infty$. De manière plus générale, pour $\nu = \frac{1}{2} + k$, le champ U sera de classe C^k. Par conséquent, définir notre champ revient à estimer l'hyperparamètre $\theta = (\tau, \alpha, \nu)$ de la fonction de covariance. On note C_θ la fonction de covariance d'hyperparamètre θ, et Σ_θ la matrice de covariance associée aux points $\{X_i, i = 1, \cdots, N\}$, donnée par:

$$\Sigma_\theta = \begin{pmatrix} C_\theta(||X_1 - X_1||) & \cdots & C_\theta(||X_1 - X_N||) \\ \vdots & \ddots & \vdots \\ C_\theta(||X_N - X_1||) & \cdots & C_\theta(||X_N - X_N||) \end{pmatrix}$$

2.4 Estimation des paramètres de la fonction de covariance

Estimation par maximum de vraisemblance. Le premier estimateur considéré est obtenu par maximum de vraisemblance. Dans notre modèle, la fonction de vraisemblance est définie par:

$$L(\theta \mid U_X) = f(U_X \mid \theta) = \frac{1}{(2\pi)^{N/2} |\Sigma_\theta|^{1/2}} e^{-\frac{U_X^T \Sigma_\theta^{-1} U_X}{2}}$$

où $U_X = \begin{pmatrix} U_1 & \cdots & U_N \end{pmatrix}$. Pour ce faire, nous devons minimiser la fonction de log-vraisemblance négative, donnée par:

$$\begin{aligned} -\log(L(\theta \mid U_X)) &= \frac{N}{2}\ln(2\pi) + \frac{N}{2}\ln(\tau) + \frac{\ln|V_{\alpha,\nu}|}{2} \\ &+ \frac{U_X^T V_{\alpha,\nu}^{-1} U_X}{2\tau} \end{aligned}$$

où $\tau V_{\alpha,\nu} = \Sigma_\theta$. Il n'existe cependant pas de solution analytique à cette équation, et devons alors utiliser des méthodes numériques pour déterminer $\hat{\theta}$. Pour ce faire, nous avons choisi de comparer les méthodes de Nelder-Mead [9], la descente du gradient et de Newton. L'optimisation sur ν étant difficile dans notre cas, nous avons estimé ce paramètre par validation croisée.

Estimation par inférence bayésienne. Malgré l'utilisation de méthodes itératives, notre estimation de l'estimateur du maximum de vraisemblance peut converger vers des maximums locaux. Afin d'éviter cela, nous avons choisi d'utiliser d'autres estimateurs basés sur l'inférence bayésienne. On va alors trouver des estimateurs à partir de la loi a posteriori de nos paramètres $f(\theta \mid U_X)$. Cette dernière est construite, avec une loi a priori $\pi(\theta)$ sur nos paramètres, à l'aide de la règle de Bayes:

$$f(\theta \mid U_X) = \frac{f(U_X \mid \theta)\pi(\theta)}{\pi(U_X)} \propto L(\theta \mid U_X)\pi(\theta)$$

Cependant, la loi de la densité a posteriori de nos paramètres n'étant pas calculable, nous allons alors échantillonner cette distribution. Pour ce faire nous utilisons une méthode de Monte-Carlo par chaîne de Markov (MCMC): l'algorithme de Metropolis-Hastings [6]. Ce dernier construit, à partir d'une loi de proposition $q(. \mid \theta)$, une chaine de Markov de la manière suivante.

- Choisir $\theta^1 \sim \pi(\theta)$
- Créer θ^{t+1} à l'aide de θ^t
  1. Générer $\theta^* \sim q(. \mid \theta^t)$
  2. Calculer la probabilité d'acceptation p:
  $$p = \min\left(1, \frac{\pi(\theta^{t+1})L(\theta^{t+1} \mid U_X)q(\theta^t \mid \theta^{t+1})}{\pi(\theta^t)L(\theta^t \mid U_X)q(\theta^{t+1} \mid \theta^t)}\right)$$
  3. Choisir $\theta^{t+1} = \theta^*$ avec probabilité p, sinon choisir $\theta^{t+1} = \theta^t$

Afin de construire notre chaine de Markov, nous devons définir les lois de $\pi(\theta)$ et $(q(\theta \mid .)$. Pour ce faire, nous supposons tout d'abord que ces lois sont séparables, c'est-à-dire que $q(\theta \mid .) = q(\tau \mid .)q(\alpha \mid .)q(\nu \mid .)$ et $\pi(\theta) = \pi(\tau)\pi(\alpha)\pi(\nu)$. N'ayant que peu d'informations a priori sur nos paramètres, nous choisissons de mettre des lois peu informatives, dont un résumé est présenté en Tableau 1.

Lois a priori $\pi(.)$	Lois de proposition $q(.\|\tilde{\theta})$
$\alpha \sim \mathbb{U}[0, 1]$	$\alpha \sim \mathbb{U}[\tilde{\alpha} - 0.05, \tilde{\alpha} + 0.05]$
$\tau \sim \mathbb{U}[0, 500]$	$\tau \sim \mathbb{U}[\tilde{\tau} - 50, \tilde{\tau} + 50]$
$\nu \sim \mathbb{U}[[\frac{1}{2}, \cdots, \frac{11}{2}]]$	$\nu \sim \mathbb{U}[[\tilde{\nu} - 1, \cdots, \tilde{\nu} + 1]]$

Tableau 1: Distribution a priori et de proposition sur le paramètre θ.

©AfIA 2018

Une fois la chaine MCMC construite, nous avons choisi d'estimer $\hat{\theta}$ à partir de la loi a posteriori $f(\alpha \mid U_X)$, plutôt que de $f(\theta \mid U_X)$. Pour ce faire, on estime tout d'abord $\hat{\alpha}$, puis on sélectionne l'hyperparamètre $\hat{\theta}$ correspondant. Il existe trois estimateurs bayésiens, que nous utiliserons par la suite: le maximum a posteriori (MAP), la moyenne et la médiane.

2.5 Interpolation sur une nouvelle position

Une fois l'hyperparamètre $\hat{\theta}$ estimé, la fonction de covariance $\hat{C}$ est connue, et ainsi la loi de U également. Il nous reste alors à interpoler le déplacement $U(X^*)$ sur une nouvelle position X^*. Pour ce faire, nous allons calculer l'espérance conditionnelle. Comme U est un processus gaussien, ceci revient à effectuer un krigeage simple[4], défini par:

$$U(X^*) = \begin{pmatrix} U_1 & \cdots & U_N \end{pmatrix} \hat{\Sigma}^{-1} \begin{pmatrix} \hat{C}(||X^* - X_1||) \\ \cdots \\ \hat{C}(||X^* - X_N||) \end{pmatrix}$$

3 Applications

Pour chaque exemple, on discrétise les courbes en 300 points. Avant de présenter nos résultats, nous présentons les critères de qualité du recalage.

3.1 Critères sur la qualité du recalage

Un bon recalage doit avoir une faible erreur d'interpolation, et doit être lisse. Pour évaluer la qualité d'interpolation de notre méthode , on utilise 200 points pour effectuer le recalage, et les 100 restants pour l'évaluation. La qualité est alors estimée entre les points d'évaluation à l'aide de la racine carré des erreurs en moyenne quadratique (RMSE). Comme on utilise des courbes pour effectuer notre recalage, on peut aussi utiliser ces dernières pour estimer la qualité d'interpolation. Nous avons choisi d'utiliser la distance de Fréchet (FD) définie par:

$$d_F(F_1, F_2) = \inf_{\gamma_1, \gamma_2 \in \Gamma} \max \{d(F_1 \circ \gamma_1, F_2 \circ \gamma_2)\}$$

où γ_1 et γ_2 sont des reparamétrisations. Concernant la régularité du champ de déformation, nous utilisons la carte de la norme du laplacien. Un maximum de la norme du laplacien (MaxLap) faible signifie alors que le champ de déformation est lisse.

3.2 Application à des données synthétiques

Afin d'évaluer la performance de notre approche, nous nous intéressons à des données synthétiques présentant différents degrés de déformation. Un exemple de résultat obtenu est présenté en Figure 2. En ce basant sur la courbe de la norme du laplacien, on remarque que la déformation est lisse et locale. De plus, d'après le Tableau 2, qui résume l'évaluation de chaque méthode sur cet exemple, le champ de déformation admet de faibles erreurs d'interpolation.

Méthode	RMSE	DF	MaxLap
Nelder-Mead	0.0713	0.5813	0.2902
Gradient	0.0699	0.5541	0.2889
Newton	**0.0698**	**0.5477**	**0.2878**
MAP	0.0718	0.5909	0.2982
Moyenne	0.0721	0.5947	0.3003
Mediane	0.0718	0.5918	0.2990

Tableau 2: Evaluation quantitative des différentes méthodes d'estimation des paramètres sur un exemple de données synthétique présenté en Figure 2.

3.3 Application sur des données réelles

Contrairement aux données synthétiques, plusieurs courbes peuvent être en correspondance pour le même recalage d'images IRM/échographie. La Figure 3 montre un exemple de ce type de données, dont les résultats obtenus sont résumés dans le Tableau 3. Sur cet exemple, les méthodes d'optimisation de Nelder-Mead et de Newton ne donnent pas de bons résultats. Pour la descente du gradient et l'inférence bayésienne, le recalage est lisse et donne de faibles erreurs d'interpolation.
Afin d'étudier la stabilité de notre méthode, nous l'avons appliqué sur 7 données réelles. Les résultats obtenus sont résumés sur la Figure 4. En moyenne, l'erreur d'interpolation est faible pour tous les estimateurs, avec un champ de déformation relativement lisse. Les estimateurs bayésiens sont plus performants car ils fournissent un champ de déformation plus lisse, tout en gardant une petite erreur d'interpolation.

Méthode	RMSE	DF	MaxLap
Nelder Mead	0.1272	0.1796	0.7550
Gradient	0.1353	**0.1132**	0.4091
Newton	**0.1268**	0.1755	0.7058
MAP	0.1383	0.1200	**0.3667**
Moyenne	0.1380	0.1190	0.3680
Mediane	0.1381	0.1201	0.3673

Tableau 3: Évaluation quantitative des différentes méthodes d'estimation des paramètres sur un exemple de données réelle présenté en Figure 3.

4 Conclusion

Nous avons construit un outil de recalage d'images, basé sur les champs gaussiens. Cette méthode est efficace car le champ estimé est lisse et admet de faibles erreurs d'interpolation.
Afin d'améliorer l'estimation des paramètres, et donc la qualité du recalage, plusieurs approches sont envisagées

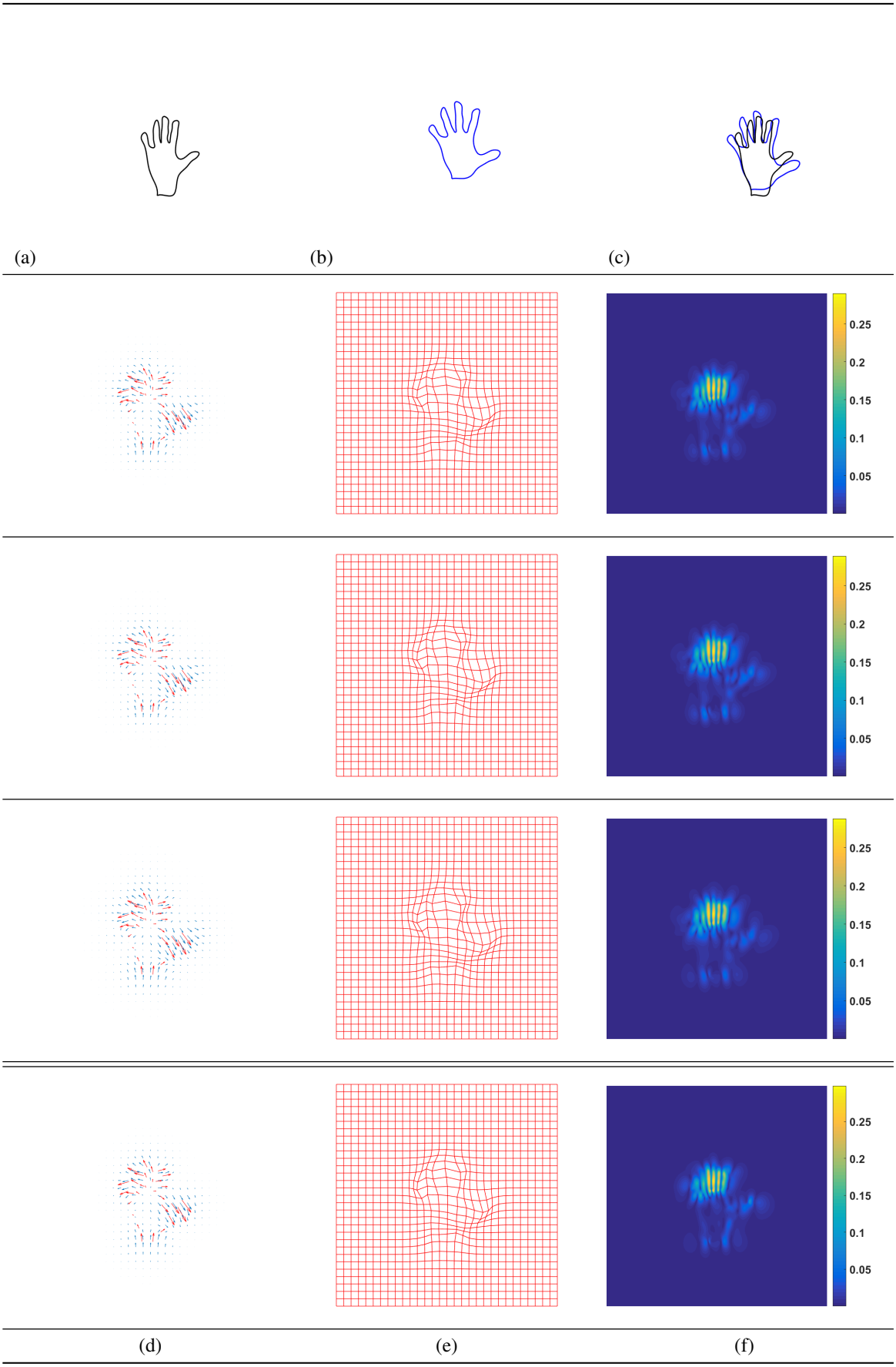


Figure 2: Exemple de champ de déformation estimé sur des données synthétiques: (a) courbes sources, (b) courbes cible, (c) courbes sources (en noir) et cible (en bleu) superposées, (d) le champ de déformation connu (rouge) et estimé (bleu), (e) la grille uniforme déformé par le champ estimé, et (f) la carte de la norme du laplacien. De haut en bas, les résultats sont obtenus après utilisation de la méthode d'optimisation suivante: Nelder-Mead, descente du gradient, Newton, Map

©AfIA 2018

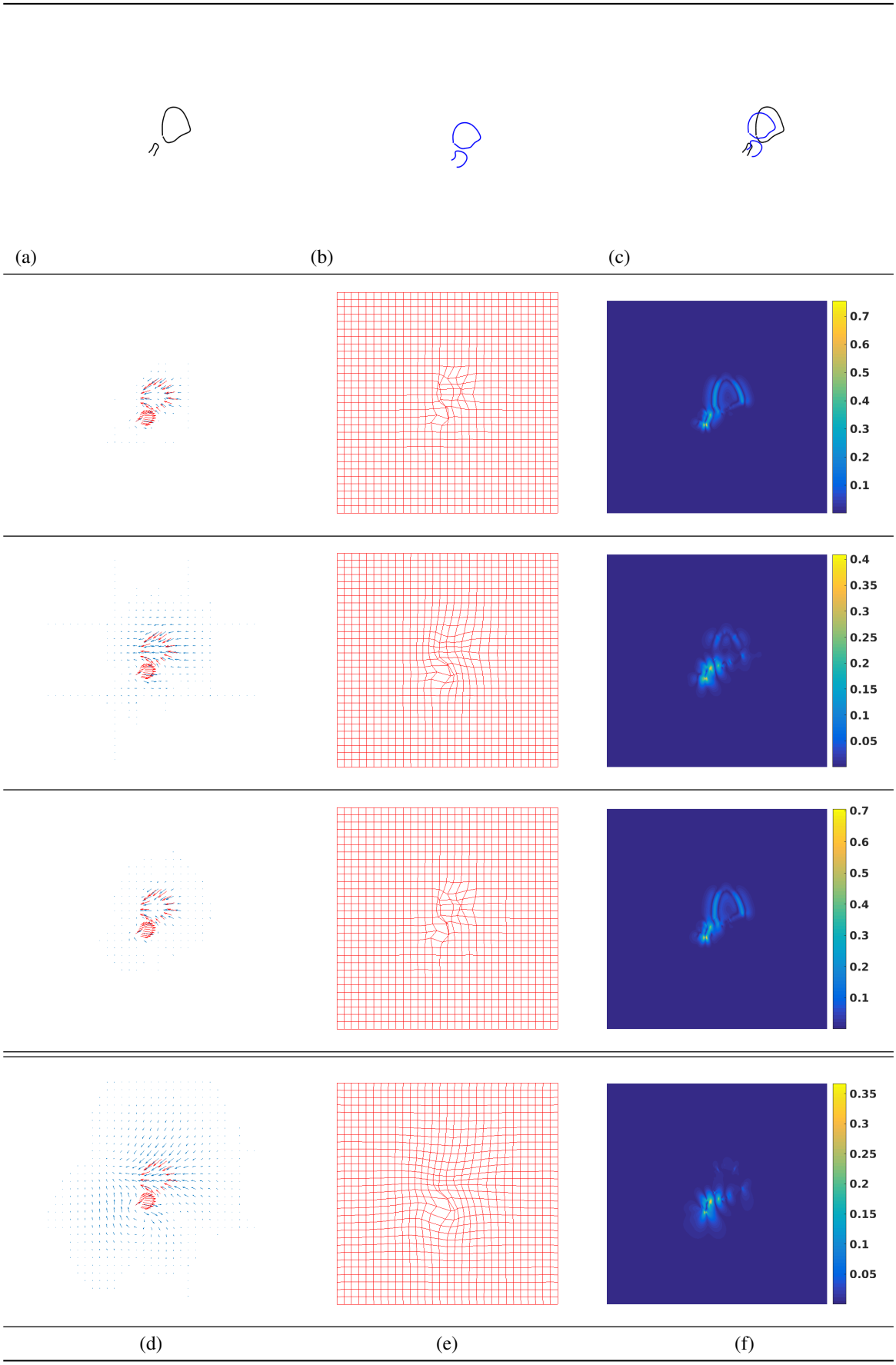
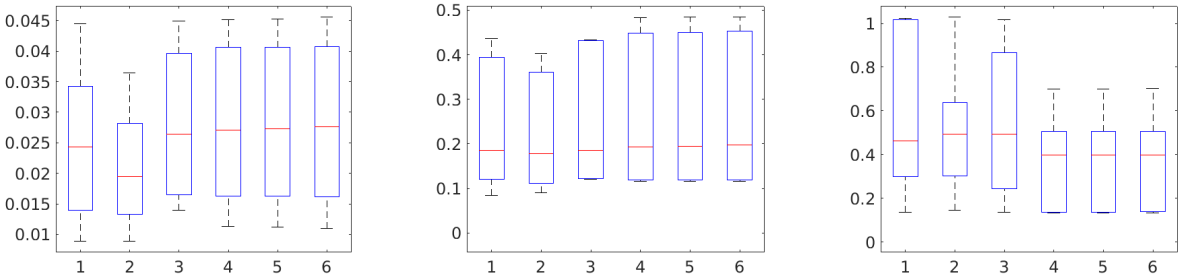


Figure 3: Exemple de champ de déformation estimé sur un recalage IRM/échographie: (a) courbes sources, (b) courbes cible, (c) courbes sources (en noir) et cible (en bleu) superposées, (d) le champ de déformation connu (rouge) et estimé (bleu), (e) la grille uniforme déformé par le champ estimé, et (f) la carte de la norme du laplacien. De haut en bas, les résultats sont obtenus après utilisation de la méthode d'optimisation suivante: Nelder-Mead, descente du gradient, Newton, Map.



(a) RMSE (b) FD (c) MaxLap

Figure 4: Évaluation quantitative des différentes méthodes d'estimation des paramètres sur 7 données réelles: (1) Nelder-Mead, (2) descente du gradient, (3) Newton, (4) MAP, (5) moyenne, et (6) médiane. Pour chaque méthode, nous présentons un boxplot de (a) la RMSE,(b) la distance de Fréchet, et (c) le maximum de la norme du laplacien.

et feront l'objet de futurs travaux. Tout d'abord, pour les méthodes de la descente du gradient et de Newton, une approximation des dérivées partielles semble nécessaire pour éviter les erreurs numériques. De plus, un pas adaptatif pourrait améliorer la convergence de ces algorithmes. Dans un second temps, on pourrait utiliser des lois a priori plus informatives sur nos paramètres, ainsi que des lois de propositions plus restreintes pour améliorer la convergence des chaines MCMC.

Références

[1] R. J. Adler and J. E. Taylor. *Random Fields and Geometry*. Springer Monographs in Mathematics, 2007.

[2] O. Arandjelović, D.-S. Pham, and S. Venkatesh. Efficient and accurate set-based registration of time-separated aerial images. *Pattern Recognition*, 48(11):3466–3476, 2015.

[3] L. P. Chamié, R. Blasbalg, R. M. A. Pereira, G. Warmbrand, and P. C. Serafini. Findings of pelvic endometriosis at transvaginal us, mr imaging, and laparoscopy. *Radiographics*, 31(4):E77–E100, 2011.

[4] N. Cressie. *Statistics for Spatial Data, Revised Edition*. Wiley, 1993.

[5] T. Gneiting, W. Kleiber, and M. Schlather. Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association*, 105:1167–1177, 2010.

[6] W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[7] S. Kurtek, A. Srivastava, E. Klassen, and Z. Ding. Statistical modeling of curves using shapes and related features. *Journal of the American Statistical Association*, 107(499):1152–1165, 2012.

[8] W. Mio, A. Srivastava, and S. Joshi. On shape of plane elastic curves. *International Journal of Computer Vision*, 73(3):307–324, 2007.

[9] J. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7, issue. 4:308–313, 1965.

[10] K. Rohr. *Landmark-based image analysis: Using geometric and intensity models*. Kluwer Academic Publishing, 2001.

[11] J. E. Roos, D. Weishaupt, S. Wildermuth, J. K. Willmann, B. Marincek, and P. R. Hilfiker. Experience of 4 years with open mr defecography: pictorial review of anorectal anatomy and disease. *Radiographics*, 22(4):817–832, 2002.

[12] A. Sotiras, C. Davatazikos, and N. Paragios. Deformable medical image registration : A survey. *INRIA Report*, september 2012.

[13] A. Srivastava, E. Klassen, S. Joshi, and I. Jermyn. Shape analysis of elastic curves in Euclidean spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33:1415–1428, 2011.

[14] M. L. Stein. *Interpolation of Spatial Data*. Springer Series in Statistics, 1999.

[15] Z. Tu, W. Xie, J. Cao, C. Van Gemeren, R. Poppe, and R. C. Veltkamp. Variational method for joint optical flow estimation and edge-aware image restoration. *Pattern Recognition*, 65:11–25, 2017.

©AfIA 2018

oogo : Ontologie des Outils utiles à la Gestion d'Ontologies

Sylvie Despres

Université Paris 13, Sorbonne Paris Cité, LIMICS, (U1142), INSERM, Sorbonne Universités, UPMC Université Paris6, 74 rue Marcel Cachin F-93017 Bobigny cedex, France

sylvie.despres@univ-paris13.fr

Résumé

Cet article présente, oogo, une ontologie ayant pour objectif la description des outils utiles à la gestion d'ontologies. A l'origine, ce travail répondait à une demande qui était de présenter à des chercheurs et/ou des industriels des outils utilisables pour élaborer une ontologie de domaine et si nécessaire leur permettre de les sélectionner en fonction de leurs besoins. Une ontologie nous a semblé constituer la forme la plus adaptée pour délivrer cette information. Le modèle conceptuel la structurant est fondé sur le cycle classique de construction d'ontologies commun à toutes méthodologies. Les activités afférentes à la construction d'ontologies servent de guide à la présentation des fonctionnalités les plus utiles de ces outils. Les scénarios d'usage sont construits à partir des besoins des acteurs intervenant autour d'une ontologie. Un état de l'art relatif à chaque catégorie d'outils sous-tend la construction du modèle conceptuel de oogo.

Mots Clef

oogo, ontologie, ingénierie des ontologies, outils pour l'ingénierie ontologique.

Abstract

This paper presents oogo an ontology for describing tools that are used for the management of ontologies.
Originally, this work responded to a request that was to present to researchers and/or manufacturers some tools used to develop a domain ontology and if necessary allow them to select among them the ones that were best fitted to their needs. An ontology seemed the most suitable form to deliver this information. The conceptual model structuring oogo is based on the classical cycle of the ontology design that is shared by all of the methodologies. Ontology building activities serve as a guide for presenting the most useful features of these tools. The usage scenarios are built from the needs of actors intervening around an ontology. A state of the art relating to each category of tools underlies the construction of the oogo conceptual model.

Keywords

oogo, ontology, ontology engineering, tools for ontology engineering.

1 Introduction

Dans cet article, nous décrivons, oogo, une ontologie ayant pour objectif la description des outils utiles à la gestion d'ontologies. A l'origine, ce travail répondait à une demande des organisateurs de la journée PDIA2017 qui était de présenter à des chercheurs et/ou des industriels des outils utilisables pour élaborer une ontologie de domaine et si nécessaire leur permettre de les sélectionner en fonction de leurs besoins.

La première question qui s'est posée pour répondre à cette requête a été de trouver la forme la plus pertinente pour renseigner ces acteurs. Il existe en effet de nombreux articles décrivant les différents outils et de nombreuses comparaisons les concernant. Ces dernières sont difficiles à exploiter car elles sont statiques, dédiées à des catégories d'outils particulières et les critères de comparaison utilisés peuvent varier d'un article à l'autre. Dans ce contexte, l'utilisation d'une ontologie nous a semblé la forme la mieux adaptée pour donner accès à l'ensemble des caractéristiques des outils actuellement disponibles. Elle permet entre autres de répondre aux questions les plus fréquemment posées lors de la sélection des outils intervenant dans la gestion d'une telle ressource. Elle est destinée à être publiée pour être réutilisée et enrichie par la communauté d'ingénierie des ontologies. Les utilisateurs concernés par oogo sont des ingénieurs de la connaissance ou des chargés de mission en « information technology » en charge de la construction d'ontologies et s'interrogeant sur les fonctionnalités des outils existants.

Le modèle conceptuel de oogo est fondé sur les activités intervenant dans le cycle classique de construction d'ontologies (scénario 1 de la méthodologie Neon) commun à toutes méthodologies de construction [37]. Les activités afférentes à la construction d'ontologies servent de guide à la présentation des fonctionnalités les plus utilisées de ces outils [38]. Le périmètre de l'ontologie est établi en adoptant la méthodologie de [41]. Les scénarios d'usage identifiés sont issus de cas réels rencontrés lors de la construction d'ontologies. A partir de ces scénarios, une liste non exhaustive de questions de compétence, auxquelles oogo doit être en mesure de répondre, a été élaborée à partir des questions posées par les acteurs devant construire une ontologie. Un état de l'art relatif à chaque outil aide à la construction du modèle conceptuel de oogo.

Dans la suite de l'article, la section 2 décrit le modèle adopté pour le concept Outil à partir duquel les modèles des outils présentés dans la section 3 sont décrits. Puis nous concluons dans la section 4 sur l'intérêt et l'usage d'une telle ontologie.

©AfIA 2018

2 Modèle adopté pour le concept Outil

Nous envisageons l'outil comme un moyen permettant d'obtenir un résultat correspondant à la mise en œuvre d'un service. Nous proposons un modèle pour le concept d'outil dédié à la construction de ressources du web sémantique que nous spécialiserons en fonction du type d'outil étudié. En utilisant plusieurs vocabulaires standards, un outil est identifié par une URI correspondant à sa homePage, une URI de documentation et une URI de téléchargement. Il peut être caractérisé par des mots-clés. Il a été conçu par un auteur ou une équipe d'auteurs et développé par une organisation. Il est exécutable sous un SE. Il est identifié par un type de licence. Il est mis à disposition selon différentes modalités. Il fournit un type de service. Il est interopérable avec d'autres outils. Il dispose d'un format de stockage pour les fichiers contenant les ressources. Il permet l'import et l'export de ressources. Il est capable de gérer les sauvegardes avec ou sans gestion de version.

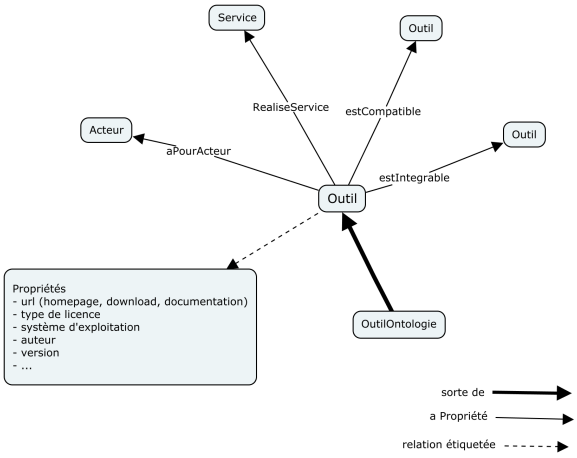


Figure 1 – Modèle du concept Outil

Plusieurs vocabulaires sont réutilisés pour la construction du modèle concept Outil : dcterms, doap, dvia, educore, foaf, org.

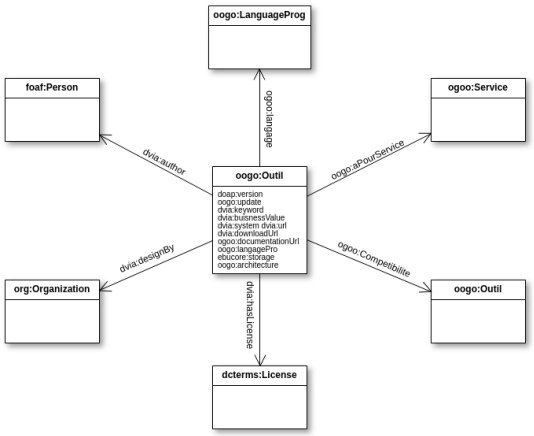


Figure 2 – Aperçu des vocabulaires réutilisés.

3 Démarche de construction de l'ontologie oogo

Nous nous intéressons aux outils dédiés à la construction d'ontologie définie comme « une spécification formelle explicite d'une conceptualisation partagée » [36]. Le langage de représentation considéré pour la formalisation de ces ontologies est OWL/OWL2 standard du W3C.

Le cycle classique de développement d'une ontologie est constitué des étapes de spécification, planification, conceptualisation, formalisation, implémentation [38]. Au cours de ce cycle, les activités mises en œuvre [39] comportent les activités de gestion, orientées développement (pré-développement, développement, post-développement) et de support à la construction. Dans oogo, nous décrivons une partie des outils supportant les activités de support et de développement.

Le cadre méthodologique dans lequel nous nous situons est celui du projet NeON (http://www.neon-project.org/). Nous adoptons le cycle classique de construction d'une ontologie (spécification, planification, conceptualisation, formalisation, implémentation) correspondant au scénario 1 de la méthodologie Neon.

La spécification des besoins auxquels doit répondre l'ontologie a été décrite dans l'introduction. La sélection des outils présentés reposent sur des critères qui sont relatifs à la disponibilité pour une installation, leur maintenance, leur actualité et leur formalisation en OWL. Elle est non exhaustive et pourra être étendue au fil de l'utilisation de l'ontologie. Une partie a été collectée via le site du w3c et de AI3 ( ) et les articles synthétisant les différents outils existants. Ces outils ont servi à l'élaboration des modèles leur correspondant et sont décrits dans oogo.

Nous avons identifié, pour chacun d'entre eux, les scénarios d'usage issus de l'expérience acquise au cours des différents projets de recherche menés au LIMICS et les questions de compétences qui leur sont associées, c'est-à-dire les questions auxquelles notre ontologie doit être en mesure de répondre. La construction des modèles de connaissances qui leur sont associés, est décrite dans la suite de ce paragraphe.

3.1 Activités de support

Les activités de support que nous considérons sont la modélisation des connaissances, la réutilisation, l'évaluation et la validation. Les outils pouvant supporter ces activités sont les outils de recherche d'ontologies (moteurs de recherche et bibliothèques d'ontologies) et les outils de validation.

3.1.1 Outils servant à la modélisation

Des outils de modélisation graphique semi-formelle peuvent être utilisés dans la phase de conceptualisation. Ils permettent de construire le modèle conceptuel du domaine qui sera ensuite formalisé avec l'éditeur d'ontologie. Leur utilisation intervient au cours de la phase d'acquisition de connaissances qui est réalisée de manière incrémentale.

3.1.1.1 Les scénarios

Scénario 1 : L'utilisateur cherche à cerner le périmètre de l'ontologie à construire. Il doit choisir une représentation sous forme de carte cognitive.

Scénario 2 : L'utilisateur cherche à représenter graphiquement le modèle du domaine en partant d'un concept central. oogo doit décrire les outils de modélisation avec leurs propriétés.

3.1.1.2 Les questions de compétence

Question 1 : J'ai identifié un concept central et je cherche à décrire les entités connexes en indiquant les relations qui les lient. Quel outil me permet de le faire ?

Question 2 : Comment déterminer la granularité de la représentation ?

Question 3 : Est-il possible de traduire automatiquement les cartes obtenues au format OWL ?

3.1.1.3 Les outils de modélisation sélectionnés

Nous avons sélectionné les outils Yed Graph Editor , XMind et CmapsTools . Ces trois outils disposent de versions exploitables sousWindows, OsX et Linux. Nous ne considérons pas dans cet article les métamodèles de représentation de OWL dans le formalisme UML et le langage GOWL de modélisation graphique, polymorphique et typé pour la construction d'ontologie [14].

Yed Graph Editor est un éditeur de graphes de bas niveau, open source et multiplateforme, personnalisable par programme. Il est souple dans les options de présentation disponibles et possède plusieurs fonctionnalités de disposition automatique des éléments du graphe. Il permet d'exporter les graphes au format graphml, SVG, png, emf, eps, etc.

XMind est un outil de carte heuristique qui permet également la construction de carte conceptuelle. Il permet de créer, gérer et partager des cartes conceptuelles et de les exporter au formats txt, html, image, SVG. XMind propose une version open source et une version payante. Dans la seconde version, il est possible d'exporter au format csv. Cette fonctionnalité peut ensuite être exploitée par d'autres outils de création semi-automatisée de l'ontologie.

CmapTools est un outil qui permet de créer et partager facilement des cartes conceptuelles. Il est utilisable en ligne via un simple navigateur ou téléchargeable sous la forme d'un exécutable. Il dispose d'options de mise en page automatique et permet d'exporter au format image, PDF, etc. Il permet de fusionner des nœuds identiques, de créer des nœuds imbriqués de valider des liens, de réparer automatiquement les liens rompus et d'annoter les nœuds. Il permet également la comparaison des cartes qui est une fonctionnalité utile pour la construction collaborative.

3.1.1.4 Le modèle pour les outils de modélisation

Le modèle du concept Outil de modélisation prend en compte les fonctionnalités qui leur correspondent. Nous avons en particulier retenu le mode de mise en forme, les formats d'export, leur licence, les options de partage, le mode d'exploitation (en ligne ou téléchargeable). Les notions de cartes heuristique et conceptuelle sont décrites.

3.1.2 Outils de recherche

3.1.2.1 Les scénarios

Scénario 1 : L'utilisateur cherche à réutiliser des ressources ontologiques pour construire sa propre ressource. Il s'interroge sur les outils lui permettant de réaliser cette recherche et le type de recherche qu'il peut effectuer. oogo doit représenter les outils avec les types de recherche associés.

Scénario 2 : L'utilisateur s'interroge sur le type de recherche utilisé par l'outil pour retrouver une ressource. oogo doit décrire les différentes modalités de recherche associées aux outils.

Scénario 3 : L'utilisateur a réalisé une recherche qui lui renvoie une liste de ressources, il cherche une explication du classement des ressources retournées. oogo doit préciser les métriques utilisée pour le classement (ranking) des ressources.

3.1.2.2 Les questions de compétence

Question 1 : Quels outils de recherche existent pour m'aider à trouver une ontologie de domaine ? (scénario 1)

Question 2 : Existe-t-il des vocabulaires dont je pourrais réutiliser les termes pour désigner les concepts ou les relations de mon ontologie ? (scénario 1)

Question 4 : Quelles sont les modalités de recherche mises en œuvre pour trouver une ressource ? (scénario 2)

Question 5 : Comment s'explique la classification des ressources proposées par l'outil de recherche ? (scénario 3)

3.1.2.3 Les outils de recherche sélectionnés

Parmi les outils de recherche permettant de trouver des ressources ontologiques répondant à des besoins de réutilisation figurent les moteurs de recherche sémantique, les portails et les bibliothèques d'ontologies.

- Les moteurs de recherche Swoogle [11], Watson [7] et Vocab.cc [32] sont dédiés à la recherche d'éléments d'ontologies dont les vocabulaires constituent un sous-ensemble. L'expression de la recherche se fait en langue naturelle avec des mots clés et les résultats produits sont classés. Les modalités de recherche sont les suivantes : - par URI ; - par mots-clés correspondant à une classe, une relation, un individu, et au périmètre de l'ontologie (Label, Commentaire, etc.) ; - par requêtes sur des triplets ((terme générique, relation, terme) (terme générique, relation, ?), (terme générique relation, ?)) [28].
- Le portail Linked Open Vocabularies (LOV) (une version indépendante est hébergée à l'OKFN), un des modules de Datalift [42], qui est destiné à cataloguer, trouver et choisir des ontologies. Chaque vocabulaire est maintenu et publié de manière indépendante tout en entretenant des liens de dépendance vis-à-vis des autres vocabulaires.
- Les bibliothèques d'ontologies (BioPortal, AgroPortal, OBOFoundry, ODP) dont la finalité est de permettre et faciliter l'interopérabilité, fournir des ontologies reconnues et testées. Elles sont caractérisées par leur domaine, leur contenu et leurs fonctionnalités.
- L'outil OntoFox prend en charge la réutilisation des ontologies. Il permet aux utilisateurs de saisir des termes, d'extraire des propriétés sélectionnées, des

©AfIA 2018

annotations et certaines classes de termes connexes à partir d'ontologies sources et de sauvegarder les résultats à l'aide de la sérialisation RDF / XML du langage OWL. Ontofox dispose également d'un algorithme d'extraction de termes fondé sur SPARQL qui extrait des termes liés à un ensemble donné de termes de signature. En outre, Ontofox fournit une option pour extraire la hiérarchie enracinée à un terme d'ontologie spécifié.

3.1.2.4 Le modèle proposé pour les outils de recherche

Le modèle pour les moteurs et le portail LOV a été construit à partir du tableau comparatif établi par [43]. Les caractéristiques retenues comportent la méthode d'accès à la ressource (automatique et/ou manuelle), le type de la ressource (Semantic Web Document, concepts, termes de vocabulaire, vocabulaires), la métrique de ranking, le filtrage du domaine, l'accès au service web, etc.).

Le modèle pour les bibliothèques a été établi à partir du tableau comparatif proposé par [22]. Les caractéristiques décrites sont le domaine, les fonctionnalités de navigation et de recherche (sujet, structure, langue), le type d'appariement supporté (entre classe ou ontologie), le processus de soumission, l'utilisation et l'accès aux ontologies, les formats de fichiers en entrée et sortie, les outils associés.

3.1.3 Outils de validation

3.1.3.1 Les scénarios

Scénario 1 : L'utilisateur cherche à valider l'ontologie qu'il a construite et s'interroge sur les outils lui permettant de réaliser cette tâche. L'ontologie doit décrire les outils permettant les différents types de validation.

Scénario 2 : Quels sont les critères garantissant la validité d'une ontologie ? oogo doit préciser les critères utilisés pour la validation et les outils leur correspondant.

3.1.3.2 Question de compétences

Question 1 : Existe-il des outils pour valider les ontologies ? (scénario 1)

Question 2 : Quel outil me permet de vérifier la structure de mon ontologie ? (scénario 1 et 2)

Question 2 : Quel outil me permet de vérifier le modèle de l'expertise représenté dans mon ontologie ? (scénario 2)

3.1.3.3 Les outils de validation sélectionnés

Nous avons utilisé l'analyse des outils de validation réalisée par [27] pour la sélection des outils de validation. La validation de la qualité d'une ontologie se définit selon deux axes principaux qui sont la validation de la structure et la validation de la sémantique. La validation de la structure correspond à une validation formelle de la logique du modèle. Elle peut être réalisée automatiquement, grâce au développement d'outils dédiés. La validation de la sémantique consiste à mesurer l'adéquation du modèle conceptuel avec la réalité qu'il modélise. Cette étape nécessite une collaboration entre les ontologues et les spécialistes du domaine modélisé dans l'ontologie.

Les outils de validation de structure

Les raisonneurs permettent de vérifier la consistance et la cohérence d'un modèle (cf. Les outils de raisonnement, infra).

OntoCheck [29] constitue un module d'extension à l'éditeur d'ontologies Protégé et vise à contrôler le respect des conventions de nommage, ainsi que l'exhaustivité des méta-données. OntoCheck vérifie les cardinalités, la complétude des méta-données, les labels, les conventions typographiques ainsi que les métriques de l'ontologie.

XD Analyzer a été développé dans le cadre du projet NeOn , pour faire un retour qua- litatif à l'utilisateur en suivant la méthodologie XD. Cette dernière fournit une liste de bonnes pratiques (concernant entre autres, les labels, les commentaires, les concepts non utilisés) à respecter pour la construction d'ontologies. Le module d'extension RaDON [1?] a été ajouté à Neon Toolkit afin de faciliter les tâches de vérification de la consistance. Il comporte deux plugins permettant de travailler sur une ontologie ou un réseau d'ontologies.

OntOLogy Pitfall Scanner! (OOPS!) [26] est un outil indépendant de tout éditeur d'ontologies. Il s'utilise uniquement en ligne. Le but de OOPS ! est l'identification des anomalies ou des mauvaises pratiques dans une ontologie. L'application prend en entrée l'URI d'une ontologie ou le code source en RDF. L'ontologie est chargée *via* l'API Jena avant d'être analysée pour en extraire les erreurs potentielles. Le résultat est une page répertoriant les erreurs selon le piège (pitfall) identifié, avec une proposition de résolution de l'erreur. Les pitfalls peuvent concerner des éléments individuels, plusieurs éléments ou toute l'ontologie.

Les outils de validation sémantique

Les outils permettant de valider la sémantique d'une ontologie nécessitent le recours à des experts du domaine modélisé par l'ontologie. Des outils supportant le développement de construction collaborative d'ontologies contribue à la validation sémantique.

Le serveur Ontolingua propose un environnement collaboratif pour parcourir, créer, éditer, modifier ou utiliser les ontologies (Farquhar et al., 1997), WebProtégé a été développé en reprenant l'architecture de Protégé. Il permet également le développement collaboratif d'ontologies et est accessible *via* n'importe quel navigateur web [40].

3.1.3.4 Le modèle pour les outils de validation

Le modèle des outils de validation décrit les fonctionnalités concernant la validation collaborative d'ontologies et précise les types de validation possible.

3.2 Activités de développement

Parmi les activités intervenant dans la phase de développement, nous considérons la conceptualisation, la formalisation, la modularisation, l'alignement. Les outils supportant ces activités sont les outils d'édition, les raisonneurs, les outils d'alignement et les outils de visualisation.

3.2.1 Outils d'édition

Un éditeur d'ontologie est un outil permettant d'inspecter, de parcourir, de codifier, de modifier les

ontologies et de soutenir de cette manière le développement de l'ontologie et la tâche de maintenance.

3.2.1.1 Les scénarios

Scénario 1 : L'ontologue est débutant et cherche un éditeur pour construire une ontologie. Il n'a pas forcément une idée précise de fonctionnalité de l'outil. oogo doit décrire les différentes fonctionnalités de la phase de développement de l'ontologie.

Scénario 2 : L'ontologue a une idée précise de la nature de l'ontologie qu'il souhaite construire et des tâches qu'elle permettra de réaliser (annotation, raisonnement).

3.2.1.2 Les questions de compétence

Question1 : Quels sont les éditeurs d'ontologie disponibles ? (scénario 1)

Question 2 : Quel est le meilleur éditeur d'ontologie pour un débutant ? (scénario 1)

Question 3 : Quelles sont les activités supportées par l'éditeur d'ontologie ? (scénario 2)

3.2.1.3 Les outils d'édition sélectionnés

Nous avons utilisé le wiki du W3C pour répertorier les éditeurs les plus utilisés actuellement pour construire des ressources ontologiques au sens où nous les avons définis dans cet article. Par conséquent nous n'avons pas retenu les éditeurs de vocabulaire. Parmi la liste des éditeurs inventoriés nous avons retenu Protégé, NeOn Toolkit, SWOOP et TopBraid Composer.

Nous avons également exploité les comparaisons des différents éditeurs pour identifier les propriétés des outils d'édition et construire un modèle de ces outils ([1], [2], [3], [16], [25], [30], [31].

3.2.1.4 Le modèle pour les outils d'édition

Nous avons retenu les caractéristiques servant à ces comparaisons pour construire le modèle des outils d'édition. Elles concernent l'architecture de l'outil et sa capacité d'évolution, l'interopérabilité, le paradigme de représentation des connaissances et le support méthodologique, les services d'inférences, les services de visualisation, la gestion des versions et l'utilisabilité. Certaines de ces caractéristiques sont déjà décrites dans le modèle du concept Outil (cf. Figure 1).

L'étude de l'utilisabilité des éditeurs réalisée par [12] pour des ontologues débutants met en évidence des propriétés relatives à la facilité de prise en main des outils. Certaines propriétés nous semble quantifiables comme accessibilité du langage utilisé à une communauté d'utilisateurs non spécialistes, la visualisation hiérarchique des classes, la navigation d'une classe vers ses instances, le filtrage des classes à partir de critères fixés par l'utilisateur

3.2.2 Outils de raisonnement

Une ontologie exprimée en OWL peut être considérée comme un ensemble de formules FOL (First Order logic), appelées axiomes, et par conséquent comme une théorie logique. Le vocabulaire exprimé en OWL est bien défini et sans ambiguïté. Cependant, il y a souvent un décalage entre la représentation visée et la représentation réelle de la connaissance décrite dans l'ontologie. Alors que la connaissance du domaine d'intérêt est généralement bien maîtrisée ou tout au moins comprise, il n'est pas trivial de prévoir et de comprendre les conséquences logiques d'un ajout, d'un retrait ou d'une modification d'axiome, notamment quand la terminologie est fortement interconnectée. Ainsi, un ontologue a besoin d'utiliser des systèmes de déduction automatisés, communément appelés raisonneurs, pour vérifier certaines formes de déduction à partir d'axiomes préétablis et de rendre explicites ceux qui sont déclarés implicitement. Les raisonneurs sont donc, à l'heure actuelle, des éléments clés pour travailler avec des ontologies [4].

3.2.2.1 Les scénarios

Scénario 1 : L'utilisateur cherche l'outil support du raisonneur (Neon, Protégé, Swoop, etc.). L'ontologie doit donc permettre de représenter les différents supports et de leur associer les raisonneurs existants.

Scénario 2 : L'utilisateur cherche à connaître les tâches supportées par un raisonneur. L'ontologie doit décrire les tâches de raisonnement et les entités sur lesquels ils portent.

Scénario 3 : L'utilisateur cherche les capacités d'explication et de gestion des erreurs du raisonneur.

Scénario 4 : L'utilisateur cherche le raisonneur adapté au contenu de l'ontologie à exploiter ? Dépend de la formalisation adoptée pour l'ontologie qui dépend elle-même des tâches à réaliser (annotation, raisonnement). Au différents profils sont associés des raisonneurs

3.2.2.2 Les questions de compétence

Question 1 : Sous quel système d'exploitation est-il exécutable ? Quel est le type de licence ? (scénario lié au modèle outil)

Question 2 : Avec quel éditeur le raisonneur est-il compatible ? Est-il standalone ?

Question 3 : Le raisonneur peut-il exploiter des règles et quel est le formalisme supporté ?

Question 4 : Le raisonneur prend t-il en charge le raisonnement sur les individus ?

Question 5 : Le raisonneur est-il adapté à la nature de l'ontologie que j'ai construite ?

Question 6 : Quels types de justification (explications sur les inconsistances) peut fournir le raisonneur ?

Question 7 : Quel type de profil OWL est pris en charge par le raisonneur ?

3.2.2.3 Les outils de raisonnement

Nous avons utilisé les comparaisons réalisées par [8], [1], [21] et [4] pour construire le modèle des outils de raisonnement.

[8] dressent une liste de caractéristiques guidant l'utilisateur dans le choix d'un raisonneur en fonction des besoins d'une application. Elle est envisagée selon deux dimensions : les caractéristiques du raisonnement et l'utilisabilité pratique. La première dimension regroupe : la procédure ou l'algorithme utilisé par le raisonneur pour le raisonnement en logique de description (algorithme Tableau, Hyper-Tableau et Consequence-Based), la robustesse et la complétude (vérifie si toutes les inférences possibles sont réalisées) qui peut aider à une accélération du temps de raisonnement, l'expressivité et la complétude (lien avec les profils OWL), la classification incrémentale (synchronisation du raisonneur), le support de l'utilisation de règles (par exemple, SWRL), la justification (explication pour un concept inconsistant, disjonction de concept, implication logique, etc.), le support des tâches de raisonnement sur la Abox

©AfIA 2018

(raisonnement sur les individus, vérification de la consistance de la Abox, etc.). La seconde a trait à la disponibilité (plugin, standalone), au type de licence et d'autres caractéristiques comprenant le type de code, les plateformes compatibles, l'interface Jena native, etc.
Après avoir décrit l'architecture des raisonneurs raisonneurs Pellet, RACER, FaCT++ [39], Snorocket SWRL-IQ, ELK, HermiT, CEL et TrOWL, [1] en dresse une comparaison selon les caractéristiques définies par [8].
[21] réalise une comparaison de 35 raisonneurs dont la plupart ont été développé entre 2010 et 2015. Elle est fondée sur des critères d'utilisabilité (services de raisonnement supporté, les niveaux d'expressivité et la complétude du raisonneur) et les algorithmes de raisonnement implémentés.

3.2.2.4 Le modèle pour les outils de raisonnement

Le modèle des outils de raisonnement a été établi à partir du tableau comparatif présenté par [4] qui est fondé sur l'article [21]. Les raisonneurs décrits sont actuellement FaCT++, Pellet, ELK, HermiT, Konclude [34] , Euler/Eye et NoHR .

3.2.3 Outils de conversion de données en OWL

La finalité de ces outils est la conversion de données contenues dans des feuilles de calcul en OWL.

3.2.3.1 Les scénarios

Scénario 1 : L'utilisateur dispose de documents existant sur la connaissance à formaliser, le plus souvent sous la forme de listes ou de feuilles de calcul. Il souhaite pouvoir les greffer facilement sur le squelette de son ontologie en utilisant les relations qu'il a prédéfinies.
Scénario 2 : L'utilisateur souhaite importer des éléments de hiérarchie à partir d'ontologies existantes avec cependant la possibilité de les amender ou de les complèter.
Scénario 3 : L'utilisateur souhaite mettre à la disposition d'intervenants extérieurs des parties de son travail en cours à des fins de vérification ou de validation.
Scénario 4 : L'utilisateur souhaite introduire dans l'ontologie d'importantes quantités de faits ou de caractéristiques sans avoir la charge de les éditer tous manuellement.

3.2.3.2 Les questions de compétence

Question 1 : L'outil permet-il l'import de données exprimées dans des tableaux (feuilles de calcul) ?
Question 2 : L'outil permet-il l'opération inverse (export vers des feuilles de calcul) d'extraits sélectionnés de son ontologie ?
Question 3 : Ces fonctionnalités sont-elles disponibles pour une utilisation en mode « ligne de commande » pour supporter un certain degré d'automatisation des tâches répétitives lors de l'enrichissement de son ontologie ?

3.2.3.3 Les outils de conversion de données

Nous avons sélectionné l'outil Cellfie un plugin Protégé 5 intégrant MappingMaster [23] et l'outil FOE [9]. Populous et WebPopulous offrent des fonctionnalités similaires ainsi que l'option d'import Excel disponible dans Topbraid Composer .

3.2.3.4 Le modèle pour les outils de conversion de données

Pour construire le modèle des outils de conversion de données, nous avons le type et d'export et les modalités d'utilisation, le degré d'automatisation autorisé et les constructions qu'il est possible de créer.

3.2.3 Outils d'alignement

Pour construire le modèle des outils d'alignement nous avons commencé à tester les outils répertoriés sur le site de AI3 , la librairie des outils d'alignement hébergée par l'INRIA et le site du projet Ontobeep. Nous inventorions également les outils de visualisation d'alignement tels que [19]. Cette partie du travail est en cours et ne sera pas présenté dans l'article.

3.2.4 Outil de gestion d'évolution

Pour construire le modèle des outils d'évolution nous avons commencé à inventorier les outils existants [35]. Nous travaillons à partir des travaux récents de [18] qui réalisent une synthèse de ces outils du point de vue de la visualisation.
Les outils que nous analysons sont CODEX [13], REX [6], NeonToolkit [24].

3.2.5 Outils de visualisation

La visualisation d'ontologies est une tâche qui aide à la compréhension des modèles représentés. Elle est utile à l'ensemble des acteurs participant à sa construction. Or, il n'est pas simple de créer une visualisation des ontologies en raison de la complexité de ces ressources. Les entités à visualiser sont la hiérarchie des concepts, les relations entre les concepts, les attributs des concepts et les individus relevant des concepts représentés [17]. Nous avons identifié deux publications faisant références à des ontologies VO formalisant les connaissances utiles la visualisation d'ontologies [33] et [44] qui décrivent les primitives utilisées dans le domaine de la visualisation. Elles ne sont pas directement exploitables mais nous avons réutilisé certaines des classes modélisées.

3.2.5.1 Les scénarios

Scénario 1 : L'utilisateur souhaite visualiser la hiérarchie des concepts.
Scénario 2 : L'utilisateur souhaite se centrer sur un concept afin de visualiser globalement l'ensemble des liens qu'il entretient avec le réseau de concepts.
Scénario 3 : L'utilisateur souhaite visualiser les explications fournies par le raisonner et disposer d'une représentation graphique.

3.2.5.2 Les questions de compétence

Question 1 : Quel outil permet de visualiser l'ontologie ?
Question 2 : Quelles sont les entités visualisable de l'ontologie ?
Question 3 : Quelles sont les actions possibles pour visualiser l'ontologie ?

3.2.5.3 Les outils de visualisation

Les outils décrits dans la version actuelle de oogo sont GraphViz, CropCircles, KC-Viz, Knoocks, LogDiffViz, OntoGraph, OntologyVisualizer, Jambalaya, OntoTGVizTab, VOWL, WebVOWL, Protupos.

3.2.5.4 Le modèle pour les outils de visualisation

Le modèle des outils de visualisation a été établi à partir de la comparaison et l'évaluation des visualisations d'ontologies présentées par [5] et des publications relatives à certains outils [45], [20].

Dans le cadre de la construction du service web de visualisation centré utilisateur Protupos, [10] ont caractérisé les outils de visualisation actuellement utilisés :

- Graphviz est un plugin de Protégé qui s'exécute dans son propre onglet. Il sert principalement à la visualisation graphique d'ontologie grâce à l'utilisation d'un graphe de nœuds et d'arcs qu'il est possible d'étendre et de réduire dynamiquement via des actions sur des boutons. Sur le principe, Graphviz est très utile pour parcourir la hiérarchie et la manipuler (expansion, respectivement réduction, des propriétés relatives à un nœud). Il permet également la visualisation de la hiérarchie des classes inférées.
- VOWL est un plugin de Protégé qui s'exécute dans son propre onglet et qui ne prend pas en compte les actions effectuées par ailleurs. Par exemple, il n'est pas possible de se centrer sur un concept et de naviguer autour de ce dernier dans l'onglet VOWL. Sur le principe, VOWL est une spécification de notations graphiques pour les ontologies écrites en OWL http://purl.org/vowl/spec/. L'outil est clairement tourné vers l'ontologue.
- WebOWL (http://visualdataweb.de/webvowl/) est un service web implémentant VOWL indépendamment de Protégé. Les actions peuvent être effectuées grâce aux boutons et aux menus. Un inconvénient majeur de WebOWL est que comme pour tout service web, l'ontologie doit être téléchargée.
- CropCircles est un outil intégré à SWOOP qui permet de visualiser la hiérarchie des concepts. Un concept est représenté par un cercle, les fils de ce concept sont représentés à l'intérieur de ce cercle et la hiérarchie est visualisé selon un représentation en coupe. Il est possible de zoomer sur une couche particulière de la hiérarchie en cliquant dans le cercle et de naviguer dans la hiérarchie en cliquant à l'extérieur du cercle. La vision en coupe par niveau hiérarchique permet notamment de bien appréhender la densité d'un sous arbre par rapport à un autre.

Le développement de Protupos a permis de définir des fonctionnalités que nous avons utilisées pour construire le modèle des outils de visualisation :

- visualisation interactive de la hiérarchie : utilisation d'une représentation circulaire "zoomable" par actions de l'utilisateur ;
- de la hiérarchie par coupe de niveau ;
- visualisation de chaîne de propriétés centrée sur un nœud avec la possibilité de développer dynamiquement la construction de la chaîne ;
- visualisation sous la forme de réseaux d'un type de propriété ;
- visualisation centrée sur un concept en utilisant une fonction « élastique (permettant de modifier le focus) » ;
- nuage de tags associés à un concept.

4. Présentation et évaluation de oogo

oogo a été éditée avec le logiciel Protégé (version 5.2.0) . L'ontologie et la description de ses ressources et propriétés seront publiées selon les principes des données liées sur le web et le schéma sera identifié par l'URI HTTP https://www-limics.smbh.univ-paris13.fr/oogo#. Plusieurs vocabulaires sont utilisés pour la construction du modèle outil : dcterms, doap, dvia, foaf, org. Dans cette version de oogo, nous ne pouvons pas réutiliser dvia:url car le type anyUri n'est pas directement supporté par Protégé 5.

4.1 Métrique de oogo

Metrics

Axiom	756
Logical axiom count	368
Declaration axioms count	208
Class count	155
Object property count	17
Data property count	12
Individual count	23
Annotation Property count	6
DL expressivity	ALCHO(D)

Figure 3 – Métrique de l'ontologie oogo

La métrique de l'ontologie est présentée Figure 3. Elle comporte actuellement 155 classes, 17 object property et 12 data property. Elle est en cours d'évolution.

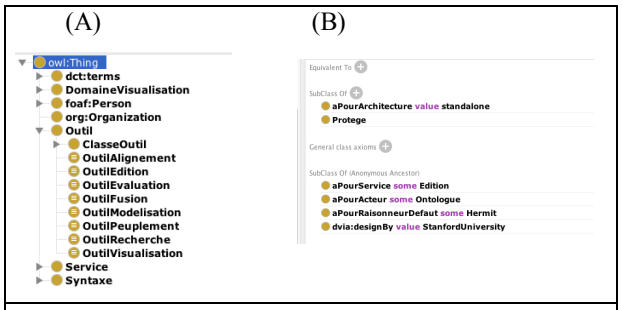


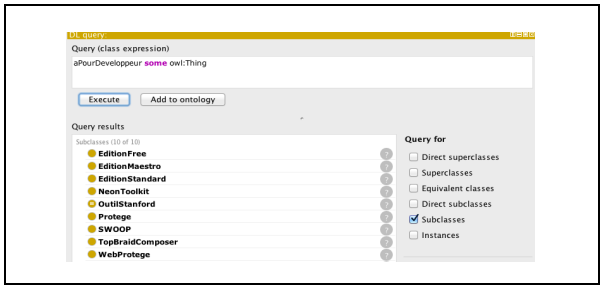
Figure 4 – Vue globale de oogo (A) et Définition d'un outil (B)

4.2 Aperçu des classes de oogo

Nous présentons une description du concept Outil dans oogo. Dans la classe ClasseOutil l'ensemble des outils étudiés sont décrits comme des sous-classes. Par exemple, la classe Protégé comporte les sous-classes Proétég3-x, Protégé4-x, Protégé5-x. Des individus relèvent de ces différentes sous-classes. Elle est également définie comme disposant d'un service d'édition. Les classes définies, comme OutilEdition de la figure 4(A), permettent d'utiliser le raisonneur pour caractériser les outils selon les services qui leur correspondent.

4.3 Requêtes DL Query

Nous montrons par le biais de quelques exemples de requêtes DL Query que les questions de compétence trouvent une réponse.



©AfIA 2018

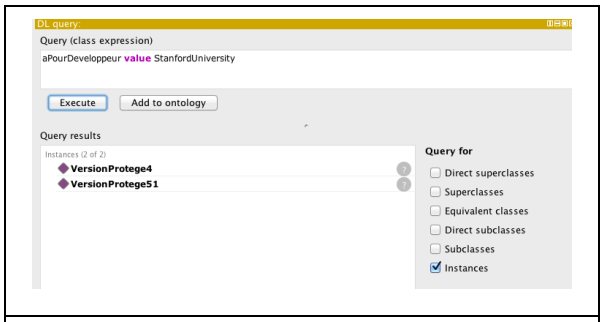


Figure 5 – Requêtes DL Query

Conclusion

Cet article présente un travail en cours qui répondait à une demande qui nous a permis de prendre conscience de la nécessité de centraliser l'expérience de la communauté de manière synthétique afin de répondre simplement et rapidement à des questions pratiques concernant la disponibilité des différents outils utilisables dans le cadre de la construction d'ontologies. L'ontologie ébauchée dans cet article est par nature destinée à évoluer régulièrement une fois qu'elle aura été mise en ligne dans sa version initiale. Actuellement, la langue utilisée pour le développement de oogo est le français néanmoins l'ensemble des entités sont décrites avec des labels et altlabel en anglais. Nous avons le projet de rendre sa construction collaborative lorsqu'elle sera publiée afin de s'assurer qu'elle ne devienne pas obsolète. De cette façon, les retours d'expérience pourront être valorisés et bénéficier à l'ensemble de la communauté.

Bibliographie

[1] Abburu, S. (2012). A Survey on Ontology Reasoners and Comparison. International Journal of Computer Applications, 57(17):33–39.
[2] Abburu, S., Babu, G. S. (2013). Survey on Ontology Construction Tools. International Journal of Scientific & Engineering Research, Volume 4, Issue 6, pp.1748-1752.
[3] Alatrish E. S. (2013). Comparison Some of Ontology Editors, in Management Information Systems, vol 8, n°2, pp. 018-024, 2013.
[4] Alaya N. (2016). Managing The Empirical Hardness of The Ontology Reasoning using The Predictive Modelling. Thèse de l'université de Tunis El Manar et de l'université Paris8.
[5] Balzer L., Do M.T., Maseluk D. (2015). Comparison and Evaluations of Ontology Visualizations. Rapport de recherche, H.5.2, I.2.4.
[6] Christen V., Gross V., Hartung M. Region Evolution eXplorer a tool for discovering evolution trends in ontology regions. Journal of Biomedical Semantics, 6:Article 26, 2015.
[7] D'Aquin M., Motta E. (2011). Watson, more than a Semantic Web search engine. Semantic Web 2(1): 55-63.
[8] Dentler K., Cornet R., & ten Teije A., & de Keizer N. (2011). Comparison of reasoners for large ontologies in the OWL 2 EL profile. Semant. web, 2(2):71–87.
[9] Despres S., Guezennec G. - Flat OWL Editor : un outil utilisant des feuilles de calcul pour découpler les tâches des acteurs impliqués dans la gestion d'une ontologie. Toth2017, (à paraître).
[10] Despres S., Nobecourt J.- Quelles fonctionnalités pour un outil de visualisation d'ontologie ? Atelier VIF Visualisation d'informations, Interaction, et Fouille de données EGC 2018 (http://gt-vif.polytech.univ-nantes.fr/egc-vif2018/atelier_VIF_EGC2018.pdf
[11] Finin T., Ding L., Pa R., Joshi A., Kolari P. (2005). Java A., Peng Y. Swoogle: Searching for knowledge on the semantic web. In Anthony Cohn, editor, Proceedings of the 20th National Conference on Artificial Intelligence - Volume 4, AAAI'05, pages 1682–1683. AAAI Press.
[12] Garcia-Barriocanal E., Sicilia M. A. & Sanchez-Alonso S. (2006). Usability Evaluation of Ontology Editors. Knowledge Organization 32(1),1-9.
[13] Hartung M., Gross A., Rahm E. CODEX: Exploration of semantic changes between ontology versions. Bioinformatics, 26(6):895–896, 2012.
[14] Héon M., Nkambou R., Langheit C. (2016). Toward G-OWL: A graphical, polymorphic and typed syntax for building formal OWL2 ontologies. Proceedings of the 25th International Conference Companion on World Wide Web. International World Wide Web Conferences Steering Committee.
[15] Jain V., Prasad S.V.A.V. (2016). Evaluation and Validation of ontology using Protégé Tool. In International Journal of Research in Engineering & Technology (IMPACT: IJRET) ISSN (E): 2321-8843; ISSN (P): 2347-4599 Vol. 4, Issue 5, May 2016, 1-12.
[1?] Ji, Q., Haase, P., Qu, G., Hitzler, P., Stadtmoeller, S.: RaDON – repair and diagnosis in ontology networks. In: Aroyo, L., Traverso, P., Ciravegna, F., Cimiano, P., Heath, T., Hyvönen, E., Mizoguchi, R., Oren, E., Sabou, M., Simper, L.E. (eds.) ESWC 2009. LNCS, vol. 5554, pp. 863–867. Springer, Heidelberg (2009).
[16] Kapoor, B., Sharma, S.A. (2010). Comparative Study Ontology Building Tools for Semantic Web Applications. International journal of Web & Semantic Technology (IJWesT), 1(3).
[17] Katifori A., Halatsis C. Lepouras G., Vassilakis C., Giannopoulou E. (2007). Ontology Visualization Methods – A Survey. ACM Computing Surveys, Vol. 39, N°4, Article 10.
[18] Lambrix P., Dragisic Z., Ivanova V., Anslow C. Visualization for Ontology Evolution, VOILA 2016 Visualization and Interaction for Ontologies and Linked Data, 2016, pp.54-67.
[19] Lanzenberger M., Sampson J. (2011). Ontology Alignment Quality: A Framework and Tool for Validation. IJISMD 2(3): 1-23
[20] Lohmann S., Link V., Marbach E., Negru S. (2016). Visualizing ontologies with VOWL. 7(4),21. Already accepted papers for the EKAW 2014 special issue, extended version.
[21] Matentzoglu N., Leo J., Hudhra V., Sattler U., Parsia B.. (2015). A survey of current, stand-alone OWL reasoners. In Proceedings of the 4th International Workshop on OWL Reasoner Evaluation, pages 68–79.
[22] Naskar D., Dutta B. (2016) - Ontology Libraries: A Study from an Ontofier and an Ontologist Perspectives. 19th International Symposium on Electronic Theses and Dissertations, Lille.

[23] O'Connor M. J., Halaschek-Wiener C., Musen M. A. (2010). Mapping master: a flexible approach for mapping spreadsheets to OWL. In ISWC'10 Proceedings of the 9th international semantic web conference on The semantic web - Volume Part II, 194-208.
[24] Palma R., Zablith F., Haase P., Corcho O. Ontology evolution. In MC Suarez-Figuero, A Gomez-Prez, E Motta, and A Gangemi, editors, Ontology Engineering in a Networked World, pages 235–255. 2012.
[25] Parveen, Kumar Sahni D.K., Khurana D., Nandal R. (2016). Ontology Development Tools and Languages: A Review. Vol. 5, Issue 6, June, 2016.
[26] Poveda-Villalo, M., Gomez-Perez, A., Suárez-Figueroa, M.C. (2014). OOPS! (OntOlogy Pitfall Scanner!): An On-line Tool for Ontology Evaluation. Int. J. Semantic Web Inf. Syst. 10(2): 7-34.
[27] Richard M. (2017). Apports de la modélisation ontologique pour le partage des connaissances en psychiatrie. Thèse de l'université Pierre et Marie Curie - Paris VI.
[28] Sabou M., D'Aquin M., Motta E. (2008). SCARLET: SemantiC RelAtion DiscoveRy by Harvesting OnLinE OnTologies. ESWC 2008: 854-858.
[29] Schober D., Tudose I., Svatek V., Boeker M. OntoCheck: verifying ontology naming conventions and metadata completeness in Protégé 4. In J Biomed Semantics. 2012; 3(Suppl 2): S4.
[30] Seongwook Y., Anchit A., Preetham C., Paavany J., Ashish M. and Shikha S. (2009). Survey about Ontology Development Tools for Ontology-based Knowledge Management, University of Southern California.
[31] Slimani, T. (2015). Ontology Development: A Comparing Study on Tools, Languages and Formalisms. In Indian Journal of Science and Technology, Vol 8(24), doi:10.17485/ijst/2015/v8i34/54249.
[32] Stadtmüller S., Harth A., Grobelnik M. (2013). Accessing information about linked data vocabularies with vocab.cc. In Juanzi Li, Guilin Qi, Dongyan Zhao, Wolfgang Nejdl, and Hai-Tao Zheng, editors, Semantic Web and Web Science, Springer Proceedings in Complexity, pages 391–396. Springer New York. doi:10.1007/978-1-4614-6880-6,34.
[33] Shu G., Avis N.J. (2008). Bringing semantic to vizualization services. Advances in Engineering Software 39. Pp.514-520.
[34] Steigmiller A., Liebig T., & Birte. G. (2014). Konclude: System description. Web Semantics: Science, Services and Agents on the World Wide Web, 27(1).
[35] Stojanovic L., Motik B. (2002) Ontology Evolution within Ontology Editors. In Conference on the Evaluation of Ontology-based Tools.
[36] Studer R., Benjamin V. R et FenselL. (1998). D. Knowledge Engineering: Principles and methods. Data & Knowledge Engineering 25, pp. 161-197.
[37] Suárez-Figueroa, M.C., Gomez-Perez, A., Fernadez-Lopez M. (2015). The NeOn Methodology framework: A scenario-based methodology for ontology development. Applied Ontology 10(2): 107-145.
[38] Suárez-Figueroa, M.C., Gomez-Perez, A., (2008). Towards a Glossary of Activities in the Ontology Engineering Field, LREC.
[39] Tsarkov D. & Horrocks I. (2006). Fact++ description logic reasoner: System description. In Proceedings of the Third International Joint Conference on Automated Reasoning, pages 292–297.
[40] Tudorache T., Nyulas C., Noy, N. F. & Musen, M. A. (2013). Webprotégé : A collaborative ontology editor and knowledge acquisition tool for the web. Semantic web, 4(1) :89–99. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3691821/.
[41] Uschold M., Gruninger M. (1996). Ontologies: Principles, methods and applications. The knowledge engineering review, vol. 11, no 2, p. 93–136.
[42] Vandenbussche P.Y., Vatant B. (2014). Linked Open Vocabularies. ERCIM news, 96:21–22.
[43] Vandenbussche P.Y., Atemezing G., Poveda-Villalon M., Vatant B. (2017). Linked Open Vocabularies (LOV): A gateway to reusable semantic vocabularies on the Web. Semantic Web 8(3): 447-452.
[44] Voigt M. & Polowinski J. (2011). Towards a Unifying Visualization Ontology. Technical Report – Technische Universität Dresden (3/2011).
[45] Kuhar S., Podgorelec, V. Ontology Visualization for Domain Experts: A New Solution, 2012 16th International Conference on Information Visualisation(IV), Montpellier, France France, 2012, pp. 363-369. doi:10.1109/IV.2012.67V.

©AfIA 2018

Categorization of B2B Service Offers: Lessons learnt from the Silex Use case

Molka Tounsi Dhouib[1 2] Catherine Faron Zucker[1] Andrea Tettamanzi[1]

[1] Université Côte d'Azur, Inria, CNRS, I3S, Sophia Antipolis, France
[2] Silex France
{dhouib, faron, tettamanzi}@i3s.unice.fr

Résumé

Dans le domaine de la recherche d'information et du traitement automatique du langage, la tâche de classification de textes est devenue une tâche cruciale. Dans cet article, nous partageons notre expérience de la classification de textes dans un contexte industriel et présentons une évaluation comparative de différents algorithmes de classification binaire et multi-label appliqués à des textes décrivant des offres de services, issus de la plateforme B2B SILEX pour la recommandation de prestataires de services. Nous montrons que dans certains cas pratiques comme celui que nous considérons, une représentation des données sous la forme de "bags of words" donne de meilleurs résultats de classification qu'une représentation réputée plus prometteuse par "word embeddings".

Mots Clef

Catégorisation de textes, Apprentissage automatique, Sac de mots, Plongement lexical

Abstract

In the domain of Information Retrieval and Natural Language Processing, text classification has become a crucial task. In this article, we share our experience of text categorization in an industrial context and we present a comparative evaluation of binary and multi-label classification algorithms applied to texts describing service offers, in the SILEX B2B platform. We show that for some use cases like the one we consider, a traditional representation of texts by "bags of words" gives better classification results than the promising representation by "word embeddings".

Keywords

Text categorization, Machine Learning, Bag of words, Word embedding

1 Introduction

The Silex France company offers a SaaS sourcing tool for identification of the service providers that are best suited to meet the service requests expressed by companies. The Silex platform is used by more than 5000 professionals to quickly identify and exchange with B2B service providers. Silex's ultimate aim is to provide a network of qualified companies to service buyers combined with the best functionality to consult it.

In the framework of a collaborative project between Silex and the I3S research laboratory, we aim at introducing semantics into the B2B platform in order to enable automatic reasoning on service requests and offers and improve the recommendation of service providers.

As a first step, we are interested in automatically categorizing the textual description of companies, service requests and service offers. With data's exponential growth, text categorization has become a crucial issue in Natural Language Processing (NLP). It consists in classifying texts according to their contents, into one or more predefined categories [2]. In recent years, the number of machine learning (ML) techniques that automatically generate text categorization has increased considerably [3].

In this paper we report some experiments we conducted to answer the Silex use case, using supervised ML techniques to classify Silex textual data into predefined categories. Meanwhile, we addressed the following questions:

(1) What is the best representation of Silex textual descriptions of service offers and requests to categorize it?

(2) What is the best ML algorithm to categorize Silex textual descriptions of service offers and requests?

This paper is organized as follows: Section 2 presents state-of-the-art text categorization methods. Section 3 introduces our approach of text categorization to answer the SILEX use case. Section 4 reports and discusses the results of our experiments of categorization of NL descriptions of service requests and offers. Section 5 draws some conclusions and directions for future work.

2 Related Works

Text categorization is the task of classifying data into a predefined set of categories. In other words, given a set of categories and a set of textual documents, text categorization is the process of automatically finding the correct category for each document [4].

2.1 Feature Vector Models

In the text categorization process, the first step consists in preprocessing textual documents, to convert them into feature vectors, a representation that can be automatically interpreted by machine. This step includes tokenization,

 ©AfIA 2018

stemming and removing of stop words before the creation of feature vectors.

Bag-Of-Words (BOW). is the most common feature vector model. In this model, the features are the frequencies of each word in the textual document. The feature space's dimension is the number of all different words in all documents [4]. The limitation of this model is that it ignores the semantic relations between words.

Word Embedding (WE). In order to overcome the weakness of BOW, a new model called Word Embedding (WE) has been successfully used in several NLP tasks. Word embeddings are projections in a continuous space of words that preserve the semantic and syntactic similarities between them [6]. There are many models that can produce a word embedding. In the following, we introduce the most important ones.
[7] proposes Word2Vec, a popular tool that produces word embeddings based on two models: *CBOW* and *Skip-gram*. *CBOW* is a Neural Network and log-linear model. It removes the non-linear hidden layer, and projects the contextual words on the same position. Word's prediction is obtained according to its past and future contexts. The process consists in computing the average of the contextual word vectors and running a log-linear classifier on the averaged vector.
Skip-gram is similar to *CBOW* and is also a Neural Network and log-linear model. Contrary to *CBOW*, *Skip-gram* predicts the contextual words given the current word [6].
Another model was introduced by [8] and is called *GloVE*. This model is based on global matrix factorization that calculates the co-occurrence of words in the corpus [6].
Finally, *FastText* [1] is a library for learning word representations and sentence classification. FastText published pre-trained word vectors for 294 languages, trained on Wikipedia. These vectors with dimension 300 were obtained using the skip-gram model [13].

2.2 Binary classification

After the construction of feature vectors, the second step of text categorization is the learning step. We present here some of the many learning algorithms for binary classification and multi-label classification.

Naive Bayes (NB) [9]. is a supervised, probabilistic algorithm based on Bayes theorem and the hypothesis of the independence between features. This algorithm is powerful due to the independence between features that ignores features' order, and consequently, the presence of a feature does not affect other features in classification tasks [14].

Support Vector Machine (SVM) [10]. is one of the most common and successful supervised classification algorithm used for text classification tasks [14]. This algorithm transforms training data into higher dimensions and searches for linear optimal separating hyperplane [1].

[1] https://github.com/facebookresearch/fastText

Neural Networks (NN). The NN is composed by many layers to perform text categorization. The perceptron is the simplest kind of a neural network and it has only two layers: the input nodes receive the feature values, the output nodes produce the categorization status values, and the link weights represent dependence relations. The NN text categorization process starts by loading feature weights into the input nodes; the categorization's final result is present in the output nodes after the propagation of the activation of the nodes forward through the network. The neural networks are trained by back propagation in order to minimize the error. In case of misclassification, the error is propagated back through the network and modifies the link weights [15].

2.3 Multi-label classification

Multi-label classification is an approach to classification problems that allows each data point to be assigned to more than one class at the same time. There are two multi-label classification methods: (i) problem transformation and (ii) algorithm adaptation [11]. Here we describe the most popular algorithms of each category.

Problem transformation. A first approach consists in transforming the multi-label problem into one or more single label classification problem. Among the methods adopting this approach, we can highlight:

- **The Binary Relevance method (BR)** is based on one-vs-all ensemble approach. BR transforms any multi-label problem into binary problems to predict the relevance of each label to a data point. All binary classifiers are then aggregated to form a set of relevant labels [11].
- **The label PowerSet method (LP)** considers the multi-label problem as a single multi-class classification problem. In order to create a transformed multi-class dataset, each combination of relevant labels is mapped to a class [11].
- **The Random K-label set (RAKEL)** aims at finding a better balance between BR and LP. RAKEL generates a series of label subsets and builds a label powerset model for each of them [11].

Algorithm adaptation. Algorithm adaptation extends specific algorithms to carry out multi-label classification.

- **Multi-label lazy algorithm (ML-KNN)** is one of the most famous adaptation algorithms. MLKNN uses the maximum a posteriori principle in addition to the K-nearest neighbour algorithm [11].
- **Multi-label decision tree (ML-DT)** is an extension of C4.5 decision tree algorithm. ML-DT allows the creation of multiple labels in the leaves, and chooses node splits based on a pre-defined muti-label entropy function [11].

- **AdaBoost** is based on the addition of simple classifiers to a pool and the use of their weights to define the final classification [16].

3 Experimental methodology

To answer the Silex use case, we conducted some experiments to compare the performance of the above described state-of-the-art feature vector models and learning algorithms for the categorization of textual descriptions of service offers or requests.

3.1 Dataset

Silex distributes a service to purchasing departments that support the sourcing system within the company. Silex users can describe their company (company description), their needs (service request description) and the description of the offers provided by the company (service offer description) in Silex web application. All descriptions are written in French.
The dataset considered in our experiments comprises 3188 company descriptions, 580 service offers descriptions and 155 service requests descriptions.
Six main categories are defined by Silex to classify these descriptions. One description can be categorized in one or more classes. The categorization phase is done manually by experts in the current version of the Silex platform. Table 1 shows the distribution of all the descriptions into the different categories. We observe that the distribution of our dataset is unbalanced, for instance the *Informatique* (Information Technology) class has 915 more companies descriptions than the *Services_industriels* (Industrial Services) class.

Categories	Companies descriptions	service offer descriptions	service request descriptions
Informatique	1043	147	42
Finance	393	40	13
Services généraux	929	149	65
Marketing	926	160	18
Ressources humaines	297	57	7
Services industriels	128	32	19

Table 1: Distribution of the textual descriptions in the corpus.

3.2 Implementation

In our implementation, we focused first on two methods to construct our feature vectors:

- The BOW representation is achieved by applying a traditional process of tokenization, stop words removal and lemmatization. Then we use the TFIDF to construct the matrix. This matrix will be used in different algorithms.

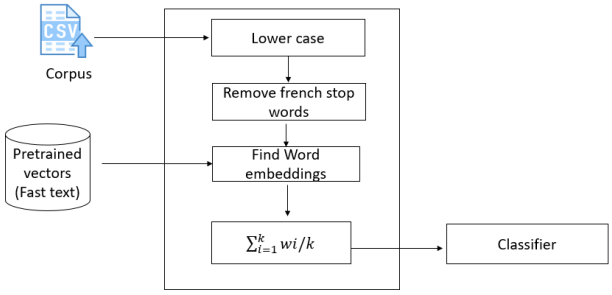


Figure 1: Building WE feature vectors representing the Silex textual descriptions from pre-trained word embeddings

- There are two methods to obtain Word embeddings representation: (i) build word embeddings from scratch using a model like word2vec or FastText. This approach needs a very large corpus to train these word embeddings; and (ii) use a pre-trained word embeddings in any desired language. For our case, we used the latter approach to obtain word embeddings because we do not have a large enough dataset. After studying the different pre-trained word embeddings, we decided to use the FastText model because it is the only model that provides French word embeddings pre-trained on a Wikipedia dump. This word embedding model [2] contains 1,152,449 tokens, which are mapped to vectors in a space with 300 dimensions.

  Our approach to produce the feature vectors representing the SILEX textual descriptions is inspired by [12] and [17]: As illustrated in Figure 1, the first step is to convert textual descriptions to lower case and remove all French stop words. Then the words are checked against the pretrained French vector. If the word exists in the dictionary, the vector representation of description is constructed by averaging the word embeddings vectors along each dimension for all the words in the description. If the dictionary does not contains the word, the zero vector is returned:

$$DescriptionWordEmbeddings = \sum_{i=1}^{k} \frac{w_i}{k}$$

  where k is the number of words in a description and $w_i \in \mathbb{R}^n$ denotes the word embeddings vector of the ith word. The vector representation of each description has exactly the same dimensions as the word embeddings vector w_i [17]. The vector representations of each description are then used as features and input for each classifier.

[2] https://fasttext.cc/docs/en/pretrained-vectors.html

©AfIA 2018

In order to test our feature vectors, we trained two methods of categorization: (i) three binary classification algorithms (NB, SVC, RNN) and (ii) three multi-label classification algorithms: two of them are problem transformation methods (Binary Relevance, Label PowerSet) and three of them are algorithm adaptation methods (Multi-label lazy algorithm (ML-KNN)). All these algorithms are implemented using scikit-learn[3], and executed on a machine with 8 GB of memory and an Intel Core i7-7500U CPU.
To deal with the unbalanced data, there are two main state-of-the-art approaches to adjust the distribution of dataset: (i) over-sampling adds copies of instances from the under-represented class. The most common technique of over-sampling is called Synthetic Minority Over-sampling Technique (SMOTE), (ii) under-sampling deletes instances from the over-represented class. In our implementation, we tested first a very large parameter with a grid function in scikit-learn to identify the best parameters for each classifier and the best method to deal with unbalanced data. We obtained the best results with SMOTE method for SVM algorithm, and RandomOverSampler method for NB. Table 2 shows all parameters used in binary classification and Table 3 shows all parameters used in multi-label classification.

Classifier	Unbalanced data	Parameters
NB	RandomOverSampler	
SGDClassifier	SMOTE	penalty=l2 alpha=0.0001
LinearSVC	SMOTE	C=0.1 penalty='l2' loss='squared hinge'
SVC	SMOTE	kernel='rbf' C=1000.0 gamma=0.0001
RNN	balanced	hidden size=32 Activation('relu') Activation('sigmoid')

Table 2: Parameters used for binary classification

Classifier	multi-label methods	Parameters
LSVC	Binary Relevance	C=1, penalty='l2'
LSVC	Label Powerset	C=1, penalty='l2
MLkNN	Algorithm adaptation	-

Table 3: Parameters used for multi-label classification

4 Experiments and discussion

4.1 Evaluation procedure

The experiments were conducted using the 5-fold cross-validation methodology. To evaluate our experiments, we calculated the classification duration and various evaluation measures inspired by [18] and described in the following.

[3]http://scikit-learn.org/stable/

Binary algorithms. To evaluate binary classification algorithms, we used the **F1 score** measure, defined as the harmonic mean between precision and recall:

$$F1score = \frac{1}{N}\sum_{i=1}^{N}\frac{2\times|h(xi)\cap yi|}{|h(xi)+yi|}$$

Multi-label algorithms. To evaluate multi_labels algorithms we used the following metrics:

- **Accuracy** allows to compute the percentage of correctly predicted labels among all predicted and true labels. It is defined as follows:

$$Accuracy(h) = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{h(xi)\cap yi}{h(xi)\cup yi}\right|$$

- **Micro-precision** is defined as the precision averaged over all the example/label pairs:

$$Micro_{precision} = \frac{\sum_{j=1}^{Q} tpj}{\sum_{j=1}^{Q} tpj+\sum_{j=1}^{Q} fpj}$$

  where tpj, fpj are the number of true positives and false positive for label λ_j .

- **Micro-recall** is defined as recall averaged over all the example/label pairs:

$$Micro_{recall} = \frac{\sum_{j=1}^{Q} tpj}{\sum_{j=1}^{Q} tpj+\sum_{j=1}^{Q} fnj}$$

  where fnj is the number of false negatives for the label λ_j.

- **Micro-F1** is the harmonic mean between micro-precision and micro-recall and is defined as:

$$Micro-F1 = \frac{2\times micro_{precision}\times micro_{recall}}{micro_{precision}+micro_{recall}}$$

- **Hamming loss** allows to evaluate how many times an example-label pair is misclassified. The performance is perfect when the value of this metric is equal to 0. Hamming loss is defined as

$$Hamming_loss(h) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{Q}|h(xi)\Delta yi|$$

  where Δ is the symetric difference between two sets, N is the number of examples and Q is the total number of possible class labels.

4.2 Results and Analysis

The results of binary classification algorithms reported in Table 4, Table 5 and Table 6 show that there is not a significant difference between the results with a BOW representation and the results with a word embeddings representation. For example, for category "Informatique" , we obtain the same precision value that equals to 0.82. The recall value equals to 0.82 for BOW representation, and is a bit better compared to WE recall value that equals to 0.79. The classification processing time is much better for WE (0.48 seconds) than for BOW (1572 seconds). This is due to the BOW expansive phases of lemmatization and matrix building compared to the vector representation of word embeddings.

Table 4: Best classifier of binary classification of companies descriptions

Category	Method	P	R	F1	Time
Informatique	LSVC + WE	**0.82**	0.79	0.80	0.48
	SGD+BOW	**0.82**	**0.82**	**0.82**	1572
Finance	LSVC+WE	0.90	**0.87**	**0.88**	0.21
	NB+BOW	**0.92**	**0.87**	**0.88**	1793
Services généraux	LSVC+ WE	0.87	0.84	0.85	0.4
	NB + BOW	**0.90**	**0.89**	**0.89**	1925
Marketing	LSVC+WE	0.82	0.80	0.81	0.41
	SGD + BOW	**0.85**	**0.85**	**0.85**	1517
Ressources humaines	LSVC + WE	0.91	0.83	0.86	0.19
	NB + BOW	**0.92**	**0.84**	**0.87**	1904
Services industriels	LSVC+WE	0.94	0.89	0.91	0.05
	NB+BOW	**0.97**	**0.94**	**0.95**	1914

Table 5: Best classifier of binary classification of service offer descriptions

Category	Method	P	R	F1	Time
Informatique	RNN+WE	**0.80**	**0.78**	**0.79**	8.53
	RNN +BOW	0.78	**0.78**	0.78	332
Finance	SVC-RBF+ WE	0.82	0.87	0.84	0.083
	SGD + BOW	**0.91**	**0.90**	**0.86**	323
Services généraux	RNN+ WE	**0.82**	**0.81**	**0.81**	6.97
	RNN+BOW	0.79	0.79	0.79	348
Marketing	SGD+WE	0.80	0.77	0.78	0.028
	NB + BOW	**0.90**	**0.90**	**0.90**	459
Ressources humaines	RNN + WE	**0.90**	**0.91**	**0.90**	7.42
	NB + BOW	0.87	0.87	0.87	419
Services industriels	SVC-RBF+ WE	0.91	0.92	0.91	0.09
	SVC-RBF+BOW	**0.96**	**0.97**	**0.96**	347

We can draw similar conclusions when comparing multi_label classification algorithms, with a slight advantage to BOW compared to word embeddings in terms of accuracy and hamming_loss but not in term of time.
If we compare binary classification and multi_label classification algorithms, we can see that the results of the first alogrithm are better than the second one. These results are as expected because first we don't have a very large corpus, and second when we use a binary classification we can play

Table 6: Best classifier of binary classification of service request descriptions

Category	Method	P	R	F1	Time
Informatique	NB +WE	**0.82**	**0.77**	**0.79**	0.006
	NB+ BOW	0.79	0.74	0.76	101.31
Finance	RNN +WE	**0.97**	0.48	0.62	2.44
	LSVC + BOW	0.94	**0.94**	**0.92**	84.82
Services généraux	NB+ WE	**0.74**	**0.74**	**0.74**	0.006
	NB + BOW	**0.74**	**0.74**	**0.74**	103.49
Marketing	RNN + WE	0.73	0.77	0.74	2.74
	NB +BOW	**0.89**	**0.81**	**0.84**	104.10
Ressources Humaines	NB+ WE	0.88	0.94	0.90	0.006
	BOW	-	-	-	-
Services industriels	RNN + WE	0.91	0.90	0.88	2.62
	LSVC+BOW	**0.94**	**0.94**	**0.94**	79.07

Table 7: Best classifier of multi_label classifications.

Classifier	Feature	Micro-F1	Accuracy	Hamming loss	Time
MLKNN	WE	0.66	0.53	0.12	7.00
	BOW	0.72	0.57	0.10	1945
BR	WE	0.61	0.35	0.17	6.54
	BOW	0.73	0.52	0.10	1850
Label Powerser	WE	0.54	0.39	0.19	9.44
	BOW	0.69	0.51	0.13	1945

on the classifier parameters for each category to enhance its performance, which is not the case for multi_label classification. The different binary classification algorithms perform more or less well in categorizing our dataset using the same representation. For example, only the LSVC algorithm categorized correctly this text as "informatique" *"Ascot accompagne les dirigeants et cadres de PME dans la définition, la maitrise et le suivi de leur transformation numérique. Plus aucune organisation ne peut ignorer l'impact global sur son métier, ses outils et ses processus. Notre expérience, de plus de 20 ans, en Conseil, Assistance technique, Développement, Ingénierie, ... est entièrement dédié à nos clients. Sous contrat ou à la carte, nos prestations leur apportent valeur et sérénité."*.

We obtained also different results when comparing the BOW and WE representations. For example, using the same LinearSVC algorithm, this text *"Agence WebMarketing 360 de nouvelle génération exploitant de nombreux leviers marketing pour optimiser vos campagnes digitales & booster votre chiffre d'affaires Kalipseo développe pour vous les stratégies digitales les plus pertinentes et s'emploie à promouvoir l'image de votre marque."* was wrongly categorized as "Informatique" when using the WE representation, but the categorization was right when using BOW representation.

As a conclusion, to best answer the Silex use case, we decided to use a binary classification algorithm with a word embedding representation. Table 8 shows the algorithms that are selected to classify each category of Silex data.

©AfIA 2018

Table 8: Selected algorithms for Silex Company

Category	Description	Representation	Algorithm
Informatique	Companies	WE	NB
	Service offer	WE	RNN
	Service provider	WE	NB
Finance	Companies	WE	RNN
	Service offer	WE	SVC-RBF
	Service provider	WE	RNN
Services généraux	Companies	WE	NB
	Service offer	WE	RNN
	Service provider	WE	NB
Marketing	Companies	WE	RNN
	Service offer	WE	SGD
	Service provider	WE	RNN
Ressources Humaines	Companies	WE	NB
	Service offer	WE	RNN
	Service provider	WE	NB
Services industriels	Companies	WE	RNN
	Service offer	WE	SVC-RBF
	Service provider	WE	RNN

5 Conclusion

In this paper we reported the results of the experimental evaluations of various vector feature models and machine learning algorithms conducted to answer the real-world use case of the Silex company: how to categorize textual descriptions of service offers and requests with the ultimate goal of recommending service providers that better answer services requests. We compared the two main state-of-the-art feature vector models. A BOW representation is a bit better than a WE representation regarding the evaluation measures. This is due to the use of a generic pre-trained vector for WE descriptions, that is based on Wikipedia dump and that does not cover well our data set. On the other hand, a WE representation has the advantage of being less time consuming because of the additional lemmatization and matrix building phases specific to the BOW representation. Based on these results, we choose to use classifiers with word embeddings data for Silex company as shown in Table 8.

As future work, we aim to define a specific B2B pre-trained vector that best covers the Silex dataset instead of using a general reference like Wikipedia. We also aim to study a combined approach based on both word embeddings and knowledge engineering approaches. The objective is to improve the representation of texts by introducing semantics to increase the performance of algorithms. To achieve this, we are going to use our built ontologies to represent the sourcing domain, and we are going to enrich the original texts with additional information based on semantics relations between concepts such as generalization or specification.

6 Aknowledgement

This work is supported by the ANRT, CIFRE I3S-SILEX 2016/0947

References

[1] Patra, Anuradha and Singh, Divaka A survey report on text classification with different term weighing methods and comparison between classification algorithms *International Journal of Computer Applications*, volume.75, pp.7,2013,Foundation of Computer Science.

[2] Sadiq, Ahmed T and Abdullah, Sura Mahmood, Hybrid intelligent technique for text categorization *Advanced Computer Science Applications and Technologies (ACSAT), 2012 International Conference on*, pp.238-245,2012,IEEE.

[3] Tan, Ah-Hwee and Lai, Fon-Lin, Text categorization, supervised learning, and domain knowledge integration *To appear, proceedings, KDD-2000 International Workshop on Text Mining, Boston*, Vol.20,2000.

[4] Feldman, Ronen and Sanger, James,The text mining handbook: advanced approaches in analyzing unstructured data,2007,Cambridge university press.

[5] Sriram, Bharath and Fuhry, Dave and Demir, Engin and Ferhatosmanoglu, Hakan and Demirbas, Murat, Short text classification in twitter to improve information filtering, Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval,pp.841–842, 2010, ACM

[6] Ghannay, Sahar and Favre, Benoit and Esteve, Yannick and Camelin, Nathalie, Word Embedding Evaluation and Combination,LREC,2016

[7] Mikolov, Tomas and Chen, Kai and Corrado, Greg and Dean, Jeffrey,Efficient estimation of word representations in vector space,arXiv preprint arXiv:1301.3781, 2013

[8] Pennington, Jeffrey and Socher, Richard and Manning, Christopher,Glove: Global vectors for word representation,Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP),pp.1532-1543,2014

[9] Schneider, Karl-Michael, Techniques for improving the performance of naive bayes for text classification, Computational Linguistics and Intelligent Text Processing, pp.682-693, 2005,Springer

[10] Cortes, Corinna and Vapnik, Vladimir,Support-vector networks, Machine learning, Vol.20,3,pp.273-297,1995,Springer

[11] Pakrashi, Arjun and Greene, Derek and MacNamee, Brian,Benchmarking Multi-label Classification Algorithms,24th Irish Conference on Artificial Intelligence and Cognitive Science (AICS'16), Dublin, Ireland, 20-21 September 2016,2016, CEUR Workshop Proceedings

[12] Bayot, Roy and Gonçalves, Teresa,Author Profiling using SVMs and Word Embedding AveragesâĂŤNotebook for PAN at CLEF 2016,2016,CEUR

[13] Bojanowski, Piotr and Grave, Edouard and Joulin, Armand and Mikolov, Tomas,Enriching Word Vectors with Subword Information, arXiv preprint arXiv:1607.04606,2016

[14] Mohare, Punam Kishor Itkar, A, A Survey on Different Types of Approaches to Text Categorization,V5,nternational Journal of Emerging Trends & Technology in Computer Science

[15] Bassil, Youssef, A Survey on Information Retrieval, Text Categorization, and Web Crawling,rXiv preprint arXiv:1212.2065,2012

[16] Benbouzid, Djalel and Busa-Fekete, Róbert and Casagrande, Norman and Collin, François-David and Kégl, Balázs,MultiBoost: a multi-purpose boosting package,V.13,pp.549-553,Journal of Machine Learning Research,2012

[17] Yang, Xiao and Macdonald, Craig and Ounis, Iadh,Using word embeddings in twitter election classification,arXiv preprint arXiv:1606.07006,2016

[18] El Kafrawy, Passent and Mausad, Amr and Esmail, Heba,Experimental comparison of methods for multi-label classification in different application domains,V.114, N.19,Foundation of Computer Science, International Journal of Computer Applications,2015

[19] Sokolova, Marina and Japkowicz, Nathalie and Szpakowicz, Stan,Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation,V.4304, pp.1015-1021,Australian conference on artificial intelligence,2006

©AfIA 2018

AI techniques to analyse a social network on text, user and group level : application on Galaxy2

Guillaume Gadek[1,2], Alexandre Pauchet[2], Stéphan Brunessaux[1], Khaled Khelif[1], Bruno Grilhères[1]

[1] Airbus, 78990 Elancourt, France
[2] Normandie Univ, INSA Rouen, LITIS, 76000 Rouen, France

guillaume.gadek@litislab.eu

Résumé

Les Réseaux Sociaux Numériques ont pris une place prépondérante dans l'espace informationnel, et sont souvent utilisés pour la publicité, le suivi de réputation, la propagande et même la manipulation, que ce soit par des individus, des entreprises ou des états.
Alors que la quantité d'information rend difficile son exploitation par des humains, des solutions sont accessibles pour aider les preneurs de décision, solutions provenant notamment de techniques d'IA pour obtenir du renseignement à partir des messages postés, pour qualifier les comportements des utilisateurs, et pour identifier la structure sociale émergente.
Dans cet article, nous illustrons l'exploitation de telles techniques sur un réseau social très particulier, nommé Galaxy2, caché sur le Dark Web.
Nous proposons une analyse de 1000 jours d'activité, extrayons des thématiques et découvrons des acteurs-clés. Nous procédons ensuite à une étape de profilage des comptes utilisateurs, basée sur leur comportement. Finalement, nous introduisons des scores d'influence et de cohésion pour les groupes d'utilisateurs, ce qui aide à les caractériser et à les évaluer.

Abstract

Online Social Networks have taken a huge place in the informational space, and are often used for advertising, e-reputation, propaganda, or even manipulation, either by individuals, companies or states.
As the quantity of information makes the human exploitation difficult, solutions to support the decision makers can only come from the use of AI techniques to extract intelligence from posted messages, to qualify the user behaviours, and to identify the social structure.
In this article, we illustrate how to exploit such techniques on a very peculiar social network, named Galaxy2, hidden in the Dark Web.
We propose an analysis of 1000 days of activity using NLP techniques to find the most interesting topics and to discover key actors. We then proceed with a ML-based profiling of the user behaviours. Finally, we introduce influence and cohesion scores for groups of users, which help their characterisation and evaluation.

Keywords

Deep Web, Social network analysis, User profiling, Group discovery, Graph analysis

1 Introduction

Web 2.0 and the social media have introduced a new paradigm in the informational space. Its specificities as a broadcast media enable anyone to be a new source of information, entertainment, or propaganda, at will. It gives power to the consumer, signalling defects on products ; it gives the opportunity to the citizen to directly chat with his/her representatives ; overall, the social web enables the user to be *more* connected and informed.
However, the presence of threats on Twitter and other social platforms is attested : the jihadist group Daesh is said to benefit from the support of 100,000 accounts on Twitter [1] ; Russia is suspected by the FBI of interferences in the 2016 US election ; in almost every election, worldwide, the intensive use of deception has been widely observed during the last years. Automation enables any party to produce and diffuse content quite effectively, if we consider the appearance of malicious accounts in the classic media a victory.
The diversity of threats implies a multiplicity of the tools to tackle all of the aspects of the problem. The quantity of messages exchanged in an Online Social Network can be astounding, with almost 500 million new tweets each day. As is possible the automation of their publishing, it is also possible to automate the reading. Opinion mining and topic detection enable the analyst to grasp the trend on a mass of texts, providing a classification along various keywords and characteristics.

©AfIA 2018

The user accounts can be classified according to various taxonomies. In this article we distinguish the profile (who the user claims to be), and the behaviour (how s/he performs actions). Unsupervised machine learning helps here to cluster the users along similar types, and enables the analyst to recognise similar situations.
The never ending list of "friendship" connections, as well as the list of interactions between user accounts, is fertile to grow graphs. This kind of graphs can be used to easily compute social influence scores [2], and to detect one of the most interesting features of a social network : people naturally gather in groups, where there is more continuous interaction. To detect these groups from graphs, we rely on community detection algorithms [3] ; specific measures help the analyst to characterise the impact and the pertinence of such groups with regard to his preoccupations [4].
Galaxy2, on Tor, is a strange place. Created in 2015 to replace a predecessor social network (simply called *Galaxy*), it has been branded as the most frequented social network on Tor. In this article, we show the added value of a variety of tools to scan this very peculiar social network, analyse the topics of discussion during Galaxy2's uptime, detect its key-actors, and discover and characterise its communities. Finally, using all this information, we perform a case study, showing the interest of dedicated tools for social network analysis.
This paper is organised as follows : Section 2 presents the website and the context of TOR, Section 3 presents the tools used, Section 4 shows the case study ; Section 5 proposes a discussion about our results and concludes this article.

2 Presentation of Galaxy2

2.1 The TOR network

Partly developed through DARPA funding in the 90s, Tor, *The Onion Router*, has been launched in 2002. Aiming to bring anonymity to the data flow, it relies on an encryption and routing protocol, named *onion*, to hide the content of the packets to the transiting servers.
Tor is used either to anonymously access the clear web, as well as to access the so-called *darknet* : some websites are only accessible through the *onion* protocol, protecting the host. Well-known examples include a Wikileaks portal, and a disrupted illegal marketplace, *the silk road*.
As anyone can use the Tor network and enjoy the increased privacy, it helps to breach censorship and can be used by journalists, political activists or anyone else. Unfortunately, Tor is also used to host illegal content, such as hacker forums, drug marketplaces, dark forums, porn and pedo-pornography [1].
However, due to the perceived complexity of use and the strong incentive to link one's Internet profile with his real life (shopping, all kinds of services...), the global usage of Tor is not very high. Moreover, some countries deploy firewalls, laws and rules to prohibit this kind of tools.

2.2 Galaxy2 : history and main features

Founded in 2015 after the disruption of a previous Tor-hosted social network ("Galaxy"), Galaxy2 is based on an open source framework named *elgg*, enabling to build small social websites. According to its anonymous founder, *Lameth*, *"The server broke down and your terrible host here (me, not the current host, mind you !) hadn't been keeping regular backups off the server."* The service is shut down since end of October 2017 [2].
The main features of Galaxy2 include The Wire, a space for microblogging posts ; blogs, polls and pages ; image and file sharing. Because of the anonymity introduced by the Tor network, the users are less keen to use their real names, and do not expose personal data. Thus, only the private direct messages are kept private, and it was possible to consult all the posts, friendship connections, images and comments performed by the users over the almost three years of uptime.
The collection of our corpus covers a time range from Galaxy2 creation, on 9th of January, 2015 to the 22nd of September, 2017. It includes all traces of activity logged and then publicly available, but does not consists of personal profiles nor photos, excluded of the collect.

3 Tools for social media analysis

3.1 Focused on text processing

The first axis of analysis deals with text processing. The tremendous amount of new messages emitted through the Online Social Networks pushed the development of machine text understanding and summary. Moreover, as users often express their feelings and opinions about products, facts and events, one of the most necessary tasks is opinion mining, which can briefly be explained as the automatic computation of a couple *(topic, sentiment)* matching a message [5].
The sentiment, or polarity, enables the evaluation of the support or reject of an entity by the author of a message. It can be either computed specifically with a topic, using a learning corpus, or generically using linguistic resources. Such a resource, SentiWordNet [6], associates polarities with sets of synonyms, sometimes considering contextual associations to modify the polarities according to the surrounding sentence. An ana-

1. https://www.fbi.gov/news/stories/playpen-creator-sentenced-to-30-years
2. https://socialmediaalternatives.org/archive/collections/show/10

lyser, Vader [7], has been made available as an open source module, and compiles various resources and a classifier to compute sentiment scores, at document level.

Topic detection can be performed either through a supervised or an unsupervised method. An analyst can describe a topic using a few keywords, or rely on text clustering to let an algorithm cluster similar texts together, based on the semantics. Instead of matching pre-established categories, this approach adapts well to new corpora, enabling the analyst to grasp the topic distribution.

Often cited as a reference for text clustering, LDA (Latent Dirichlet Allocation) [8] can briefly be resumed as a dimensionality reduction applied on a matrix where each line consists of the tf*idf representation of a text of the training corpus. It is said adapted to short text topic classification and can be associated with a module of visualisation of the detected topics, to better depict their distribution [9].

3.2 User profiling

One of the main tasks of social network analysis, or social listening, aims to spot the key actors, i.e. the most important users of the network. The focus is commonly directed either towards the *influence* [10, 11], defined as the ability of a user to trigger actions from other users, or towards community detection [12, 13].

Influence is commonly defined as the ability to *engage*, i.e. to make the other users do something (typically, sending a message or clicking on a URL). A complete method has been exposed to build a social graph from Twitter follow links [2]. This enables the computation of a daily PageRank, as a score of influence. However, using the public free API, a very long time (counted in months) is needed to obtain the complete social graph. To evaluate influence, various indicators have been compared, among which the number of retweets, number of followers and PageRank computed on a graph of friendship links [10, 14].

Sometimes, the focus is not set on the most influential accounts, but on a general profiling to help the analyst understand "who" is a given account : what its typical behaviour is. Machine learning tools can answer this question using numeric features, through clustering or classification. When the data represent a graph, specific algorithms, such as RolX [15], enable a tagging of the nodes based on their graph features.

3.3 Group detection

Studies in psychology highlight the power of the "group" as a structure, and the variety of its impacts on individual behaviour [16, 17]. In computer science, a study on polarisation of the attitudes on online debates concludes similarly : groups can be considered as influencers, too [18].

Detection. To detect these groups, the main approach relies on community detection from friendship relation graphs. A community is a set of nodes, in a graph, more densely linked than a random graph. An intuitive way to discover groups is to look for communities in a graph built from the social network data.

However it has been pointed out that the social graph ("follow" graph) does not illustrate user interaction [19]. As a response, random walks community detection have been applied on a *graph of interactions* between Twitter users (retweets and mentions) [20].

To help choosing the best community detection method, a comparison of various state-of-the-art algorithms [21] on large networks (around 400k nodes) has been performed between FastGreedy [22], InfoMap [23] and Louvain/Blondel [3]. In this experiment, the Louvain method seems to be, by far, the fastest, for similar results in terms of quality.

Characterising the strength of the groups. Topological graph features such as the modularity value or the density of the detected groups can be considered to evaluate a community detection algorithm. A comprehensive review of scoring functions for community evaluation details the following measures [24].

Internal density (d) follows the intuition that communities are more dense than a random set of nodes. However, for large, non-complete communities, its value may be very low. On the other side, the smallest possible communities (pairs of nodes) have an internal density of 1.

Triad participation ratio (TPR) evaluates the number of nodes belonging to triads, or triangles, in community S. A value of 1 means that the group is strongly internally linked. Communities should include numerous triangles, showing interaction between the members.

Conductance (c), evaluates the quantity of edges linked with the other communities [25]. It illustrates the community behaviour, ranging from 0 (not linked towards the exterior) to 1 (strongly connected to other groups).

Modularity (Q) focuses on the number of internal edges (which link members of a community), compared to a "normal" quantity if the graph was random. Introduced by [26], a high modularity Q, close to 1, denotes a good partitioning at the graph level.

These topological metrics enable the analyst to check whether a community is sufficiently connected internally and/or with its boundaries : internal interaction and reach towards the exterior are measured.

Characterising the topical cohesion and relevance of the groups. To measure each of the obtained groups based on the topics expressed by the users, two topical metrics ξ and ρ, inspired from machine learning (ML) precision and recall, have been proposed [4].

©AfIA 2018

These two ML inspired topical measures, which enable to better attribute a weight of the importance of a community on a given topic (ρ), and to evaluate the internal cohesion of a group (ξ), make the link between the graph and the semantic of the texts. Another measure ($\theta f.igf$) [4], similar to the natural language processing domain $tf * idf$, evaluates the specificity and the importance of a topic in a group.
The analysis of these scores yields interesting insights on a social and textual corpus. For marketing purposes, one can identify then target a customer group, helping to induce a positive opinion about a new product.

3.4 Discussion and contribution

Although much work has been done to tackle specific challenges in the domain of Online Social Networks, only a few systems and frameworks propose a complete analysis of a website. Most of the platforms focus on the graph and miss the texts, such as UCINET, NetDraw, Pajek and ORA, presented in [27], or a work on graph characterisation [28].
As in a mirror, a set of efficient textual analysis tools is available, enabling specific platforms for news monitoring [29] or for disaster management and evaluation [30]. However they lack of user profiling features and social graph exploration and analysis.
In this paper, we propose a full-stack processing pipeline, integrating various open-source modules and providing our own model for specific aspects. We perform the textual analysis on the text, model the users and propose an influence ranking, and conclude with exploration through graph analysis and community detection and characterisation.

4 Case study

In this section, the Galaxy2 network is exposed through the explanation of the collected data type. Then, results are proposed to analyse the texts, the user accounts and the groups, which are an emerging social feature.

4.1 Types of actions logged

Fortunately for the privacy, not every action can be retrieved. In the collected data, actions are regrouped to better grasp their diversity : first are the comments. Users can add some words to react to the other's actions : publication of photos, files, new pages or polls, and of course blog posts. A total of 4032 comments on blog posts are stored ; much less (375) on photos.
A second kind of actions gathers the upload of files and images : those are deemed to be shared, and frequently include pictures from the Anonymous. As Tor and Galaxy2 philosophies are not based on personal holidays pictures publication, the quantity of images is quite low (around 600 pictures, in total).
Creation of pages, groups and polls and their modification (new topics of discussion) are seen as an aspect of the network life. This step of page creation is to be linked with the cluster of votes, replies and publication on said pages/groups. Overall, 1944 blog posts have been emitted, and 548 groups created.
Blogging or "micro"-blogging features were the predominant ones, with 29,000 posts on The Wire, the message feed. Message size was not limited, though frequently short.
Finally, the last cluster includes some connection information : account creation notifications (19233 "joined the site"), link establishment between a user and a page, a group or another user (61027 friendship links).

4.2 Text : topic and sentiment repartition

LDA for topic detection. LDA is commonly used with a high number of topics ; a number of 200-500 is usually recommended as a first try [3]. This method clusters documents with similar $tf.idf$ representations, resulting in a predetermined number of emerging topics, seen simultaneously as sets of documents and as $tf.idf$ vectors. As a trade-off between computation time and representativeness of the topics, a number of 40 topics is chosen. Compared to other values, this choice results in a distribution of the documents over the topics, also limiting the number of too-close topics.

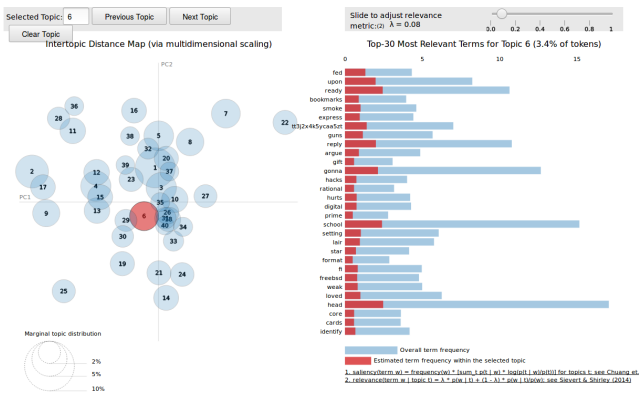


FIGURE 1 – Visualisation of the topics

Figure 1 shows a tool, proposed by [9], which enables to navigate along the 40 detected topics through the word frequencies in a given topic (in red), compared to word frequencies over the whole corpus (in blue), in the right panel. A principal component analysis (PCA) is performed to place the topics in a 2D-space, in the left of the diagram. The area of the circles is proportional to the quantity of documents clustered in the topics. Thanks to this tool, the analyst can visualise the topic distribution of, here, more than 30,000 docu-

3. http ://radimrehurek.com/gensim/models/ldamodel.html

ments, while having the ability to see which words are more frequent or more relevant in each of the topics.

Sentiment exploitation. Each text is also analysed with regard to the *sentiment*, representing the overall polarity of a text, as perceived by a reader. The implementation relies on the *Vader* module [7]. Even if some of the sentiment predictions may be inaccurate, it results in a global view. For instance, Figure 2 proposes to compare the polarity of the texts containing some keywords : *onion, Tor, Snowden, Trump and Syria*[4]. Instead of an average value, and as one message is enough to set either a *max* or *min* sentiment value, we represent it as a boxplot, showing the statistical distribution. *Onion* is globally well perceived, with positive messages. Edward Snowden seems to enjoy a good reputation, while Galaxy2 members are more mitigated about Donald Trump. Finally, the polarity is more negative about Syria, because these messages are susceptible to mention the current conflict.

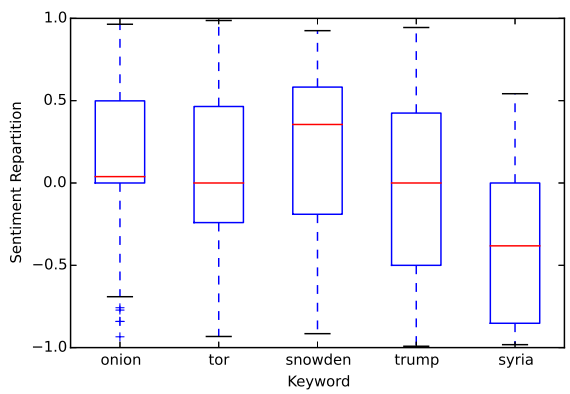


FIGURE 2 – Sentiment distribution for some keywords

The sentiment view can also be exploited to characterise a user activity. As an example, Figures 3 and 4 expose the sentiment repartition of the messages of two active authors : *XL33t* and *Fenris*. On both figures, an artefact, for a sentiment of 0.0, exists, representing the messages where no polarity could be extracted by the analyser (no known sentimental word). However, the repartition shows some signals. On the first figure, *XL33t* seems to mostly emit very positive messages. The second author, *Fenris*, covers the whole range of sentiment, avoiding to publish too positive messages. This view let the analyst grasp one aspect of a user behaviour, as well as to estimate his activity[5].

4. No distinction has been made between uppercase and lowercase.

5. Thanks to the number of messages, indicated on the y-axis.

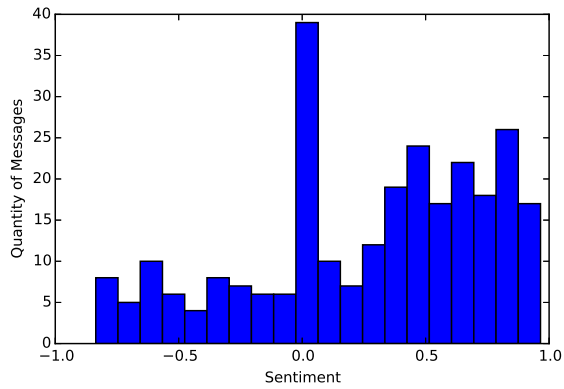


FIGURE 3 – Visualisation of the sentiment emitted by the user *XL33t*

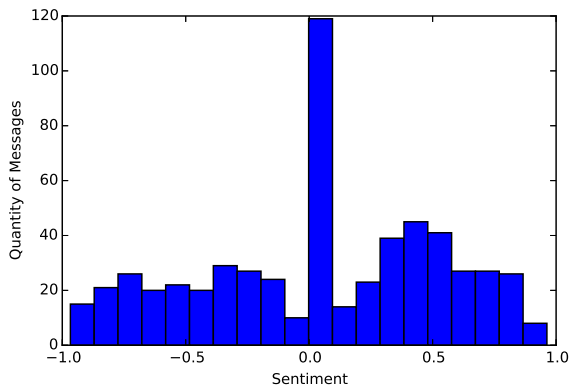


FIGURE 4 – Visualisation of the sentiment emitted by the user *Fenris*

4.3 Influential users and types of behaviour

Top-5 key users. Influence can be measured through a variety of prisms. In a first glance on a social network, one can look for the key actors, detecting a "top5". Table 1 compares the most connected users (having the highest number of friends, noted #Friends) ; the most popular users (the most mentioned, noted #Mentions) ; the users of reference (a score based on the mentions graph, in the column *Sc(Mentions)*, explained in the following paragraph) ; and the most active ones (performing the highest number of actions on Galaxy2).

Inspired from the literature [10, 14], a score of influence is computed from G_M the graph of mentions. The intuition is the following : a mention gives some value, some social capital to the mentioned user ; this value is higher if the author of the mention already has some influence him/herself. The PageRank algorithm also follows this intuition and has been shown to

©AfIA 2018

empirically match this representation of influence. In consequence, ScoreMentions is the ranking resulting from the Pagerank scores of each user-node in G_M.

Table 1 – Top5 of the users along three dimensions

#Friends		#Mentions		Sc(Me.)	#Actions	
XsyntaX	8187	XsyntaX	514	Oxyy	XsyntaX	7974
prozac	3182	Oxyy	164	XsyntaX	kheper	1846
Spooky	3012	Lameth	121	Lameth	prozac	1783
xl33t	2140	Fenris	119	ChatTor	ChatTor	1583
kheper	1922	cpnnemo	96	MahaKali	Spooky	1569

In a few words about Table 1 : *XsyntaX* outperforms every other account, with 7974 actions performed during the scope of the study ; the second, *kheper*, performed 1846 actions. *Lameth* is the founder of this social network ; as such he is often mentioned, either for thankful messages for hosting, or for administration purposes. *ChatTor* provides a chat service through Tor ; this account promotes the service and publishes news and updates.
Although some names are present on every influence score *Top5*, they do not all indicate the same kind of influence. *Lameth* is mentioned a lot, due to his status of host. However, he did not intend to be the friend of everyone else, or to be the most active. *Xl33t* connected to many accounts, but is not mentioned as much as the other influential users. Influence and key-actors take many different forms, while only the symptoms can be measured.

Types of behaviour. The social network analyst cannot be expected to know who is every user on a platform. More precisely, the data of interest can be narrowed along a few aspects to grasp the essence of the behaviour.
Biography contains the user identity : username, ID, date of creation and other available data. *Style* explains how the user writes : length of messages and quantity of punctuation usually make a difference, and can be combined with the main topic of interest and global sentiment polarity. The *social* aspect details the number of friends, of mentions, and the resulting influence of the user, both from the graph of friendship and the graph of mentions. *Media* groups the features explaining the type of actions and objects posted (be it photos, texts, etc). Finally, the *temporal* aspect covers the global rhythm of publication of the user : average number of posts on TheWire per day, and average number of actions per day.
All these numerical features enable the repartition of the users along a few typical clusters, whose reduced number (commonly 4 to 6, depending of the data) enables the analyst to get used to this unsupervised labels.
Technically, this repartition is made in a few steps. First the data is cleaned, so as to avoid zero divisions and exponential distribution of the data : some features are converted to their log-values to reduce dispersion, and normalised to follow a 0-mean, 1-standard deviation distribution. A PCA reduces the dimension of the problem, switching from 34 variables to 5 and still keeping most of the dataset variance. Finally, a k-means finds the groups ; a number of 4 groups is a good trade-off between dataset dispersion and cluster shape.

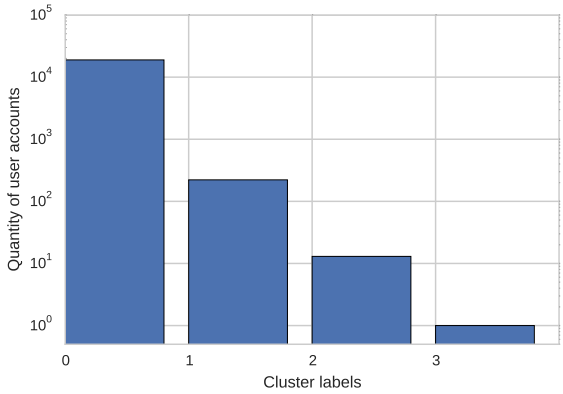


Figure 5 – Repartition of the user profiles

Figure 5 shows the repartition of the users along the 4 types of profiles. The first type (0) is the most common, including more than 90% of the profiles : most of the accounts are created, perform one or two actions, and then are forgotten. A second type, profile (1), characterises the regularly active users, even though they are not the central users of the network. Finally, profile (2) and the sole user in profile (3) are the central users, very active, those who produce most of the contents on the network. A similar repartition had already been remarked on networks such as Twitter [10].

To picture the shape of the profile clusters, Figure 6 represents each of the 19,177 Galaxy2 users along the first two dimensions of the PCA. These dimensions are linear combinations of the behavioural features, and consequently are quite abstract. However, positions near (0,0) are linked with very low levels of activities (the red cluster (0) being the most populated). In blue, cluster (1) shows the many different ways to be a standard active user. The green cluster (2) is very dispersed but matches high levels of publications, be it by posts, comments or creation of pages. Finally, the purple cluster (3) is constituted of only one account, *Spooky*, performing an average of 2.5 actions each day.

When the analyst reviews the posts, thanks to this step of profiling, he can quickly evaluate if the messages comes from well-established, influential users, or whether a given topic is dominated by usually inactive accounts, or show a unusual repartition.

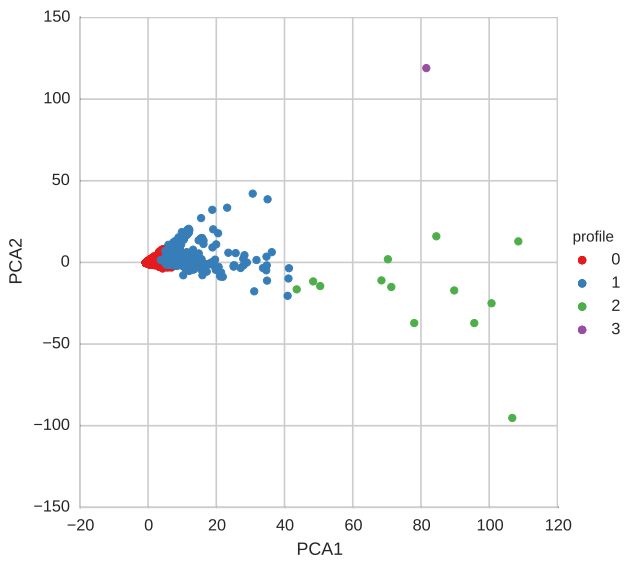


FIGURE 6 – Visualisation the clusters of users in a PCA space

4.4 Interactive groups, their cohesion and structure

About the user mentions. A common practice on social networks consists to mention people, using an "@" sign before their usernames in the text of the messages. At first this was only a social behaviour without any technical feature in the code of the social platform ; nowadays most platforms recognise such *mentions*.

As Galaxy2 is based on the old-fashion social framework *elgg*, such a feature is not hard-coded, but is still used by some members. Following the intuition that this interaction (the mention) would bring us valuable information, a mention extraction module has parsed each of the publications in search for "@" signs followed by usernames.

From the set of mentions M associated with the authors A, a graph of mentions, $G_M = (V, E)$ is built. V is the set of the union of the authors and the mentioned users ; edges $e \in E$ links the author of a message to the eventual mentioned user. This graph G_M is composed of 968 nodes, linked by 2342 edges [6]. The Louvain community detection [3] results on a set of 31 communities, amongst which only 11 contain more than 3 members.

Figure 7 illustrates one of these communities. Red nodes are part of the detected community ; blue nodes are the only external contacts, the boundary of the community in the graph G_M. A central user, *Bishop*, is the target of a few mentions, notably from an audacious account, *Nishikino_Maki* [7]. In a glance, we see that this community is centred on *Bishop*, who is not the most active in terms of community life (he does not mention the other members of the group).

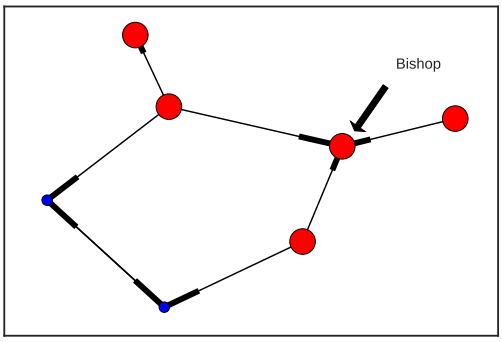


FIGURE 7 – Visualisation of a small community and its contacts through *mentions*

This community view helps the analyst to see the relations between a user and the rest of the network, and can guide his exploration along the interaction between accounts. However, as this graph of mentions is built from a social usage (using "@" to mention a user), the tool probably missed some differently written mentions and thus may be incomplete [8].

The friendship graph. The most known feature of social networks is the friendship connection. On Galaxy2, one can claim his friendship with another user, which is considered reciprocal on Galaxy2. This action is part of the list of the collected activities, and enables to link users along these friendship relations. Let G_F be the graph of friendship, where nodes are user accounts. Edges represent a claim of friendship.

From this graph, communities emerge, using the Louvain algorithm to reveal them. These communities may vary in number, based on the algorithm used. Here, 32 communities were obtained, amongst which 11 include more than 3 users. Figure 8 shows the distribution of the communities along two topological measures : x-axis gives the size of the group while y-axis represents the conductance, that is the proportion of boundary edges, linking the community to its environment, showing both its influence or its isolation.

On Figure 8, a group attracts the attention : with size=125 and conductance=0.74, this community is less tied to the rest of the network than it normally should. Its topical scores, based on [4], require some interpretation. $\xi = 0.064$ means that only 6,4% of its

6. there was a total of 5481 mentions : some accounts have been mentioned by a same user various times

7. He claimed to be "building a porn site cuz i can"

8. This graph G_M is *one of the* graphs of interaction, not the only one.

©AfIA 2018

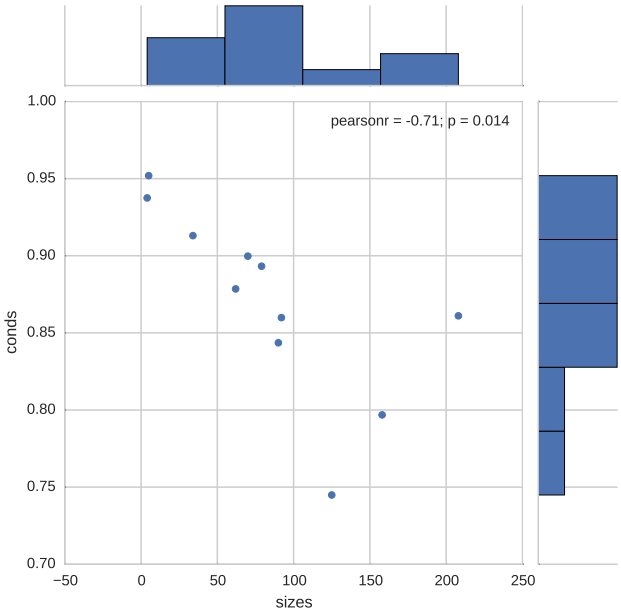


FIGURE 8 – Visualisation of the topological features of the communities

members, as a maximum, have been active about the same topic ; however, compared with the duration covered on the corpus, this situation is not surprising. In particular, this "main topic" clusters mainly around two sets of texts at different dates. The first refers to a small pornographic discussion group ; the second concerns the use of peer-to-peer mail through Tor.
These topical measures may benefit from improvements on the topic detection part, notably covering the temporality of topics and their labelling. This would lead to a more usable topic modelling, and a more precise group characterisation for the analyst.

5 Conclusion

The emergence of modern social networks reveals the inherent complexity surrounding us. Although most of the analysis is focused on the most known Online Social Networks (OSN), such as Twitter or Facebook, other instances exist ; Galaxy2 is much smaller than Twitter, but big enough to turn its analysis tedious due to the amount of exchanged messages.
We presented a whole system capable of analysing any type of OSN, showing its capabilities on a Tor-hosted social network. The functionalities include text analysis for topic and sentiment, user profiling and graph analysis, both on the friendship links as well as on the interaction between users. We showed that the tools, mostly developed for Twitter, can also be exploited for smaller networks. This approach has easily been adapted to the specific links and interactions of Galaxy2, as it could be for other platforms such as Reddit or Facebook.
Future work will be oriented on real-time processing, to analyse a social network as it lives. We have the intuition that a stream of messages would better fit the requirements in scaling and response times. We also desire to publicly release the Galaxy2 corpus, as one of the first complete social network archives.

Références

[1] G. Bertolin, *Digital Hydra : Security Implications of False Information Online.* NATO Strategic Communications Centre of Excellence, 2017.

[2] P. Noordhuis, M. Heijkoop, and A. Lazovik, "Mining twitter in the cloud : A case study," in *Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on.* IEEE, 2010, pp. 107–114.

[3] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of statistical mechanics : theory and experiment*, vol. 2008, no. 10, p. P10008, 2008.

[4] G. Gadek, A. Pauchet, N. Malandain, K. Khelif, L. Vercouter, and S. Brunessaux, "Measures for topical cohesion of user communities on twitter," in *Proceedings of the International Conference on Web Intelligence.* ACM, 2017, pp. 211–218.

[5] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up ? : sentiment classification using machine learning techniques," in *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10.* Association for Computational Linguistics, 2002, pp. 79–86.

[6] S. Baccianella, A. Esuli, and F. Sebastiani, "Sentiwordnet 3.0 : An enhanced lexical resource for sentiment analysis and opinion mining." in *LREC*, vol. 10, 2010, pp. 2200–2204.

[7] C. J. Hutto and E. Gilbert, "Vader : A parsimonious rule-based model for sentiment analysis of social media text," in *Eighth International AAAI Conference on Weblogs and Social Media*, 2014.

[8] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of machine Learning research*, vol. 3, no. Jan, pp. 993–1022, 2003.

[9] C. Sievert and K. E. Shirley, "Ldavis : A method for visualizing and interpreting topics," in *Proceedings of the workshop on interactive language learning, visualization, and interfaces*, 2014, pp. 63–70.

[10] H. Kwak, C. Lee, H. Park, and S. Moon, "What is twitter, a social network or a news media ?" in *Proceedings of the 19th international conference on World wide web.* ACM, 2010, pp. 591–600.

[11] A. Rao, N. Spasojevic, Z. Li, and T. DSouza, "Klout score : Measuring influence across multiple social networks," in *Big Data (Big Data), 2015 IEEE International Conference on.* IEEE, 2015, pp. 2282–2289.

[12] A. Java, X. Song, T. Finin, and B. Tseng, "Why we twitter : understanding microblogging usage and communities," in *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis.* ACM, 2007, pp. 56–65.

[13] T. Aynaud and J.-L. Guillaume, "Static community detection algorithms for evolving networks," in *Modeling and optimization in mobile, ad hoc and wireless networks (WiOpt), 2010 proceedings of the 8th international symposium on*. IEEE, 2010, pp. 513–519.

[14] C. Lee, H. Kwak, H. Park, and S. Moon, "Finding influentials based on the temporal order of information adoption in twitter," in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 1137–1138.

[15] K. Henderson, B. Gallagher, T. Eliassi-Rad, H. Tong, S. Basu, L. Akoglu, D. Koutra, C. Faloutsos, and L. Li, "Rolx : structural role extraction & mining in large graphs," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2012, pp. 1231–1239.

[16] S. Milgram, "Group pressure and action against a person." *The Journal of Abnormal and Social Psychology*, vol. 69, no. 2, p. 137, 1964.

[17] G. L. Cohen, "Party over policy : The dominating impact of group influence on political beliefs." *Journal of personality and social psychology*, vol. 85, no. 5, p. 808, 2003.

[18] S. Yardi and D. Boyd, "Dynamic debates : An analysis of group polarization over time on twitter," *Bulletin of Science, Technology & Society*, vol. 30, no. 5, pp. 316–327, 2010.

[19] K. H. Lim and A. Datta, "An interaction-based approach to detecting highly interactive twitter communities using tweeting links," in *Web Intelligence*, vol. 14, no. 1. IOS Press, 2016, pp. 1–15.

[20] Y. Tyshchuk, W. A. Wallace, H. Li, H. Ji, and S. E. Kase, "The nature of communications and emerging communities on twitter following the 2013 syria sarin gas attacks," in *Intelligence and Security Informatics Conference (JISIC), 2014 IEEE Joint*. IEEE, 2014, pp. 41–47.

[21] R. Cazabet and F. Amblard, "Simulate to detect : a multi-agent system for community detection," in *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2011 IEEE/WIC/ACM International Conference on*, vol. 2. IEEE, 2011, pp. 402–408.

[22] A. Clauset, M. E. Newman, and C. Moore, "Finding community structure in very large networks," *Physical review E*, vol. 70, no. 6, p. 066111, 2004.

[23] M. Rosvall and C. Bergstrom, "Maps of information flow reveal community structure in complex networks," Technical report, Tech. Rep., 2007.

[24] J. Yang and J. Leskovec, "Defining and evaluating network communities based on ground-truth," *Knowledge and Information Systems*, vol. 42, no. 1, pp. 181–213, 2015.

[25] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 8, pp. 888–905, 2000.

[26] M. E. Newman, "Modularity and community structure in networks," *Proceedings of the national academy of sciences*, vol. 103, no. 23, pp. 8577–8582, 2006.

[27] S. F. Everton, *Disrupting dark networks*. Cambridge University Press, 2012, no. 34.

[28] D. Ediger, K. Jiang, J. Riedy, D. A. Bader, and C. Corley, "Massive social network analysis : Mining twitter for social good," in *Parallel Processing (ICPP), 2010 39th International Conference on*. IEEE, 2010, pp. 583–593.

[29] J. Sankaranarayanan, H. Samet, B. E. Teitler, M. D. Lieberman, and J. Sperling, "Twitterstand : news in tweets," in *Proceedings of the 17th ACM SIGSPATIAL international conference on advances in geographic information systems*. ACM, 2009, pp. 42–51.

[30] F. Toriumi and S. Baba, "Real-time tweet classification in disaster situation," in *Proceedings of the 25th International Conference Companion on World Wide Web*. International World Wide Web Conferences Steering Committee, 2016, pp. 117–118.

©AfIA 2018

Étude de l'évolution du modèle de l'utilisateur des systèmes de construction collaborative d'ontologies

Alain Giboin

Inria, UCA, CNRS, I3S
Équipe Wimmics
2004 route des Lucioles
06992 Sophia Antipolis, France

alain.giboin@inria.fr

Résumé

Cet article rend compte d'une étude en cours sur l'évolution du modèle de l'utilisateur de systèmes de construction collaborative d'ontologies. Par modèle de l'utilisateur *(ou* modèle du contributeur*), nous entendons la représentation que les concepteurs se font des utilisateurs de leurs systèmes et plus généralement des contributeurs à la construction des ontologies. Nous décrivons : 1) la méthode que nous utilisons pour étudier l'évolution du modèle de l'utilisateur ; 2) l'évolution de ce modèle (en termes de types d'utilisateurs, de caractérisations de l'utilisateur et de caractérisations de l'environnement de l'utilisateur) ; 3) les évolutions parallèles : a) des méthodes de conception des systèmes collaboratifs ; b) des systèmes eux-mêmes ; et c) des méthodes de construction collaborative des ontologies. Nous mentionnons quelques perspectives d'évolution envisagées par les concepteurs eux-mêmes. Cette étude vise à faire ressortir l'importance d'acquérir une meilleure connaissance des contributeurs potentiels à la construction collaborative des ontologies afin d'obtenir des outils collaboratifs mieux adaptés à ces contributeurs*

Mots Clef

Systèmes de construction collaborative des ontologies ; Ingénierie ontologique participative ; Méthodes pour l'ingénierie ontologique ; Modèle de l'utilisateur ; Modèle du contributeur ; Evolution des modèles.

Abstract

This paper reports on an ongoing study on the evolution of the user model of collaborative ontology construction systems. By a user model (or a contributor model), we refer to the designers' representation of the users of their systems, and more generally of the contributors to the collaborative construction of ontologies. We describe: (1) the method we use to study the evolution of the user model; (2) the evolution of the model (in terms of user types, user characterizations; and user's environment characterizations); (3) the parallel evolutions of: (a) the design methods of collaborative systems, (b) the systems themselves; and (c) the methods of collaborative ontology construction. We mention some evolution perspectives envisioned by the designers themselves. This study aims to highlight the importance of acquiring a better knowledge of the potential contributors to the collaborative construction of ontologies in order to get collaborative tools better adapted to these contributors.

Keywords

Collaborative ontology construction systems; Participatory ontology engineering; Ontology engineering methods; User model; Contributor model; Model evolution..

1 Introduction

Dans le domaine de la conception et de l'évaluation de l'Interaction Humain-Machine (IHM), on désigne par *modèle de l'utilisateur* : ou bien *1)* le *modèle-de-l'utilisateur du concepteur* (la représentation que le concepteur se fait de l'utilisateur de son système) ; ou bien *2)* le *modèle-de-l'utilisateur du système* ou *modèle embarqué dans le système* (la représentation que « se fait » le système de son utilisateur au fur et à mesure de son interaction avec ce dernier) ; ou bien *3)* le *modèle-mental-du-système de l'utilisateur* (la représentation que l'utilisateur se fait du système) (cf. Kelly & Colgan, 1992)[47].

On s'intéresse ici au premier type de modèle et, plus précisément, au *modèle-de-l'utilisateur des concepteurs de systèmes de construction collaborative d'ontologies*, ou *modèle-du-contributeur* de ces concepteurs (les utilisateurs contribuant d'une façon ou une autre à la construction des ontologies). Cet article rend compte d'une étude en cours sur l'évolution de ce modèle entre le début des années 1990 et aujourd'hui, où l'on verra qu'initialement destinés à des groupes d'ontologues ou d'ingénieurs de la connaissance, les systèmes de construction collaborative d'ontologies se sont ouverts progressivement à d'autres types d'utilisateurs / contributeurs.

Par cette étude nous souhaitons montrer qu'une meilleure connaissance des utilisateurs/contributeurs – reflétée par un modèle de l'utilisateur/contributeur plus réaliste – est importante pour concevoir des systèmes adaptés. Cette étude se veut aussi une invitation à pratiquer plus avant l'analyse et la modélisation de l'utilisateur/contributeur de systèmes de construction collaborative d'ontologies.

Dans la suite de cet article, nous fournissons une première description : *1)* de la méthode que nous utilisons pour étudier l'évolution du modèle de l'utilisateur/contributeur ; *2)* de l'évolution de ce modèle

 ©AfIA 2018

(en termes de types d'utilisateurs, de caractérisations de l'utilisateur et de caractérisations de l'environnement de l'utilisateur) ; *3)* des évolutions parallèles : a) des méthodes de conception des systèmes de construction collaborative d'ontologies ; b) des systèmes eux-mêmes ; c) des méthodes de construction collaborative des ontologies et d) des perspectives d'évolution. Nous mentionnons pour terminer quelques perspectives d'évolution envisagées par les concepteurs eux-mêmes.

2 Méthode d'étude de l'évolution du modèle utilisateur

Dans cette section nous décrivons la méthode que nous utilisons pour étudier l'évolution du modèle de l'utilisateur des systèmes de construction collaborative d'ontologies.

2.1 Sources consultées

Nous consultons principalement les types suivants de sources : *1)* des travaux décrivant les systèmes de construction collaborative d'ontologies ;qu'ils s'appellent {éditeurs | outils | boîtes à outils | plates-formes | environnements | …} de {construction | conception | ingénierie | …} d'ontologies outils, boites à outils ou autre (voir Table 1) ;.*2)* des travaux décrivant des systèmes spécifiques et leurs utilisateurs (voir table 1, une liste non exhaustive de systèmes) ; *3)* des états de l'art sur les systèmes (ex. : Simperl & Luczak-Rösch, 2014 ; Correndo & Alani, 2007) [73][19] ou des articles comportant un état de l'art détaillé (ex. : Mangione et al., 2011)[57] ; *4)* des travaux décrivant des méthodes de construction collaborative d'ontologies supportées par un système ; *5)* des travaux rapportant des études d'utilisation des systèmes de construction collaborative d'ontologies.

Table 1 – Systèmes de construction collaborative d'ontologies recensés (liste non exhaustive).
[Voir Annexes.]

2.2 Analyse des sources

Dans ces différentes sources, nous cherchons à repérer plusieurs indicateurs de l'évolution du modèle de l'utilisateur des systèmes de construction collaborative d'ontologies, ainsi que des indicateurs des évolutions parallèles des systèmes, des méthodes de conception de ces systèmes, des méthodes de construction collaborative des ontologies et des perspectives d'évolution. Ces indicateurs sont principalement : 1) les termes désignant les types d'utilisateurs des COEs ; 2) les caractérisations de ces utilisateurs ; 3) les caractérisations de l'environnement de l'utilisateur/contributeur ; 4) la méthode d'analyse et de modélisation de l'utilisateur de COEs ; 5) le format de représentation du modèle de l'utilisateur ; les fonctions des COEs découlant des caractérisations des utilisateurs ; 6) la méthode de construction des ontologies en rapport avec les caractérisations des utilisateurs (recherche de correspondances entre éléments de la méthode et caractéristiques des utilisateurs) ; 7) les perspectives d'évolution – ce que les concepteurs envisagent de réaliser pour adapter davantage leurs systèmes aux utilisateurs ; 8) la méthode d'évaluation du COE utilisée : évaluation avec ou sans des utilisateurs (cf. inspection de l'interface) (quelle représentation a-t-on de l'utilisateur lors de cette phase du cycle de conception du système).

3 Évolution du modèle de l'utilisateur / contributeur

Dans cette section, nous décrivons plusieurs directions d'évolution du modèle de l'utilisateur/contributeur que nous avons repérées. Nous avons regroupé ces directions en trois catégories : *1)* évolution des types d'utilisateurs / contributeurs impliqués ; *2)* évolution des caractérisations de l'utilisateur/contributeur ; et *3)* évolution des caractérisations de l'environnement de l'utilisateur/contributeur.

3.1 Évolution des types d'utilisateurs / contributeurs impliqués

Une première évolution concerne les types d'utilisateurs/contributeurs impliqués dans la construction collaborative des ontologies.

On constate un élargissement des types d'utilisateurs/contributeurs : de l'ontologue ou ingénieur de la connaissance (celui qui structure et formalise la connaissance) à toute partie prenante (*stakeholder*), c'est-à-dire à *« quiconque est activement impliqué dans le développement d'une ontologie et dans son usage, et quiconque dont les intérêts peuvent être affectés par le développement de cette ontologie » (*di Maio, 2011)[27] (voir Table 2a). Notons cependant que, pour certains auteurs, les ontologues ou ingénieurs de la connaissance sont exclus des parties prenantes : les parties prenantes incluent les utilisateurs pas ou peu experts en ingénierie ontologique ; d'où la remarque que ces parties prenantes ont acquis désormais un « statut de première classe » dans le processus de construction des ontologies (cf. Shosha et al., 2015)[72].

On constate également un passage de l'individu ou du groupe restreint au collectif ouvert : de l'ontologue ou ingénieur de la connaissance ou de l'équipe d'ontologues à la communauté, voire à la foule (*crowd*) (voir Table 2b).

Table 2 – Types d'utilisateurs des systèmes de construction collaborative d'ontologies (liste non exhaustive) : a) Individus ; b) Collectifs
[Voir Annexes.]

3.2 Évolution des caractérisations de l'utilisateur / contributeur

Une autre catégorie d'évolutions concerne les caractérisations de l'utilisateur / contributeur. Ces caractérisations deviennent plus complètes et plus détaillées. Ces évolutions montrent le besoin de mieux identifier et comprendre les types possibles d'utilisateurs/contributeurs afin de mieux adapter les systèmes à ces derniers. L'évolution des caractérisations s'observe du point de vue du contenu et du format de ces dernières.

3.2.1 Contenu des caractérisations de l'utilisateur

Du point de vue du contenu des caractérisations, on constate ainsi :

- *une description de plus en plus fondée sur des observations de contributeurs réels* ; par analogie avec la distinction chère aux ergonomes entre tâches prescrites et tâches effectives – ces dernières étant fondées sur des observations –, on pourrait introduire la distinction entre *modèle prescrit* et *modèle effectif* ; d'ailleurs, Van Laere et al. (2014)[87] montrent que les types et rôles prescrits de parties prenantes ne suffisent pas à caractériser les contributeurs d'un projet ontologique et qu'il est nécessaire de les établir grâce par exemple une méthode de profilage à partir d'une analyse des interactions réelles entre ces parties prenantes et entre chaque partie prenante et le système ;
- *une description de plus en plus fine des rôles de chaque contributeur* (rôle explicite, rôle implicite ; rôle reconnu, rôle non (encore) reconnu ; etc.). Sont apparus par exemple les rôles de : *Super ontologist* | *Tacit ontologist* ; *Ontology author* ; *Ontology committer* (*a proxy for author*) | *Unique committer* ; *Ontology manager* ; *Ontology editor*; *Ontology submitter* ; *Ontology curator* | *Central curator* | *Single curator* | *Alpha curator* | *Content curator*; *Validator*; *Scribe* | *Single scribe*; *Modeler* | *Individual modeler*; *Facilitator* | *Moderator* ; *Panelist* ; *Authoritative party* ;
- *une description plus poussée des motivations, buts, engagements, connaissances, compétences, préférences, exigences, attitudes, émotions… des utilisateurs/contributeurs et de leur impact sur les systèmes et sur les méthodes*. Par exemple, Mezghani et al. (2016)[58] ont observé que les utilisateurs ayant une formation en génie logiciel préféraient une approche asynchrone de l'ingénierie ontologique collaborative alors que les utilisateurs ayant une formation en science informatique préféraient une approche synchrone ;
- *une description plus complète des tâches, modes opératoires des différentes catégories d'utilisateurs/contributeurs* ; des activités jusque-là considérées comme secondaires ou les activités auxquelles on n'avait pas pensé sont mises en valeur (ex. : les retours des utilisateurs des ontologies). En rapport avec cette prise en compte de ce type d'activités est apparue la notion d'« ontologues tacites » (*tacit ontologists*), c'est-à-dire ceux qui contribuent par des mails ou par des annotations (Malone & Stevens, 2013)[56]. On se dirige ainsi vers une reconnaissance de toutes les activités contributrices ;
- *une description plus poussée des modes de collaboration et d'interaction entre contributeurs* (ces derniers participant au même projet).

3.2.2 Format des caractérisations de l'utilisateur

Du point de vue du format des caractérisations, on observe entre autres : un *passage d'une représentation abstraite à une représentation plus concrète de l'utilisateur/contributeur* (ex. : représentation sous forme de Personas – voir section 4.1.1) ; une *description plus fréquente de scénarios impliquant les utilisateurs/contributeurs* (voir section 4.1.1) ; une *inclusion plus fréquente d'une représentation de l'utilisateur/contributeur dans les schémas d'architecture* des systèmes de construction collaborative d'ontologies (cf. de Sainte Marie et al., 2011).

3.3 Évolution des caractérisations de l'environnement de l'utilisateur / contributeur

Une troisième catégorie d'évolutions concerne l'environnement physique et social (ou cadre [*setting*] ou contexte) de l'utilisateur/contributeur, cet environnement exerçant une influence sur l'activité des utilisateurs/contributeurs. De manière générale, on observe une description de l'utilisateur/contributeur incluant de plus en plus de caractéristiques de cet environnement. On constate par exemple :

- une *description incluant de plus en plus des caractéristiques sociales* : mode de répartition du travail ou mode d'attribution d'une contribution (en relation avec les rôles des contributeurs ; cf. Li et al., 2005)[51], mode de communication (Toppano, 2010)[82], particularités culturelles (en particulier lorsque la collaboration est interculturelle ; cf. Anticoli & Toppano, 2011a et b)[3][4]… ;
- une *description de la contribution de l'utilisateur/contributeur en rapport avec la tâche principale ou courante de ce dernier* (la tâche principale ou courante motivant la contribution à l'ontologie ; voir par exemple l'association entre *tagging* (folksonomie) et construction de l'ontologie ; Huyinh-Kim-Bang et al., 2008 ; Limpens et al., 2008)[44][53].

4 Évolutions parallèles

Nous décrivons maintenant des évolutions qui ont eu lieu parallèlement à l'évolution du modèle de l'utilisateur / contributeur : *1)* l'évolution des méthodes de conception des systèmes (en particulier des méthodes d'analyse et de modélisation des utilisateurs) ; *2)* l'évolution des systèmes de construction collaborative d'ontologies ; et *3)* l'évolution des méthodes de construction collaborative d'ontologies. L'objectif de cette section est de montrer les conséquences ou les causes de l'évolution du modèle de l'utilisateur / contributeur.

4.1 Évolution des méthodes de conception des systèmes

Les méthodes de conception des systèmes, en particulier les méthodes d'analyse et de modélisation des utilisateurs, se sont davantage centrées sur l'utilisateur. Les concepteurs se sont ainsi beaucoup inspirés des méthodes utilisées dans les communautés IHM et CSCW (*Computer-Supported Cooperative Work*). L'objectif était de mieux comprendre les utilisateurs/contributeurs pour mieux adapter les

©AfIA 2018

systèmes à ces derniers et guider les concepteurs. On observe un développement des approches empiriques et théoriques.

4.1.1 Approches empiriques

On constate ainsi par exemple :

- un fort développement des ***études empiriques de l'activité de construction collaborative des ontologies*** dans des projets réels avec des systèmes existants : analyse et modélisation des *patterns (logs) d'usage* dans des projets de taille et d'envergure différentes (Walk et al., 2014)[89] ; étude des *patterns collaboratifs* dans des grands projets de développement d'ontologies (Falconer et al., 2011)[30] ; étude de la *dynamique de collaboration* lors de la construction d'ontologies avec des environnements différents – WebProtégé et MoKi (Rospocher et al., 2014)[70] ; étude des *processus d'ingénierie collaborative distribuée* et des capacités correspondantes de l'outil Collaborative Protégé (Schober et al., 2009)[71] ; étude de plusieurs projets de développement d'ontologies sous l'angle des *processus et* des *coûts* (Simperl & Tempich, 2006)[74] ; étude des problèmes rencontrés dans les différentes tâches de construction d'ontologies (Vigo et al., 2014)[88] : étude de la *construction distribuée d'ontologies comme pratique professionnelle* (Randall et al., 2011)[67] ; étude exploratoire de l'*ingénierie ontologique pour la documentation en architecture* logicielle (de Graaf et al., 2014)[22] ; étude de l'*impact sur la qualité de l'ontologie de l'implication d'experts du domaine dans l'annotation sémantique* de leurs articles, par comparaison avec l'annotation par des ingénieurs de la connaissance (Tatarintseva & Ermolayev, 2013)[79] ; étude des *activités de co-création d'une ontologie* par différentes parties prenantes (Bleumers et al., 2011)[9] ; étude du *rôle de la répartition géographique*, etc., sur la construction d'ontologies (Pinto et al., 2009)[66] ; étude de la *motivation des utilisateurs/contributeurs* – c'est ainsi que les experts du domaine (auteurs d'articles) ne sont pas motivés pour s'engager à raffiner l'ontologie car ils ne sont pas impliqués dans le processus d'annotation, lequel est pris en charge par l'ontologue (Tatarintseva & Ermolayev, 2013)[9] ; étude des *facteurs déterminant l'implication des différents contributeurs* dans la construction collaborative des ontologies (Randall et al., 2011)[67], des facteurs tels que le moment auquel participer ou le lien avec ses propres objectifs ou attentes (cf. l'attitude de « satisfaction suffisante » ou *satisficing attitude* : ce que font les contributeurs est « suffisamment bon » pour atteindre leurs objectifs) ;
- l'utilisation de la *technique des défis*; voir par exemple le *Collaborative Knowledge Construction (CKC) Challenge* (Noy et al., 2008)[61] organisé dans le but de faire tester différents outils de construction collaborative de connaissances à différents utilisateurs et d'apprendre ce que ces utilisateurs attendent de ce type d'outils ;
- l'utilisation de *techniques d'analyse et de définition de profils utilisateurs* en ingénierie ontologique collaborative (Van Laere et al., 2014)[87] ;
- l'utilisation de *techniques d'analyse et de modélisation de l'utilisateur* classiques en IHM, comme la *technique des Personas* d'Alan Cooper (1999) ; les personas sont des archétypes *concrets* d'utilisateurs élaborés à partir de données d'entretiens et/ou d'observations d'utilisateurs potentiels du dispositif à concevoir ou à évaluer ; cf. par exemple l'utilisation de cette technique par de Bonis et al. (2011 ; de Sainte Marie et al., 2011)[21][23] pour évaluer l'utilisabilité du composant de création d'ontologies de la plate-forme OntoRule par des chefs d'entreprise, des analystes d'affaire et des développeurs ;
- l'utilisation des *techniques d'analyse des parties prenantes* pour rendre compte de la collaboration entre multiples parties prenantes (Kozaki et al., 2011)[50]. NB : la technique des Personas peut être considéré comme une technique d'analyse de parties prenantes dans la mesure où elle fait la distinction entre personas primaires, personas secondaires, personas tertiaires et *ante-personas* ;
- l'utilisation de *techniques de définition de rôles* (cf. Li et al., 2005 ; Walk et al., 2014)[51][90] ;
- l'utilisation des *techniques de scénarios d'usage ou de tâches* (cf. Giboin et al., 2002 ; Palma et al., 2011)[39][64]. On notera que la technique des Personas suppose l'élaboration de scénarios mettant en scène les personas utilisant le système en cours de conception ;
- l'utilisation d'*indicateurs centrés utilisateurs* tels que le degré de participation ou le degré d'agrément/accord avec la représentation du domaine (*degree of community grounding*) (Siorpaes & Hepp, 2007)[75] ;
- l'élaboration d'*outils de visualisation des processus de construction collaborative de différents contributeurs*, tels que l'outil Pragmatix (Walk et al., 2013 ; Walk et al., 2014 ; Walk et al., 2016)[89][90][91].

4.1.2 Approches théoriques

A côté de ces approches empiriques, se développent également des approches plus théoriques. Ces approches font appel à des théories des sciences humaines et sociales pour modéliser les utilisateurs/contributeurs. Par exemple :

- *Communication entre concepteurs et utilisateurs des ontologies :* Toppano (2010)[82] utilise un modèle basé sur la communication pour décrire la conception et l'utilisation ou la réutilisation des ontologies ; dans ce modèle les ontologies sont considérées comme des objets sémiotiques et l'accent est mis sur la relation entre l'interprétation de ces objets par les concepteurs des ontologies et leur interprétation par les utilisateurs.
- *Influence des contributeurs :* Considérant l'ontologie comme un objet social, Aimé et Charlet (2012, 2016) [1][2]font appel à la psychologie sociale pour déterminer les influences des différents

contributeurs (ontologue, expert du domaine) ; ils font référence en particulier à la psychologie socio-sémiotique de Chabrol (1984)[14] et à la conceptualisation du Web socio-sémantique de Zacklad (2005)[95].

- *Traits de personnalité des contributeurs :* Mezghani et al. (2016)[58] se basent sur des travaux de psychologie montrant que les comportements de partage de connaissance entre individus sont influencés par les traits de personnalité de ces individus. Pour déterminer les personnalités des utilisateurs/contributeurs et éviter les problèmes de communication entre ces derniers, ils utilisent le modèle des *Big Five*, qui décrit les cinq traits centraux de la personnalité (Ouverture, Conscienciosité, Extraversion, Agréabilité, Névrosisme).

4.2 Évolution des systèmes de construction collaborative d'ontologies

Les systèmes (leurs fonctionnalités, leurs interfaces utilisateurs, etc.) ont évolué vers une plus grande adaptation aux différents types d'utilisateurs/contributeurs possibles, à leurs caractéristiques et à leurs activités, en particulier collaboratives. On observe entre autres :

- l'*élargissement du périmètre fonctionnel des systèmes*. Au départ on parlait d'*éditeurs* ; on parle davantage maintenant d'*environnements* ou de *plates-formes*. Cet élargissement peut se traduire par une connexion à d'autres outils que ceux contenus dans l'environnement proprement dit de construction collaborative d'ontologies ;
- le *développement de fonctionnalités pour la réalisation de tâches amont et aval de construction de l'ontologie*, tâches auxquelles peuvent contribuer les parties prenantes autres que les ontologues. Par exemple : *a)* préconisation d'une phase de pré-conceptualisation dans le processus d'ingénierie ontologique collaborative (Braun et al., 2007)[11] ; *b)* aide à la modélisation, par les experts du domaine, des différences de points de vue sur le sens des termes (Towntology ; Keita et al., 2006)[46] ; *c)* aide à l'extraction contextuelle des termes candidats (ECCO ; Giboin & Durville, 2007)[38] ;
- l'i*mplémentation de* plug-ins *collaboratifs pour des environnements non collaboratifs*, par exemple pour Protégé (d'où Collaborative Protégé ; Tudorache & Noy, 2007)[83], pour WebOnto (d'où Hozo ; Kozaki et al., 2007)[50] ou pour Terminae (*plug-in* en projet ; Ressad-Boudghaghen et al., 2013)[68] ;
- la *connexion à des outils de représentation moins formelle des connaissances* (langage semi-formel de graphes, langue naturelle..) ou le *développement de fonctionnalités de visualisation des ontologies* : cf. CMapTools (Cañas et al., 2004)[13], VocBench (Stallato et al., 2015)[76], Bio-Mixer (Fu et al.,)[32], Hozo (Kozaki et al., 2007)[50], WebOnto et Tadzebao (Domingue, 1998)[28], ces deux derniers outils permettant de mettre en œuvre une « ingénierie ontologique graphique », par exemple d'exprimer des arguments à l'aide de textes, d'images Gif et même de croquis dessinés à la main ;
- l'*utilisation de traducteurs pour passer d'un langage à un autre* : comme les traducteurs vers Prolog, CLIPS et LOOM dans le cas d'Ontolingua (Farquhar et al., 1997)[31] ou les traducteurs vers RDF(S) et OWL dans le cas de WebODE (Corcho et al., 2002)[18] ;
- le *développement de fonctionnalités de communication entre contributeurs* ; pour cela on se base beaucoup sur les wikis (cf.. Siorpaes & Hepp, 2007, MyOntology ; Mezghani et al., 2016 ; Semantic Media Wiki)[75][58] ;
- l'*intégration d'une fonctionnalité d'argumentation*, pour aider par exemple aux délibérations entre contributeurs, comme dans Compendium (Buckingham Shum et al. 2002)[12] ou dans WebOnto et Tadzebao (Domingue, 1998)[28] ;
- l'*intégration d'une fonctionnalité de recommandation*, pour recommander par exemple des concepts aux experts du domaine (Walk et al., 2013)[89].

4.3 Évolution des méthodes de construction collaborative d'ontologies

Les méthodes de construction collaborative d'ontologies ont également évolué vers une plus grande adaptation aux différents types d'utilisateurs/contributeurs possibles, à leurs caractéristiques et à leurs activités, en particulier collaboratives. Là encore, les « méthodologues » ont cherché à centrer leurs méthodes sur les utilisateurs/collaborateurs réels de ces méthodes. Deux de ces « méthodologues » ont d'ailleurs intitulé leur méthode « méthodologie d'ingénierie ontologie centrée sur l'humain » – *Human-Centered Ontology Engineering. Methodology*, HCOME (Kotis & Vouros, 2006)[49].

Plusieurs directions d'évolution « centrées utilisateurs / contributeurs » apparaissent, parmi lesquelles :

- *simplifier la démarche de construction des ontologies*, comme dans le *« Just enough » ontology engineering* de di Maio (2011)[27], qui permet aux utilisateurs ayant peu ou pas d'expertise en ingénierie ontologique de participer à la construction d'ontologies légères ; ou comme dans UPON Lite, la méthodologie d'ingénierie ontologique rapide (De Nicola & Missikof, 2016)[24]. Voir aussi Siorpaes & Hepp (2007)[75];
- *impliquer l'ensemble des utilisateurs/contributeurs dans toutes les étapes du cycle de construction des ontologies*, comme dans HCOME (Kotis & Vouros, 2006)[49] ou DILIGENT (Pinto et al., 2009)[66]. On notera que pour certains « méthodologues » comme Ongenae et al. (2013)[62], la méthode HCOME est une méthode « extrémiste » car elle privilégie le rôle des experts du domaine dans les différentes phases de conception des ontologies ; Ongenae et al. proposent en conséquence une méthode occupant une position médiane entre HCOME et l'autre type de méthode extrémiste qui,

©AfIA 2018

elle, privilégie le rôle de l'ingénieur de la connaissance (les experts du domaine n'étant impliqués que dans la phase de spécification de l'ontologie). Voir également l'approche dite d'« ingénierie ontologique participative » (Giboin et al., 2008 ; Ongenae et al., 2011, 2013)[38][63][62] ;

- *intégrer toutes les tâches contributives* comme dans la méthode de développement collaboratif inter-organisationnel d'ontologies de Palma et al. (2011)[64], laquelle applique une approche holistique ;
- *s'appuyer sur le* tagging *collaboratif ou l'élaboration de folksonomies*, activités réalisées « naturellement » voire « sauvagement » par les utilisateurs/contributeurs autres que les ontologues (cf. par exemple : Halpin et al., 2007 ; Zacharias & Braun, 2007 ; Huynh-Kim-Bang & Dané, 2008 ; Limpens et al., 2008 ; Gandon & Giboin, 2008)[41][94][44][53][33] ;
- *développer les tâches de* crowdsourcing *ontologique* (cf. Mortensen, Musen & Noy, 2013)[60] ;
- *s'appuyer sur une meilleure connaissance des caractéristiques cognitives et/ou affectives des utilisateurs / contributeurs*, comme dans la méthode de Gavrilova & Leschcheva (2014)[34], qui s'appuie sur les styles cognitifs des utilisateurs/contributeurs (Dépendance vs Indépendance à l'égard du Champ, Impulsivité vs Réflexivité, Catégorisation en Largeur vs en Profondeur) ou comme dans la méthodologie de Mezghani et al. (2016)[58] qui s'appuie sur les traits de personnalité (Ouverture, Conscienciosité, Extraversion, Agréabilité, Névrosisme) ;
- *s'appuyer sur une meilleure connaissance des disciplines des utilisateurs/contributeurs* (cf. Kotis & Vouros, 2006 ; Bourcier et al., 2006 ; Ressad-Bouidghaghen et al., 2013)[49][10][68] ;
- *exploiter de manière plus systématique la méthode des scénarios*, par exemple dans la phase d'évaluation des ontologies, cette méthode étant utilisée en complément de la technique des « scénarios motivants » de la méthode TOVE de Grüninger et ses collègues (cf. Giboin et al., 2002)[39]. Voir également Mezghani et al. (2016)[58] ;
- *favoriser l'interactivité* comme dans la méthode interactive LOVMI de validation structurelle et sémantique des ontologies de Richard et al. (2015)[69] ;
- *tracer la logique de conception des ontologies* de façon à améliorer l'intercompréhension des décisions de conception (Dellschaft et al., 2008)[23].

5 Perspectives d'évolution envisagées par les concepteurs eux-mêmes

Nous venons de voir comment a évolué le modèle de l'utilisateur/contributeur de systèmes de construction collaborative d'ontologies et comment ont évolué en parallèle les méthodes de conception de ces systèmes, les systèmes eux-mêmes ainsi que les méthodes de construction collaborative d'ontologies. Quelles sont maintenant les perspectives d'évolutions futures qu'envisagent les concepteurs en rapport avec ces différents axes d'évolution ? Certaines de ces perspectives ont été réalisées et présentées dans les sections précédentes. On décrit ici quelques perspectives non encore réalisées à ce jour, ou pas entièrement.

Perspectives d'évolution du modèle de l'utilisateur / contributeur.– *a) Evolution des types d'utilisateurs / contributeurs impliqués :* préciser mieux les types de collectifs ; *b) Evolution des caractérisations de l'utilisateur/contributeur :* Enrichir les descriptions de tâches/stratégies/… des différents contributeurs; de leurs rôles ; des intérêts ; de leurs exigences sur les interfaces (Tudorache et al., 2008, 2013)[84][85] ; mieux définir la notion de communauté : sa dénomination, sa composition, son hétérogénéité, les relations entre ses membres (Randall et al., 2011)[67] ; *c) Evolution des caractérisations de l'environnement de l'utilisateur / contributeur :* définir les cadres réels d'utilisation des systèmes collaboratifs (Tudorache et al., 2008, 2013)[84][85] ; décrire les mauvais usages et les actes de vandalisme afin de les prévenir (Mainz et al., 2008)[55].

Perspectives d'évolutions parallèles.– *a) Évolution des méthodes de conception des systèmes* : découvrir davantage les utilisateurs/contributeurs ; découvrir de nouveaux résultats surprenants sur ces utilisateurs/contributeurs ; analyser la dynamique du développement collaboratif d'ontologies ; développer la connaissance des rôles ; évaluer le système dans d'autres cadres réels que ceux dans lesquels le système a déjà été testé (Tudorache et al., 2008, 2013)[84][85] ; *b) Évolution des systèmes de construction collaborative des ontologies :* ajouter des mécanismes supplémentaires de collaboration ; accorder aux utilisateurs/contributeurs des privilèges différents à différents niveaux de granularité en fonction des rôles (Tudorache et al., 2008, 2013)[84][85] ; c) *Évolution des méthodes de construction collaborative d'ontologies :* supporter des workflows différents selon les différents types d'utilisateurs/contributeurs (Tudorache et al., 2008, 2013)[84][85].

6 Conclusion

Nous venons de rendre compte d'une étude en cours sur l'évolution du modèle de l'utilisateur de systèmes de construction collaborative d'ontologies et sur les évolutions parallèles des méthodes de conception des systèmes de construction collaborative d'ontologies, des systèmes eux-mêmes et des méthodes de construction collaborative des ontologies. Cette étude se poursuit actuellement. Elle mériterait d'ailleurs d'être poursuivie de manière collaborative avec les membres de la communauté IC.

Quoi qu'il en soit, notre intention en présentant cette étude en cours était de montrer qu'une meilleure connaissance des utilisateurs/contributeurs – reflétée par un modèle de l'utilisateur/contributeur plus réaliste – est

importante pour concevoir des systèmes adaptés. En 2013 Tudorache et al. [85] écrivaient : *« À mesure que nous en apprenons plus sur la manière dont les experts du domaine construisent des ontologies dans un environnement distribué, nous pouvons ajuster les outils pour améliorer la collaboration. »* Avec la présente étude, on pourrait écrire : À mesure que nous en apprenons plus sur la manière dont tous les types de contributeurs construisent des ontologies dans un environnement distribué, nous pouvons ajuster les outils pour améliorer la collaboration entre ces contributeurs.

Cette étude a fourni par ailleurs des indications sur : a) les méthodes à utiliser pour acquérir cette meilleure connaissance de l'utilisateur ; b) les ajustements des méthodes de conception des systèmes de construction collaborative d'ontologies, des systèmes eux-mêmes et des méthodes de construction collaborative des ontologies, auxquels conduit cette connaissance.

Bibliographie

[1] Aimé X, Charlet J. (2013). IC : Ingénierie des Connaissances ou Ingénierie du Conformisme ? In: Troncy R, editor. *Actes des 24es journées francophones d'Ingénierie des Connaissances (IC'2013)*. Lille, France ; 2013.

[2] Aimé X, Charlet J. (2016). Social Psychology Insights into Ontology Engineering. *Future Generation Computer Systems* 54, p. 348–351.

[3] Anticoli, L., Toppano, E. (2011a). The role of culture in collaborative ontology design. ISWSA 2011: 4.

[4] Anticoli, L., Toppano, E. (2011b). How Culture May Influence Ontology Co-Design: A Qualitative Study. *IJITWE 6(2)*, p. 1-17.

[5] Auer S., Dietzold S., & Riechert T. (2006) OntoWiki – A Tool for Social, Semantic Collaboration. In Cruz I. et al. (eds) *The Semantic Web - ISWC 2006. ISWC 2006*. Lecture Notes in Computer Science, vol 4273, p. 736-749, Springer, Berlin, Heidelberg.

[6] Bao, J. & Honavar V.(2004). Collaborative ontology building with wiki@nt - a multiagent based ontology building environment. In *Proc. of 3rd International Workshop on Evaluation of Ontology-based Tools*, located at ISWC 2004, 8th November 2004, Hiroshima, Japan, p. 37–46.

[7] Bao, J., Hu, Z., Caragea, D., Reecy, J., & Honavar, V. (2006). A tool for collaborative construction of large biological ontologies. In: Bressan, S., Küng, J., Wagner, R. (eds.) *DEXA 2006*. LNCS, vol. 4080, pp. 191–195. Springer, Heidelberg.

[8] Bao, J., Hu, Z., D., Reecy, J., Honavar, V.G. (2006). A Proposal for Collaborative Ontology Editor for Animal Trait Ontology.

[9] Bleumers, L., Jacobs, A., Ongenae, F.,Ackaert, A., Sulmon, N., Verstraete, M., van Gils, M., & De Zutter, S.. (2011). Towards Ontology Co-creation in Institutionalized Care Settings. *5th International Conference on Pervasive Computing Technologies for Healthcare and Workshops, PervasiveHealth 2011*, p. 559-562.

[10] Bourcier, D., Dulong de Rosnay, M., & Legrand, J. (2006). Susciter la construction interdisciplinaire d'ontologies juridiques : bilan d'une expérience. In M. Harzallah, J. Charlet et N. Aussenac-Gilles. *Semaine de la Connaissance, journée Ontologies et textes juridiques*, Juin 2006, Nantes, France. 3, p. 50-59.

[11] Braun, S., Schmidt, A., Walter, A., Nagypal, G., & Zacharias, V. 2007. Ontology maturing: a collaborative Web 2.0 approach to ontology engineering. In *Proceedings of the Workshop on Social and Collaborative Construction of Structured Knowledge* at the 16th International World Wide Web Conference (WWW 2007), Edinburgh.

[12] Buckingham Shum, S., Motta, E., & Domingue, J. (2002) Augmenting Design Deliberation with Compendium: The Case of Collaborative Ontology Design. *Workshop on Facilitating Hypertext-Augmented Collaborative Modelling* ACM Hypertext Conference, Maryland, June 11th-12th, 2002.

[13] Cañas, A. J., Hill, G., Carff, R., Suri, N., Lott, J., Eskridge, T., et al. (2004). CmapTools: A knowledge modeling and sharing environment. In A. J. Cañas, J. D. Novak & F. M. González (Eds.), *Concept maps: Theory, methodology, technology. Proceedings of the first international conference on concept mapping* (Vol. I, pp. 125-133). Pamplona, Spain: Universidad Pública de Navarra.

[14] Chabrol C. (1984). Psycho-socio-sémiotique : Définitions et propositions. *Langage et société 28(1)*, p. 53-71

[15] Chen, Y., Zhang, S., Peng, X., & Zhao, W. (2008). A Collaborative Ontology Construction Tool with Conflicts Detection. In *2008 Fourth International Conference on Semantics, Knowledge and Grid*, Beijing, 2008, p. 12-19.

[16] Cooper, A. (1999). *The Inmates Are Running the Asylum: Why High Tech Products Drive Us Crazy and How to Restore the Sanity*. Macmillan Publishing Co., Inc. Indianapolis, IN, USA.

[17] Corby, O. & Durville, P. (2009). ECCO (Editeur Collaboratif et Contextuel d'Ontologies). *Journée "Composition logicielle"*, Sophia Antipolis, 17 avril 2009.

[18] Corcho, O., Fernández-López, M., Gómez-Pérez, A., & Vicente, O. (2002): WebODE: An Integrated Workbench for Ontology Representation, Reasoning, and Exchange. *EKAW 2002*, p. 138-153.

[19] Correndo, G. & Alani, H. (2007). Survey of Tools for Collaborative Knowledge Construction and Sharing. *Web Intelligence/IAT Workshops 2007*, p. 7-10.

[20] Dabrowski, M., Kruk, S.R., Piotrowski, P., Szczecki, P. & Wozniak, M. (2007). Collaborative Ontology Development with MarcOnt Portal, *ESWC 2007*.

[21] De Bonis, S., Bellino, C., & El Ghali, A. (2011) *Final Usability Report: Evaluation and Conclusions*. OntoRule Project Deliverable D2.5, December 2011, 79 p.

[22] de Graaf, K.A. Liang, P., Tang, A., van Hage, W.R., & van Vliet, H. (2014). An exploratory study on ontology engineering for software architecture

©AfIA 2018

documentation. *Computers in Industry 65(7)*, p. 1053-1064 .

[23] Dellschaft, K., Engelbrecht, H., Barreto, J.M, Rutenbeck, S., & Staab, S.(2008). Cicero: Tracking Design Rationale in Collaborative Ontology Engineering. ESWC 2008. p. 782-786.

[24] De Nicola, A. & Missikoff, M.: (2016). A lightweight methodology for rapid ontology engineering. *Commun. ACM 59(3)*, p. 79-86.

[25] de Sainte Marie C., Iglesias Escudero M., & Rosina P. (2011) The ONTORULE Project : Where Ontology Meets Business Rules. In: Rudolph S., Gutierrez C. (Eds.) *Web Reasoning and Rule Systems* (p. 24-29), RR 2011. Lecture Notes in Computer Science, vol 6902. Springer, Berlin, Heidelberg.

[26] Diamantini, C., Genga, L., Potena, D., & Storti, E. (2014). Collaborative Building of an Ontology of Key Performance Indicators. *OTM Conferences 2014*, p. 148-165

[27] di Maio, P. (2011). 'Just enough' ontology engineering. In *Proceedings of the International Conference on Web Intelligence, Mining and Semantics (WIMS '11)*, Article No. 8.

[28] Domingue, J. (1998). Tadzebao and WebOnto: Discussing, Browsing, and Editing Ontologies on the Web. In *Eleventh Workshop on Knowledge Acquisition,Modeling and Management, EKAW 98* (18-23 April 1998, Banff, Alberta, Canada)

[29] Erdmann M., Waterfeld W. (2012) Overview of the NeOn Toolkit. In: Suárez-Figueroa M., Gómez-Pérez A., Motta E., Gangemi A. (eds) Ontology Engineering in a Networked World. Springer, Berlin, Heidelberg.

[30] Falconer, S.M., Tudorache, T., & Noy, N.F. (2011). An analysis of collaborative patterns in large-scale ontology development projects. *K-CAP 2011*, p. 25-32

[31] Farquhar, A., Fikes, R., & Rice, J. (1997). The ontolingua server: A tool for collaborative ontology construction. *International journal of human-computer studies 46 (6)*, p. 707-727.

[32] Fu, B., Grammel, L., & Storey, M-A.D. (2012). BioMixer: A Web-based Collaborative Ontology Visualization Tool. *ICBO 2012*.

[33] Gandon, F. & Giboin, A. (2008). Vers des ontologies à l'état sauvage. In *Atelier IC 2.0* (associé aux 19èmes Journées Francophones d'Ingénierie des Connaissances, IC2008), Nancy.

[34] Gavrilova, T. A., & Leshcheva, I. A. (2014). Collective Ontologies Design and Development. In Proceedings of 2014 *Eighth International Conference on Complex, Intelligent and Software Intensive Systems (CISIS– 2014)*, Birmingham, United Kingdom, IEEE Computer Society, Conference Publishing Services (CPS) (p. 564-569).

[35] Gennari, J.H. , Musen, M.A. , Fergerson, RW., Grosso, W.E., Crubézy, M., Eriksson, H., Noy, N.F., & Tu, S.W. (2003) The Evolution of Protégé: An environment for knowledge-based systems development. *International Journal of Human-Computer Studies 58 (1)*, p. 89-123.

[36] Ghidini, Ch., Rospocher, M., & Serafini, L. (2010). MoKi: a Wiki-based Conceptual Modeling Tool. EKAW (Posters and Demos) 2010

[37] Giboin, A., Durville, P. (2007). [ECCO:] Editeur collaboratif et contextuel d'ontologie. In INRIA Sophia Antipolis, EADS, LISI (Eds.), Outils et Services de gestion des ontologies, Section 2 (p. 9-59), Rapport de projet ANR-RNTL « e-WoK_Hub ».

[38] Giboin, A., Durville, P., & Gandon, F. (2008). Ingénierie ontologique participative : essai de mise en œuvre avec l'éditeur collaboratif ECCO. In *Atelier IC 2.0* (associé aux 19èmes Journées Francophones d'Ingénierie des Connaissances, IC2008), Nancy.

[39] Giboin, A., Gandon, F., Corby, O., Dieng, R. (2002). User Assessment of Ontology-based Tools: A Step Towards Systemizing the Scenario Approach, In *Proceedings of EON'2002: Evaluation of Ontology-based Tools, OntoWeb-SIG3 Workshop at the 13th International Conference on Knowledge Engineering and Knowledge Management EKAW 2002*, Sigüenza (Spain), September 30, 2002, p. 63-73.

[40] Giboin, A., Grataloup, S., Morel, O., & Durville, P. (2013). Building Ontologies for Analyzing Data Expressed in Natural Language, In M. Perrin, J.-F. Rainaud (Eds.), *Shared Earth Modeling, Knowledge Based Solutions for Building and Managing Subsurface Structural Models*, (p. 232-259), Technip Editions, Paris (France).

[41] Halpin, H., Robu V., Shepherd H. (2007). The Complex Dynamics of Collaborative Tagging. In *WWW 2007*, ACM Press, p. 211-220.

[42] Hayes, P., Saavedra, R. & Reichherzer, Th. (2003). A Collaborative Development Environment for Ontologies (CODE).

[43] Hayes, P., Eskridge, T.C., Mehrotra, M., Bobrovnikoff, D., Reichherzer, Th., & Saavedra, R. (2005). COE: Tools for collaborative ontology development and reuse.

[44] Huynh-Kim-Bang, B. & Dané, E. (2008). Social bookmarking et tags structurés. *19es Journées Francophones d'Ingénierie des Connaissances (IC 2008)*, Juin 2008, Nancy, France. p.111-122.

[45] Jiménez Ruiz, E., Cuenca Grau, B., Horrocks, I., & Berlanga, R. (2011). Supporting Concurrent Ontology Development: Framework, Algorithms and Tool. *Data & Knowledge Engineering 70(1)*, p. 146-164.

[46] Keita, A., Roussey, C., & Laurini, R. (2006). Un outil d'aide à la construction d'ontologies pré-consensuelles : le projet Towntology. xxivème Congrès INFORSID, May 2006, Hammamet, Tunisie. pp.911-926, 2006.

[47] Kelly, C. & Colgan, L. (1992). User Modeling and User Interface Design. In A. Monk, D. Diaper and D. Harrison (Eds.). *People and Computers VII* (p. 227-239). Cambridge : Cambridge University Press.

[48] Kotis K., Vouros G.A., & Alonso J.P. (2005) HCOME: A Tool-Supported Methodology for Engineering Living Ontologies. In Bussler C., Tannen V., Fundulaki I. (eds) *Semantic Web and*

Databases. SWDB 2004. Lecture Notes in Computer Science, vol 3372. Springer, Berlin, Heidelberg

[49] Kotis, K. & Vouros, G.A. (2006). Human-Centered Ontology Engineering: the HCOME Methodology. *International Journal of Knowledge and Information Systems (KAIS) 10(1)*, p. 109–131.

[50] Kozaki, K., Saito, O., & Mizoguchi, R. (2012). A Consensus-Building Support System based on Ontology Exploration. IESD 2012, International Workshop at EKAW 20, Galway, Ireland

[51] Li M., Wang D., Du X., & Wang S. (2005). Ontology Construction for Semantic Web: A Role-Based Collaborative Development Method. In: Zhang Y., Tanaka K., Yu J.X., Wang S., Li M. (Eds.). *Web Technologies Research and Development - APWeb 2005*. Lecture Notes in Computer Science, vol 3399. Springer, Berlin, Heidelberg.

[53] Limpens, F., Gandon, F., & Buffa, M. (2008). Rapprocher les ontologies et les folksonomies pour la gestion des connaissances partagées : un état de l'art.. *19es Journées Francophones d'Ingénierie des Connaissances (IC 2008)*, Juin 2008, Nancy, France. p.123-134.

[54] Loskyll, M.; Heckmann, D.; & Kobayashi, I. (2009). UbisEditor 3.0: Collaborative Ontology Development on the Web. In *Proceedings of Workshop on Web 3.0: Merging Semantic Web and Social Web 2009 (SW)²*, Turin, Italy, CEUR Workshop Proceedings.

[55] Mainz, I., Weller, K. Paulsen, I. Mainz, D., & Kohl, J. (2008).Ontoverse: Collaborative Ontology Engineering for the Life Sciences. *Information & Praxis 59(2)*, p. 91-99.

[56] Malone, J. & Stevens , R. (2013). Measuring the level of activity in community built bio-ontologies, *Journal of Biomedical Informatics 46(1)*, p. 5-14,

[57] Mangione, G.R., Mazzoni, E., Orciuoli, F., & Pierri, A. (2011). A Pedagogical Approach for Collaborative Ontologies Building. In T. Daradoumis et al. (Eds.). *Technology-Enhanced Systems and Tools for Collaborative Learning Scaffolding*, p. 135-166.

[58] Mezghani, E., Expósito, E., & Drira, K. (2016).. A Collaborative Methodology for Tacit Knowledge Management: Application

[59] Missikoff, M., Navigli, R., Velardi, P.,: The Usable Ontology: An Environment for Building and Assessing a Domain Ontology. In *Proceedings of the First International Semantic Web Conference, (ISWC 2002)*, p. 39-53.

[60] Mortensen JM, Musen MA, & Noy N.F. (2013). Crowdsourcing the Verification of Relationships in Biomedical Ontologies. *Proceedings of the AMIA Annual Symposium*.

[61] Noy N.F., Chugh A, & Alani H. (2008). The CKC Challenge: Exploring Tools for Collaborative Knowledge Construction. *IEEE Intell Syst. 23(1)*, p. 64-68.

[62] Ongenae, F. Duysburgh, P., Sulmon, N., Verstraete, M., Bleumers, L., De Zutter, S., Verstichel, S., Ackaert, A. Jacobs, A., & De Turck, F. (2013). An Ontology Co-design Method for the Co-creation of a Continuous Care Ontology, *Applied Ontology 2013*, p. 1–40

[63] Ongenae, F., Bleumers, L. Sulmon, N., Verstraete, M., Van Gils, M., Jacobs, A., De Zutter, S., Verhoeve, P., Ackaert, A., & De Turck, F. (2011). Participatory Design of a Continuous Care Ontology: Towards a User-Driven Ontology Engineering Methodology

[64] Palma, R., Corcho, O., Gómez-Pérez, A., & Haase, P. (2011). A holistic approach to collaborative ontology development based on change management. *J. Web Sem. 9(3)*, p. 299-314.

[65] Petersen, N., Coskun, G., & Lange, Ch. (2016). TurtleEditor: An Ontology-Aware Web-Editor for Collaborative Ontology Development. *ICSC 2016*: 183-186. [see also: Petersen, N., Similea, A., Lange, Ch., & Lohmann, S. (2017). TurtleEditor: A Web-Based RDF Editor to Support Distributed Ontology Development on Repository Hosting Platforms. Int. J. Semantic Computing 11(3): 311-324 (2017)].

[66] Pinto, H.S. Tempich, Ch. & Staab, S. (2009). Ontology engineering and evolution in a distributed world using DILIGENT. In *Handbook on ontologies* (p. 153-176), Springer, Berlin, Heidelberg

[67] Randall, D., Procter, R., Lin, Y., Poschen, M., Sharrock, W., & Stevens, R. (2011). Distributed ontology building as practical work. . *International Journal of Human-Computer Studies 69*, p. 220-233.

[68] Ressad-Bouidghaghen, O., Szulman, S., Zargayouna, H. & Paul, E. (2013). Construction collaborative d'une Ressource Termino-Ontologique (RTO) pour le droit des collectivités territoriales. *IC - 24èmes Journées francophones d'Ingénierie des Connaissances*, Jul 2013, Lille, France.

[69] Richard M., Aimé X., Krebs M.O., & Charlet J. (2015). LOVMI : vers une méthode interactive pour la validation d'ontologies. In: Actes des 26es journées francophones d'Ingénierie des Connaissances (IC'2015).

[70] Rospocher M., Tudorache T., Musen M.A. (2014) Investigating Collaboration Dynamics in Different Ontology Development Environments. In: Buchmann R., Kifor C.V., Yu J. (Eds.) *Knowledge Science, Engineering and Management. KSEM 2014*. Lecture Notes in Computer Science, vol 8793. Springer, Cham

[71] Schober, D., Malone, J., & Stevens, R. (2009). Practical experiences in concurrent, collaborative ontology building using Collaborative Protégé. *Nature Precedings* <http://hdl.handle.net/10101/npre.2009.3517.1> (2009)

[72] Shosha R., Debruyne C., & O'Sullivan D. (2015) Towards an Adaptive Tool and Method for Collaborative Ontology Mapping. In: Ciuciu I. et al. (Eds.) *On the Move to Meaningful Internet Systems: OTM 2015 Workshops*. Lecture Notes in Computer Science, vol 9416. Springer, Cham.

[73] Simperl, E. & Luczak-Rösch, M. (2014). Collaborative ontology engineering: a survey. *Knowledge Eng. Review 29(1)*, p. 101-131.

[74] Simperl, E. and Tempich, C., (2006) Ontology engineering: A reality check Meersman, R. and Tari,

©AfIA 2018

Z. (eds.) *The 5th International Conference on Ontologies, DataBases, and Applications of Semantics (ODBASE2006)*, France. 31 Oct - 02 Nov 2006. , p. 836-854.

[75] Siorpaes, K. & Hepp, M. (2007).myOntology: The Marriage of Ontology Engineering and Collective Intelligence, *Proceedings of the ESWC 2007 Workshop "Bridging the Gap between Semantic Web and Web 2.0"*, June 7, 2007, Innsbruck, Austria.

[76] Stellato, A., Rajbhandari, S., Turbati, A., Fiorelli, M., Caracciolo, C., Lorenzetti, T., Keizer, J., & Pazienza, M.T. (2015). VocBench: A Web Application for Collaborative Development of Multilingual Thesauri. *ESWC 2015*, p. 38-53

[77] Stellato, A., Turbati, A., Fiorelli, M., Lorenzetti, T., Costetchi, E., Laaboudi, Ch., Van Gemert, W., & Keizer, J. (2017). Towards VocBench 3: Pushing Collaborative Development of Thesauri and Ontologies Further Beyond. *NKOS@TPDL 2017*; p. 39-52

[78] Sure, Y., Erdmann, M., Angele, J., Staab, S., Studer, R., & Wenke, D. (2002). OntoEdit: Collaborative ontology development for the semantic web. In: Horrocks, I., Hendler, J. (Eds.) *ISWC 2002*. LNCS, vol. 2342, p. 221. Springer, Heidelberg.

[79] Tatarintseva, O. & Ermolayev, V. (2013). Refining an Ontology by Learning Stakeholder Votes from their Texts. In *ICTERI: International Conference on ICT in Education, Research, and Industrial Applications* (p; 64-78), CEUR-WS.org.

[80] Tennison, J., O'Hara, K., & Shadbolt, N. (2002). APECKS: using and evaluating a tool for ontology construction with internal and external KA support. *Int. J. Hum.-Comput. Stud. 56(4)*, p. 375-422.

[82] Toppano, E. (2010). A communication-based model of ontology design and (re)use. ISWSA '10 Proceedings of the 1st International Conference on Intelligent Semantic Web-Services and Applications, Article No. 6 , Amman, Jordan — June 14 - 16, 2010.

[83] Tudorache, T. & Noy, N. (2007). Collaborative Protégé. In *Workshop on Social and Collaborative Construction of Structured Knowledge (CKC 2007)* at WWW 2007, Banff, Canada, 2007.

[84] Tudorache, T., Noy, N. F., Tu, S., & Musen, M. A. (2008). Supporting Collaborative Ontology Development in Protégé. In A. Sheth et al. (Eds.): *ISWC 2008*, LNCS 5318, p. 17–32.

[85] Tudorache, T., Nyulas, C., Noy, N. F., & Musen, M. A. (2013). WebProtégé: A Collaborative Ontology Editor and Knowledge Acquisition Tool for the Web. *Semantic Web*, *4*(1), p. 89–99.

[86] Valo, A., Hyvönen, E., & Komulainen, V. (2005). A Tool for Collaborative Ontology Development for the Semantic Web. In *Proceedings of International Conference on Dublin Core and Metadata Applications (DC 2005)*, Nov, 2005.

[87] Van Laere S., Buyl R., & Nyssen M. (2014) A Method for Detecting Behavior-Based User Profiles in Collaborative Ontology Engineering. In: Meersman R. et al. (Eds.) *On the Move to Meaningful Internet Systems: OTM 2014 Conferences. OTM 2014*. Lecture Notes in Computer Science, vol 8841. Springer, Berlin, Heidelberg.

[88] Vigo, M; Bail, S; Jay, C; & Stevens, R. (2014). Overcoming the Pitfalls of Ontology Authoring: Strategies and Implications for Tool Design. *International Journal of Human-Computer Studies 72(12)*, p. 835-845.

[89] Walk, S., Pöschko, J., Strohmaier, M., Andrews, K., Tudorache, T., Noy, N.F., Nyulas, C., & Musen, M.A. (2013). PragmatiX: An Interactive Tool for Visualizing the Creation Process Behind Collaboratively Engineered Ontologies. *Int. J. Semantic Web Inf. Syst. 9(1)*, p. 45-78.

[90] Walk, S., Singer, Ph., Strohmaier, M., Tudorache, T., Musen, M.A., & Noy, N.F. (2014). Discovering Beaten Paths in Collaborative Ontology-Engineering Projects using Markov Chains, *Journal of Biomedical Informatics 51*, p. 254-271.

[91] Walk, S., Tudorache, T., Musen, M.A. (2016). Visualizing User Editing Behavior in Collaborative Ontology-engineering Projects. *VOILA@ISWC 2016*, p. 68-79.

[92] Wohlgenannt, G., Sabou, M., & Hanika, F. (2016). Crowd-based ontology engineering with the uComp Protégé plugin. *Semantic Web 7(4)*, p. 379-398.

[93] Xexéo, G., Vivacqua, A., de Souza, J.M., Braga, B., D'Almeida, Jr, J.N., Almentero, B.K., Castilho, R., & Miranda, B. (2005). COE: A collaborative ontology editor based on a peer-to-peer framework. *Advanced Engineering Informatics 19(2)*, p.113-121.

[94] Zacharias, X. & Braun, S. (2007).. Soboleo – social bookmarking and lightweight engineering of ontologies. In *Proc. WWW 2007 Workshop on Social and Collaborative Construction of Structured Knowledge*, Banff, Canada, May 2007.

[95] Zacklad, M. (2005). Introduction aux ontologies sémiotiques dans le Web Socio Sémantique. In M.-C. Jaulent (Ed.) Actes des 16èmes journées francophones d'Ingénierie des Connaissances, Grenoble: PUG.

Annexes

TABLE 1 – Systèmes de construction collaborative d'ontologies recensés (liste non exhaustive).

Nom	**Description**
APECKS	APECKS = Adaptive Presentation Environment for Collaboration Knowledge Structuring A tool for ontology construction with internal and external KA support *Destiné aux experts du domaine* (Tennison et al., 2002)[80]
CoB Editor ATO Editor	COB = Collaborative Ontology Building An ontology editing tool which exploits the notion of modular ontologies to support sharing, reuse, and collaborative editing of partial order (i.e., DAG-structured) ontologies (Bao et al., 2006)[7][8]
	ATO = Animal Trait Ontologies Cet édieur est une adaptation de l'éditeur CoB (Bao et al., 2006)[7][8]
COE	COE = Collaborative Ontology Editor (Xexéo et al., 2005)[93]
CODE-COE	COE = Collaborative Ontology Environment Based on CMapTools *Tool dedicated to domain experts* (Hayes et al., 2003, 2005)[42][43]]
Collaborative Protégé Web Protégé uComp Protégé Plugin	Collaborative Protégé : An extension of the existing Protégé system that supports collaborative ontology editing (Tudorache & Noy, 2007)[83] (Tudorache et al., 2008)[84]
	Web Protégé: a collaborative ontology editor and knowledge acquisition tool for the Web (Tudorache et al. 2013)[85]
	uComp Protégé Plugin: A plug-in aimed at crowd-based ontology engineering (Wohlgenannt et al., 2016)[92]
Compendium	A tool for augmenting design deliberation in Collaborative Ontology Design (BuckinghamShum et al., 2002)[12]
ContentCVS	A tool for supporting concurrent ontology development (Jiménez Ruiz et al., 2010°[45]
ECCO Folkon	Editeur collaboratif et contextuel d'ontologies (Giboin & Durville, 2007; Giboin et al., 2008, 2013 ; Corby & Durville, 2009)[37][38][40][17]
	Folkon : Editeur de folksonomies Adil El Ghali (non publié)
HCONE et SharedHCONE	Human-centered ontology management environments supporting the HCOME methodology (Kotis et al., 2005, 2006)[48][49]
Hozo	A tool for distributed and collaborative construction of ontologies (Kozaki et al., 2007)[50]
KPI Onto Editor	Collaborative building of an ontology of key performance indicators (KPI) (Diamantini et al., 2014)[26]
MarcOnt Portal WMap Portal	MarcOnt Portal : Outil collaboratif de développement et de gestion des ontologies (Dabrovski et al., 2007)[20]
	WMap Portal (variante de MarcOntPortal) : outil collaboratif de mise en correspondance d'ontologies (Dabrovski et al., 2007)[20]
MoKi	A collaborative MediaWiki-based tool for modeling ontological and procedural knowledge in an integrated manner (Ghidini et al., 2010)[36]
MyOntology	A wiki-based ontology engineering system marrying ontology engineering and collective intelligence (Siorpaes & Hepp, 2007)[75]
NeOn Toolkit	An open-source, multiplatform ontology engineering environment, which provides comprehensive support for the ontology engineering life cycle of networked ontologies (Erdmann & Waterfeld, 2012)[29]
ONKI	A Tool for Collaborative Ontology Development for the Semantic Web (Valo et al., 2005)[86]

©AfIA 2018

[OntoCommand]	A collaborative ontology construction tool with conflicts Detection (Chen et al., 2008)[15]
OntoEdit	An collaborative ontology editor that integrates numerous aspects of ontology engineering (Sure et al., 2002)[78]
Ontolingua	A tool for collaborative ontology construction (Farquhar et al., 1997)[31]
OntoVerse	An ontology wiki supporting all phases of collaborative ontology engineering. (Mainz et al., 2008)[55]
OntoWiki	A tool to develop ontologies collaboratively The tool converts human readable maps into a machine readable ontology *Tool dedicated to domain experts* (Auer et al., 2006)[5]
Plug-ins collaboratifs associés à Terminae	Outil de collaboration ajouté à *Terminae*, l'outil de construction de ressources termino-ontologiques (RTO)*(en projet)* (Ressad-Bouidghaghen et al., 2013)[68]
Semantic Media Wiki	A. semantic wiki engine that supports collaborative ontology development. Mezghani et al., 2016)[58]
Soboleo	A tool combining social bookmarking and lightweight engineering of ontologies (Zacharias & Braun, 2007)[94]
Tadzebao et WebOnto	Discussing, browsing and editing ontologies on the Web (Domingue, 1998)[28]
Towntology	Un outil d'aide à la construction d'ontologies pré-consensuelles (Keita et al., 2006)[46]
TurtleEditor	An ontology-aware Web editor for collaborative ontology development (Petersen et al., 2016)[65]
UbisEditor 3	A tool for collaborative ontology development on the Web (Loskyll et al., 2009)[54]
Usable Ontology	Environnement de construction et d'évaluation d'une ontologie (Missikof et al., 2002)[59]
VocBench VocBench 2 VocBench 3	A Web application for collaborative development of multilingual thesauri and ontologies complying with Semantic Web standards (SKOS, OWL) (Stellato et al., 2015, 2017)[76][77] (autre nom pour VocBench : ACSW - Agrovoc Concept Server Workbench)
WebODE	An integrated workbench for ontology representation, reasoning, and exchange (Corcho et al., 2002)[18]
Wiki@nt	Environnement multi-agents de construction collaborative d'ontologies (Bao & Honavar, 2004)[6]

TABLE 2 – Types d'utilisateurs des systèmes de construction collaborative d'ontologies (liste non exhaustive) : a) Individus ; b) Collectifs

a) INDIVIDUS	
Noms	**Variantes ou Instances**
Ontologist *Ontologue*	Ontology expert | | Ontology engineering expert | Ontology engineer | Ontology designer | Ontology builder | Ontology developer | Ontology author | …
Knowledge Engineer *Ingénieur de la connaissance*	Knowledge expert | … Ontology engineer | Linguist) | … Cogniticien
Domain Expert *Expert du domaine*	Expert | Instructional design expert | External expert | Subject matter expert |Alpha subject matter expert | Domain specialist | Domain connoisseur | …
Developer	Knowledge-based system developer
Researcher	Knowledge-engineering researcher
User *Utilisateur*	End User | Targeted end-user | Potential end-user | Ontology user | Application user | Web user | Non expert builder … Direct user |Indirect user Professional | Practitioner | Researcher |Social Scientist | …
Stakeholder *Partie prenante*	Interested Party | Ontology stakeholders | Domain knowledge stakeholder… Users |Sponsors | Investors | Technology providers | Industry associations | Standardization bodies | Other people and roles

b) COLLECTIFS	
Noms	**Variantes ou Instances**
Dyad *Dyade*	Dyads having different cognitive styles
Group *Groupe*	Single research group | Group of contributors | Group of people | Group of non experts | Group of users with diverse levels of ontology expertise and training… |Beginner user g roupe | Expert user group | … Stakeholder group |Stakeholder subgroup | Ontology- co-creation stakeholder group | … Group of editors |Editor group | Guest group |
Board *Comité*	Expert board | Board of ontology stakeholders | …
Team *Equipe*	Design team |Team of ontology engineers and domain experts | …
Project *Projet*	Project participants | Participants to an ontology engineering project | …
Organization *Organisation*	International organization | …
Community *Communauté*	Community members |Community of learners |Large user community | Community of domain experts | Intra-community | Co-evolving communities…
Consortium *Consortium*	Consortium of ontology and software developers | …
Network *Réseau*	Réseau interdisciplinaire |…
Population *Population*	Large population of non experts | …
Crowd *Foule* *[Grand Public]*	Crowd worker | …

©AfIA 2018

Bandits-Manchots Contextuels : Précision Globale Versus Individuelle

Nicolas Gutowski[1,2] Tassadit Amghar[2] Olivier Camp[1] Fabien Chhel[1]

[1] ESEO-TECH
10 boulevard Jean Jeanneteau, 49100 Angers, France

[2] LERIA, Université d'Angers (UBL)
2 boulevard Lavoisier, 49000 Angers, France

nicolas.gutowski@eseo.fr

Résumé

Dans la littérature, la plupart des travaux sur les bandits manchots sont évalués à l'aide d'une mesure de la précision globale. Concernant les bandit-manchots contextuels, les approches existantes ont pour objectif d'atteindre une personnalisation individuelle. Ainsi, leur précision globale devrait refléter la précision individuelle pour chacun des utilisateurs. Afin de mesurer le niveau de personnalisation atteint par ces approches, nous avons défini une nouvelle évaluation comparant les précisions individuelles des recommandations faites à chaque utilisateur avec la précision globale. Sur la base de cette comparaison, démontrant des disparités entre la précision individuelle et la moyenne de précision globale, nous proposons *Sliding Window LinUCB (SW-LinUCB)*. *SW-LinUCB* est une combinaison de *LinUCB* (CMAB) et d'un mécanisme de diversification pénalisant les bras sélectionnés trop fréquemment. Notre approche, inspirée d'applications réelles, comme les systèmes de recommandation, ne nécessite pas uniquement d'atteindre une bonne précision globale mais doit aussi tenir compte de la précision individuelle. Nous expérimentons et discutons nos résultats sur plusieurs jeux de données réelles.

Mots Clefs

Apprentissage par renforcement, LinUCB, Bandits-manchots contextuels, Système de recommandation

Abstract

Most works on Multi-Armed Bandits (MAB) focus the evaluations of their methods on a global accuracy performance metric. In the case of Contextual Multi-Armed Bandit (CMAB), the existing algorithms claim to eventually provide full personalization, which might suggest that their global accuracy metric should reflect each user's individual accuracy. In order to verify this, we consider a novel approach of CMAB assessment focused on the evaluation of individual accuracy and compare it to global accuracy. Based on the results of this comparison highlighting some users far from the average global accuracy, we propose *Sliding Window LinUCB* (*SW-LinUCB*), a combination of the original *LinUCB* (CMAB) and a diversification mechanism penalizing arms which are pulled too frequently. It is motivated by the requirements of different real-world applications such as clinical trials or recommender systems, which must converge to a good global accuracy and should equally distribute it among individuals. We experiment and discuss the benefits and losses of the proposed method on several real-world datasets.

Keywords

Recommendation System, Reinforcement learning, LinUCB, Contextual Multi-armed Bandits

1 Introduction

De nos jours, les bandits-manchots contextuels (Contextual Multi-Armed Bandit : CMAB) sont très largement considérés par de nombreuses applications se heurtant à des problèmes de décision séquentielle e.g., les systèmes de recommandation [1], ou encore les essais cliniques [2]. À chaque itération, les algorithmes d'apprentissage de CMAB ont pour objectif de choisir une action optimale (tirer le bras optimal) parmi un ensemble de possibilités, en tenant compte du contexte donné et des récompenses passées obtenues en regard de ces actions. Dans la majorité des cas, les récompenses retournées sont égales à 1 si l'algorithme réalise une bonne classification du contexte par rapport à l'action choisie, et à 0 sinon [1]. Ainsi, la plupart des études évaluent la performance de leurs méthodes, jusqu'à un horizon final, à travers une mesure de précision globale e.g., récompense moyenne, cumul des récompenses, ou nombre de regrets total [1, 3, 4]. Néanmoins, de telles métriques semblent inadéquates pour déterminer la précision à associer à chaque contexte [5]. Ainsi, si nous prenons l'exemple des systèmes de recommandation pour lesquels les utilisateurs peuvent être des visiteurs réguliers, ou des abonnés (exemple : applications mobiles), il semble

 ©AfIA 2018

essentiel de tenir compte de leurs retours individuels en regard des recommandations qui leur sont faites.
De telles considérations nous ont amené à définir une méthode permettant de mesurer la précision individuelle, et à évaluer différents algorithmes existants au regard de notre nouvelle métrique. Le problème se pose alors comme étant double objectif puisqu'il convient de maximiser la précision individuelle tout en maintenant une bonne précision globale.
Ainsi, nous introduisons une mesure complémentaire de la performance qui est fondée sur l'observation de la fonction de distribution cumulative (Cumulative Distribution Function : CDF) de la précision individuelle. De plus, nous proposons l'adaptation d'une méthode de CMAB que nous nommons *SW-LinUCB*. Cette méthode est bâtie à partir de l'algorithme classique *LinUCB* combiné avec une fenêtre glissante. De ce fait, *SW-LinUCB* a pour objectif d'améliorer la précision individuelle en incorporant un mécanisme de diversification, tout en conservant une précision globale satisfaisante.
En nous appuyant sur des jeux de données d'applications réelles, nous étudions les performances de différentes méthodes existantes (MAB et CMAB) à travers notre mesure de CDF sur l'ensemble des contextes observés. Nous montrons que pour les deux méthodes — pour les CMAB, spécifiquement selon le niveau de parcimonie du vecteur de contexte — un fossé se creuse entre les utilisateurs pour lesquels nous obtenons un haut niveau de précision et ceux défavorisés par un très faible niveau de précision, malgré une précision globale satisfaisante. Nos expériences montrent que *SW-LinUCB* réussit à combler ce fossé tout en maintenant une précision globale qui ne décroit pas plus de 10% par rapport à celle d'origine. En outre, on constate que notre algorithme parvient à atténuer les effets négatifs d'un contexte insuffisamment renseigné.
La contribution de notre article est double : 1) nous proposons une nouvelle mesure de l'évaluation basée sur la précision individuelle ; 2) nous présentons *SW-LinUCB* qui est une adaptation de *LinUCB* dont l'objectif est de maximiser la précision individuelle tout en conservant une précision globale satisfaisante.
Cet article est organisé comme suit : la section 2 présente un état de l'art sur les problèmes de MAB et les différentes métriques utilisées. La section 3 introduit les travaux connexes sur les problématiques de CMAB, l'algorithme *LinUCB* et l'usage de fenêtres glissantes. La section 4 dresse notre problématique et la méthode que nous mettons en place. Dans la section 5 nous exprimons et discutons les résultats de nos expérimentations. Enfin, nous concluons et présentons les perspectives de notre travail dans la section 6.

2 Contexte

Le problème du Bandit-Manchot (Multi-Armed Bandit : MAB) est un sujet qui a suscité de nombreuses recherches depuis sa première formalisation en 1952 [6]. De nombreuses formulations ont pu être proposées : stochastiques [7, 8, 4], ou encore Bayésiennes [9]. Plus précisément, le défi pour tout problème de MAB consiste à construire une stratégie visant à tirer le bras optimal sans connaissance préalable de la rentabilité de chacun des bras disponibles. La résolution de ce problème consiste à trouver un compromis entre l'exploration de l'ensemble des bras pour en déduire leurs rentabilités et l'exploitation de ce qui a été inféré pour favoriser la sélection des bras optimaux. Une version étendue de ce problème prend en compte le contexte. Il s'agit du problème de Bandit-Manchot Contextuel (Contextual Multi-Armed Bandit : CMAB) [10, 3]. Ainsi, dans une approche CMAB, le défi visant à déterminer le bras optimal reste le même que pour un problème de MAB mais doit tenir compte du contexte des utilisateurs.

Dans la littérature, le critère le plus fréquemment observé pour mesurer la performance d'un algorithme de bandit reste la précision globale — i.e. le nombre de fois qu'une récompense positive a été obtenue en tirant les différents bras [7, 8, 1, 3, 4]. Néanmoins, en fonction du domaine dans lequel les bandits sont appliqués, l'évaluation de leur performance peut nécessiter de s'ouvrir à d'autres critères. En effet, comme c'est tout particulièrement le cas pour les systèmes de recommandations, il a été observé dans certaines études que les mesures de précision ne sont pas suffisamment adaptées et pourraient être préjudiciables et nuire à la satisfaction des utilisateurs [5]. De ce fait, même si des algorithmes de CMAB tels que *LinUCB* [1] ou encore *Contextual Thompson Sampling* [3] permettent à terme une personnalisation complète auprès de chaque utilisateur, une autre étude soutient en revanche que ceux-ci nécessitent un si grand nombre d'itérations pour atteindre cette personnalisation qu'ils risquent de causer la frustration des utilisateurs avant d'y parvenir [11].

De telles constats ont conduit des recherches sur les CMAB et les systèmes de recommandations vers deux directions : 1) Tenter de réduire le nombre d'itérations nécessaires pour atteindre la personnalisation pour chaque utilisateur [11, 12], 2) Prendre en considération d'autres critères d'évaluation de la performance comme : la qualité [13], la diversité et la nouveauté [14], la couverture et la sérendipité [15], ou encore la satisfaction utilisateur [16].

La majorité des travaux tendent à montrer que la diversification serait l'un des points-clé pour améliorer la satisfaction utilisateur. Par exemple cela permettrait de mieux répondre aux besoins éphémères des utilisateurs [17], de les aider à découvrir de nouveaux éléments [13], ou d'éviter les recommandations redondantes [18].

À notre connaissance, aucune approche n'aborde spécifiquement le problème de la recherche d'un compromis entre précision individuelle et précision globale pour les CMAB à travers l'usage de techniques de diversification. Ceci constitue l'objectif principal de notre travail.

3 Travaux Antérieurs

Cette section présente les concepts clés sous-jacents à notre approche : le problème de CMAB, l'algorithme *LinUCB* qui le résout, et un principe de diversification s'appuyant sur une fenêtre glissante.

3.1 Bandits-Manchots Contextuels

Les approches contextuelles du problème de bandits-manchots (CMAB) [10] ont été très largement étudiées via des méthodes telles que *LinUCB* [1], Contextual Thompson Sampling *CTS* [3] ou encore Neural Bandit [19]. Ces méthodes résolvent le problème de CMAB en supposant une dépendance linéaire entre la récompense attendue d'une action et son contexte. Selon les travaux de Langford [10], le problème de CMAB peut être défini comme suit : Soit $A = \{a_1, ..., a_k\}$ un ensemble donné de k bras indépendants. Soit $X \subseteq \mathbb{R}^d$ l'ensemble de vecteurs de contexte de dimension d caractérisant un utilisateur et son environnement e.g., $x \in X$ est un vecteur binaire codant les caractéristiques telles que : l'âge, le sexe, le métier, les préférences, les spécialités, la localisation ou encore les caractéristiques des bras eux-mêmes. Soit l'horizon $T \in \mathbb{N}^*$, à chaque itération $t \in [1, T]$, le contexte x_t incluant l'utilisateur, est pris en considération afin de permettre la sélection du bras optimal compte tenu des récompenses obtenues lors des itérations précédentes. Pour chaque itération t, soit $r_t = (r_{t,a_1}, ..., r_{t,a_k})$ le vecteur de récompense où r_{t,a_i} correspond à la récompense obtenue après avoir sélectionné le bras a_i et $r_{t,a_i} \in \{0, 1\}$ dans notre cas où les récompenses sont tirées depuis des distributions de Bernoulli. Soit $\mathcal{D}_{x,r}$ la distribution conjointe entre les contextes x et les récompenses r, et soit $\theta_{t,a}$ le vecteur de coefficients inconnu (restant à déterminer) associé au bras a à l'itération t. Nous supposons que les récompenses attendues d'un bras a à l'itération t est une fonction linéaire du vecteur de contexte x_t de dimension d tel que $\mathbb{E}[r_{t,a}|x_t] = \widehat{\theta}_a^\top x_t$ où $\widehat{\theta}_a$ représente le vecteur de coefficients estimé associé au bras a. Ainsi, soit $\Pi : X \to A$ l'ensemble des politiques possibles où la politique optimale devant être déterminée est $\pi^* = \arg\max_{\pi \in \Pi} \mathbb{E}_{r,x}[r_{t,\pi(x)}]$. Alors, soit $\pi_t \in \Pi$ la politique empruntée par un algorithme de CMAB $\mathcal{A}$ à l'itération t. Par conséquent, dans le cadre d'un environnement stationnaire où $\mathcal{D}_{x,r}$ ne varie pas, le pseudo-regret instantané à l'itération t peut alors être défini tel que $\rho_t(\mathcal{A}) = \mathbb{E}_{r,x}[r_{t,\pi^*(x_t)} - r_{t,\pi(x_t)}]$ et le pseudo-regret cumulé tel que $\rho(\mathcal{A}) = \sum_{t=1}^{T} \rho_t(\mathcal{A})$.

Les algorithmes tels que *LinUCB* [1] ou *Contextual Thompson Sampling* (*CTS*) [3] ont été modélisés et largement étudiés afin de résoudre ce problème de CMAB. Aussi, à la section suivante nous rappelons l'un des plus populaires d'entres eux : *LinUCB*.

3.2 LinUCB

Nous avons d'abord décidé de bâtir et d'expérimenter notre approche à partir de *LinUCB* [1] qui reste l'un des algorithmes de CMAB les plus célèbres présentés dans la littérature.

LinUCB [1] est un algorithme contextuel à bornes supérieures de confiance qui renforce rapidement la sélection des bras optimaux en ajoutant un *bonus* (l'écart de la récompense) au gain total calculé. À chaque itération t, *LinUCB* sélectionne le bras $a \in A$ avec le gain calculé $p_{t,a}$ maximum parmi l'ensemble des bras disponibles. $p_{t,a}$ est construit à partir d'une combinaison linéaire du coefficient $\theta_{t,a}$ et du vecteur de caractéristiques x_t auxquels vient s'ajouter l'écart de récompense qui représente la valeur d'action optimiste du gain obtenu. Le vecteur de coefficient $\widehat{\theta_a}$ est construit à partir de la matrice D_a de dimension $n \times d$ (n recommandations en correspondance de d caractéristiques), et $b_a \in \mathbb{R}^d$ représente le vecteur de réponse correspondant, dont les poids pour chaque dimension sont fonction des récompenses obtenues. Plus précisément, $\widehat{\theta_a} = (D_a^\top D_a + I_d)^{-1} b_a$ où I_d représente la matrice identité de dimension $d \times d$. Par conséquent, à chaque itération t, *LinUCB* sélectionne le bras a_t tel que $a_t = \arg\max_{a \in A} p_{t,a}$ où $p_{t,a} = \widehat{\theta_a}^\top x_t + \alpha\sqrt{x_t^\top (D_a^\top D_a + I_d)^{-1} x_t}$. Ainsi, $\widehat{\theta_a}^\top x_t$ représente l'espérance de récompense et $\alpha\sqrt{x_t^\top (D_a^\top D_a + I_d)^{-1} x_t}$ l'écart de récompense où α est un paramètre pouvant être considéré comme un critère de robustesse face au bruit. De plus, selon [20], il y a une probabilité d'au moins $1 - \delta$ que $\mid \widehat{\theta}_a^\top x_t - \mathbb{E}[r_{t,a}|x_t] \mid \leq \alpha\sqrt{x_t^\top (D_a^\top D_a + I_d)^{-1} x_t}$ avec $\alpha = 1 + \sqrt{ln(2/\delta)/2}$. Si un ensemble de bras contient k bras, alors la borne supérieure du regret sera en $\tilde{\mathcal{O}}\left(\sqrt{kdT}\right)$.

Néanmoins, même si le regret total est ici bien identifié et que sa borne supérieure a été démontrée, il est encore nécessaire de surmonter les problématiques de faible précision individuelle possiblement induite par des environnements non stationnaires ou par des contextes trop pauvres en informations.

Dans la section suivante, nous rappelons plusieurs mécanismes de diversification reposant sur l'utilisation de fenêtres glissantes.

3.3 Mécanismes de diversification

Dans les systèmes de recommandation, la diversité est pertinente pour la satisfaction individuelle. La diversification peut aussi trouver son intérêt dans le cadre d'environnements non-stationnaires afin de permettre à l'algorithme de rester à jour et favoriser les observations les plus récentes. À l'aide d'une fenêtre glissante, des algorithmes tels que *SW-UCB* [21], ou encore *Windows Thompson Sampling with Restricted Context (Windows TSRC)* [12] permettent d'atténuer les effets résultant de la non-stationnarité. De plus, pour résoudre ces mêmes problèmes induits par la non-stationnarité et plus particulièrement dans le cadre d'une problématique de bandits de type *restless* [22], il existe une approche utilisant également une fenêtre glissante et dont l'objectif est de pénaliser les bras qui ont été

©AfIA 2018

tirés trop souvent [23]. Cette approche intéressante a inspiré notre proposition.

4 Problématique et Méthodes

Dans cette section, nous posons notre problème, puis nous définissons notre nouvelle approche *SW-LinUCB*. Notre méthode est basée sur la combinaison de l'algorithme original *LinUCB* [1] et l'utilisation d'une fenêtre glissante inspirée de [23].

4.1 Énoncé du problème

Soit $\mathcal{U} = \{u_1, ..., u_n\}$ l'ensemble des n agents disponibles dans un problème de bandits, et pouvant par exemple correspondre dans le cadre d'applications réelles, à des utilisateurs ou encore des patients. Inspiré par [24], nous supposons pour chaque bras $a \in A$ et étant donné $x \in X \subseteq \mathbb{R}^d$, que $\mathcal{U}$ peut être partitionné en un nombre $m_a(x)$ de clusters $\mathcal{U}_{1,a}(x), \mathcal{U}_{2,a}(x), ..., \mathcal{U}_{m_a(x),a}(x)$ d'utilisateurs partageant les mêmes comportements vis à vis des récompenses qu'ils octroient à chaque bras a. Faisant maintenant l'hypothèse de l'existence d'un vecteur de contexte optimal $x^* \in X^*$ qui posséderait toutes les caractéristiques pertinentes associées, avec une confiance de 100%, au bras optimal correspondant et cela pour chaque contexte disponible. Comme *LinUCB* suppose une dépendance linéaire entre la récompense attendue d'une action et son contexte tel que $\mathbb{E}[r_{t,a}|x_t] = \widehat{\theta_a}^\top x_t$, alors lorsque x^* est fourni, *LinUCB* converge vers une précision de 100% et offre une personnalisation pour chaque individu. Cela signifie que toute précision individuelle convergera également vers 100%. Cependant, dans les situations réelles, x peut manquer d'informations et rester incomplet pour différentes raisons telles que : un manque d'information sur les caractéristiques des bras, une mauvaise modélisation du contexte c'est-à-dire un contexte spécifié de manière incomplète, des restrictions dues à des problématiques de confidentialité et de protection de la vie privée, un profil mal renseigné, des informations manquantes sur l'environnement de l'utilisateur (par exemple, une localisation temporairement indisponible). Dans les cas où $x \neq x^*$, les algorithmes de CMAB doivent faire face à des contraintes de parcimonies dans les données ou d'incomplétude sur les caractéristiques disponibles puisque les caractéristiques de x^* manquantes dans x ne peuvent pas être prises en compte. En effet, avec un vecteur de contexte insuffisamment décrit, les clusters associés à x^* ne seront pas pris en compte par *LinUCB*, qui peut finalement être incapable de tirer le bras optimal pour différentes situations. Les utilisateurs affectés par cette parcimonie vectorielle pourraient donc se retrouver insatisfaits de la sélection des bras qui leur est proposée par *LinUCB*. Cela entraîne une diminution de la précision globale mais également de la précision individuelle ciblant ces utilisateurs. Ces problématiques nous ont conduits à construire une nouvelle approche visant, à la fois, à garder une bonne précision globale et à atténuer la diminution de la précision individuelle. La sous-section suivante présente notre méthode qui utilise un mécanisme de diversification afin de contrer le manque d'information contextuelle et favoriser la sérendipité.

4.2 Sliding Window LinUCB : SW-LinUCB

Notre Fenêtre Glissante : Notre nouvelle approche combine *LinUCB* et l'utilisation d'une fenêtre glissante permettant de pénaliser la sélection des bras optimaux (tirés plus fréquemment), afin de favoriser l'exploration des bras moins optimaux que nous appellerons ici l'*ensemble des bras sous-optimaux*. Les méthodes utilisant des fenêtres glissantes appliquent généralement un coefficient dit de *discount* pondérant les récompenses obtenues par leurs bras afin de favoriser les observations les plus récentes. Ainsi, il est possible de définir un coefficient de *discount* qui pondère les récompenses cumulées obtenues pour chaque bras tel que $\sum_{t=1}^{T} \gamma_t r_{t,a}$ [23]. Avec $\gamma_t = 1 - \frac{Occ_w(a,t)}{w}$ où w correspond à la taille de la fenêtre glissante et $Occ_w(a,t)$ représente le nombre de fois qu'un bras a a été sélectionné durant les t dernières itérations. $Occ_w(a,t) = \#_1(E_{t,a})$ où $E_{t,a} = \{0..(2^{(w+1)} - 1)\}$ représente les w dernières sélections d'un bras a donné e.g., pour une taille de fenêtre $w = 6$, $E_{t,a} = 101001$ signifie que a a été sélectionné aux itérations $t-6$, $t-4$ et $t-1$. Néanmoins, même si il pourrait être intéressant de combiner une telle méthode à *LinUCB*, celle-ci reste un processus mettant en œuvre une mémoire à court-terme qui, dans notre cas, ne permettra pas de diversifier suffisamment. Dans notre cas, nous devons conserver un processus d'élimination des mauvaises solutions sur le long terme tout en diversifiant suffisamment parmi l'ensemble des bras sous-optimaux. Ainsi, nous proposons une nouvelle méthode de calcul du gain basée sur le $p_{t,a}$ originel tel que

$$p_{t,a}^w = \gamma_t \widehat{\theta_a}^\top x_t + \alpha \sqrt{x_t^\top M_a^{-1} x_t} \tag{1}$$

où $M_a = (D_a^\top D_a + I_d)$. Ce calcul permet à la fois de garder la confiance (élimination à long terme) grâce au *bonus*, et de diversifier suffisamment parmi l'ensemble des bras sous-optimaux en pénalisant temporairement l'espérance calculée des récompenses pour les bras sélectionnés trop fréquemment.

L'algorithme SW-LinUCB : l'objectif de *SW-LinUCB* est de déterminer la politique π qui maximise les récompenses cumulées à l'horizon T tandis qu'une fenêtre glissante force la diversification parmi l'ensemble des bras sous-optimaux. Notre hypothèse est la suivante : en fonction du niveau de parcimonie du vecteur de contexte ($x \neq x^*$), diversifier parmi l'ensemble des bras sous-optimaux atténuera la perte de précision individuelle pour les utilisateurs pour lesquels la méthode d'origine obtient une très faible précision. Notre méthode est décrite dans l'algorithme 1.

5 Expérimentations et Résultats

Jeux de données : L'évaluation de notre proposition se base sur quatre jeux de données. Tout d'abord un jeu de

Algorithme 1 *Sliding Window LinUCB (SW-LinUCB)*

Require: L'ensemble des k bras $a \in A$ disponibles, $\alpha \in \mathbb{R}^+$, l'horizon T, et l'ensemble des n contextes fixes disponibles X

1: $w \leftarrow k$
2: **for** $t = 1$ to T **do**
3: Considérer $x_t \in X$: un utilisateur et son contexte
4: **for all** $a \in A$ **do**
5: **if** a n'a pas encore été sélectionné **then**
6: $Occ_w(a,t) \leftarrow 0$; $M_a \leftarrow I_d$; $b_a \leftarrow 0_{d\times 1}$
7: **end if**
8: $\widehat{\theta}_a \leftarrow M_a^{-1} b_a$
9: **if** $t > w$ **then**
10: Calculer $Occ_w(a,t) = \#_1(E_{t,a})$
11: **end if**
12: $p^w_{t,a} \leftarrow \left(1 - \frac{Occ_w(a,t)}{w}\right) \widehat{\theta}_a^\top x_t + \alpha\sqrt{x_t^\top M_a^{-1} x_t}$
13: **end for**
14: Sélectionner le bras $a_t = \arg\max_{a_t \in A}(p_{t,a})$ et observer la récompense r_t retournée par l'utilisateur
15: $M_{a_t} \leftarrow M_{a_t} + x_t x_t^\top$; $b_{a_t} \leftarrow b_{a_t} + r_t x_t$
16: $\forall a \neq a_t$, mettre à jour toutes les sous-séquences $E_{t,a}$ en ajoutant un bit 0
17: Mettre à jour la sous-séquence E_{t,a_t} en ajoutant un bit 1
18: **if** $t > w$ **then**
19: Réaliser un décalage logique vers la gauche (Left Shift) de $E_{t,a}$ et E_{t,a_t}
20: **end if**
21: **end for**

données a été artificiellement généré afin d'obtenir un x^* garantissant une équiprobabilité entre chacun des bras. Il servira de jeu de contrôle dans nos expérimentations. Enfin, nous avons utilisé trois autres jeux de données d'applications réelles : Recommendation System for Angers Smart City (RS-ASM)[1], Covertype et Poker Hand[2]. Chacun des jeux de données considéré est constitué d'un nombre d'instances, s'appuie sur un contexte d'une dimension donnée et propose un nombre défini de bras (voir Tableau 1).

Jeu	Instances	Dim	Bras	Source
Contrôle	1000	4	4	Generated
RS-ASM	2152	56	18	Kaggle
Covertype	581 012	95	7	UCI
Poker Hand	1 025 010	11	9	UCI

Tableau 1 – Jeux de données

Mesure de Précision Globale : La précision globale est un critère de performance basé sur le total des récompenses positives cumulées à l'horizon T. De ce fait, pour obtenir la précision globale, nous calculons le gain c'est à dire le nombre total de récompenses positives $g(T)$ puis nous calculons enfin la précision (Accuracy : Acc, voir Tableau 2) tel que : $Acc(T) = \frac{g(T)}{T}$ où $g(T) = \sum_{t=1}^{T} r_t$ et $r_t = \{0, 1\}$.

Mesure de Diversité : La diversité de sélection parmi un ensemble fini et fixe de k bras peut être définie comme un critère de dispersion découlant du coefficient de variation (C_v) des bras sélectionnés. Il est par conséquent possible de calculer la diversité (Div, voir Tableau 2) comme suit : $Div(N) = 1 - \frac{c_v(N)}{\sqrt{k}}$ où $N = \{n_{a_1}, ..., n_{a_k}\}$, et n_{a_i} correspond au nombre de fois qu'un bras $a_i \in A$ a été sélectionné. Ainsi, la dispersion de la sélection tend à son maximum quand $c_v(N) \to 0$ alors qu'elle tend vers son minimum quand $c_v(N) \to \sqrt{k}$ [25].

Mesure de Précision Individuelle : La précision individuelle par utilisateur $Acc_u(T)$ peut être définie comme étant $\forall u \in \mathcal{U}, Acc_u(T) = \frac{\sum_{t=1}^{T} r_{t,u}}{T_u}$ où T_u représente le nombre de fois qu'un utilisateur avec son contexte a été sélectionné à l'horizon T, et $r_{t,u}$ correspond à la récompense retournée par u à l'itération t. Les mesures de précision individuelle de chaque utilisateur peuvent être représentées par une fonction de distribution cumulative (CDF). La CDF nous permet ainsi d'observer la distribution de la précision individuelle sur l'ensemble des utilisateurs $\mathcal{U}$ avec leur contexte à l'horizon T (voir Figure 1).

Comparaison des Algorithmes[3] : Nous comparons notre algorithme *SW-LinUCB* avec les méthodes suivantes : *UCB* standard (MAB) [8], et *LinUCB* classique (CMAB) [1]. Notons que *LinUCB* et *SW-LinUCB* auront la même valeur du paramètre α calculé avec $\delta = 0.1$.

Protocole Expérimental : Pour chaque algorithme et pour chaque jeu de données, nous simulons 2 cas différents : 1) Avec le vecteur de contexte complet (vc), 2) Avec une partie tronquée à 25% du vecteur de contexte d'origine (vt). Ici, le terme tronqué représente la proportion (en pourcentage) des caractéristiques, sélectionnées aléatoirement, que nous décidons de perdre au début de l'expérience. Ainsi, pour chacun des différents cas et pour chaque algorithme, nous simulons 20 expériences de $10,000,000$ d'itérations pour Poker Hand et Covertype, et $100,000$ concernant RS-ASM et le jeu de données de contrôle. Comme le nombre d'instances de chaque jeu de données est plus ou moins important, nous devons mettre à l'échelle l'horizon T pour chacun d'entre aux afin d'obtenir une mesure suffisante de la précision individuelle.

De plus, pour simuler un flux de données d'utilisateurs se présentant pour recevoir une recommandation (voir ligne 3 de l'algorithme 1), nous sélectionnons séquentiellement et aléatoirement les contextes disponibles dans l'ensemble du jeu de données. Ensuite, nous déterminons les moyennes et écart-types de précision globale et de diversité de sélection des bras sur l'ensemble des 20 simulations. De plus, nous calculons la précision individuelle et déduisons sa CDF dont les données et la représentation sont représentées Figure 1 et Tableau 2. Enfin, nous réalisons un test

[1] https://www.kaggle.com/
[2] http://archive.ics.uci.edu/ml/
[3] Voir notre étude préliminaire https://git.io/vxCcv

©AfIA 2018

		Mesures globales		Distribution de la Précision Individuelle				
		Acc	Div	10%	Q_1	Med	Q_3	90%
UCB	Contrôle	$0.25_{\pm\varepsilon}$	$10^{-3}{}_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.25_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$
	RS-ASM	$0.52_{\pm 0.08}$	$10^{-3}{}_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.80_{\pm 0.40}$	$1.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$
	Poker Hand	$0.47_{\pm 0.04}$	$10^{-3}{}_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.60_{\pm 0.49}$	$1.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$
	Covertype	$0.41_{\pm 0.05}$	$10^{-3}{}_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.90_{\pm 0.30}$	$0.90_{\pm 0.30}$
LinUCB (*vc*)	Contrôle	$1.0_{\pm\varepsilon}$	$0.997_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$
	RS-ASM	$0.78_{\pm\varepsilon}$	$0.86_{\pm\varepsilon}$	$0.04_{\pm 0.02}$	$0.77_{\pm 0.02}$	$0.95_{\pm\varepsilon}$	$0.99_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$
	Poker Hand	$0.53_{\pm\varepsilon}$	$0.06_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.90_{\pm 0.01}$	$1.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$
	Covertype	$0.72_{\pm\varepsilon}$	$0.44_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$
SW-LinUCB (*vc*)	Contrôle	$0.991_{\pm\varepsilon}$	$0.997_{\pm\varepsilon}$	$0.98_{\pm\varepsilon}$	$0.99_{\pm\varepsilon}$	$0.99_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$
	RS-ASM	$0.76_{\pm\varepsilon}$	$0.88_{\pm\varepsilon}$	$0.06_{\pm 0.02}$	$0.68_{\pm 0.01}$	$0.92_{\pm 0.01}$	$0.98_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$
	Poker Hand	$0.48_{\pm\varepsilon}$	$0.34_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.22_{\pm 0.02}$	$0.50_{\pm\varepsilon}$	$0.72_{\pm 0.02}$	$0.87_{\pm 0.02}$
	Covertype	$0.69_{\pm\varepsilon}$	$0.47_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.40_{\pm\varepsilon}$	$0.89_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$
LinUCB (*vt*)	Contrôle	$0.749_{\pm\varepsilon}$	$0.88_{\pm 0.07}$	$0.35_{\pm 0.09}$	$0.52_{\pm 0.04}$	$0.88_{\pm 0.04}$	$1.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$
	RS-ASM	$0.629_{\pm\varepsilon}$	$0.33_{\pm 0.01}$	$0.01_{\pm\varepsilon}$	$0.06_{\pm 0.01}$	$0.92_{\pm 0.01}$	$0.97_{\pm\varepsilon}$	$0.99_{\pm\varepsilon}$
	Poker Hand	$0.50_{\pm\varepsilon}$	$0.01_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.84_{\pm 0.04}$	$1.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$
	Covertype	$0.60_{\pm\varepsilon}$	$0.35_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$
SW-LinUCB (*vt*)	Contrôle	$0.746_{\pm\varepsilon}$	$0.96_{\pm 0.02}$	$0.43_{\pm 0.02}$	$0.50_{\pm\varepsilon}$	$0.82_{\pm 0.01}$	$0.99_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$
	RS-ASM	$0.567_{\pm\varepsilon}$	$0.69_{\pm\varepsilon}$	$0.08_{\pm 0.01}$	$0.33_{\pm 0.01}$	$0.61_{\pm 0.01}$	$0.85_{\pm 0.01}$	$0.95_{\pm 0.01}$
	Poker Hand	$0.48_{\pm\varepsilon}$	$0.33_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.26_{\pm 0.01}$	$0.50_{\pm\varepsilon}$	$0.67_{\pm\varepsilon}$	$0.85_{\pm\varepsilon}$
	Covertype	$0.56_{\pm\varepsilon}$	$0.41_{\pm\varepsilon}$	$0.00_{\pm\varepsilon}$	$0.25_{\pm\varepsilon}$	$0.62_{\pm\varepsilon}$	$0.88_{\pm\varepsilon}$	$1.00_{\pm\varepsilon}$

Tableau 2 – Résultats sur plusieurs jeux de données avec vecteur complet (*vc*) et vecteur tronqué (*vt*) ($\varepsilon = 0.0009$)

de *Kruskal-Wallis* pour vérifier l'inégalité des moyennes obtenues sur les critères observés sur l'ensemble des algorithmes, puis nous complétons ces tests par des comparaison deux à deux en réalisant des tests de rang de *Wilcoxon* pour mettre en évidence la significativité statistique de ces inégalités.

5.1 Analyse Globale

Les analyses ci-dessous s'appuient sur les résultats présentés dans la Tableau 2 dont les CDFs sont illustrées Figure 1.

Tests Statistiques : Un test de Kruskal-Wallis pour chaque expérience nous indique qu'il y a une différence significative entre les mesures de précision de chacun des 3 algorithmes ($p < 0.01$). De plus, le test des rangs signés de Wilcoxon met en évidence une différence significative entre chaque paire d'algorithmes ($p < 0.01$).

Diversité : En ce qui concerne les expériences sur le jeu de données de contrôle nous observons comme attendu que lorsque nous fournissons un vecteur optimal x^* les deux algorithmes de CMAB diversifient à 100%. Néanmoins, pour chaque jeu de données, lorsque nous perdons 25% de l'information du vecteur d'origine, alors la diversité décroit pour les deux algorithmes de CMAB. En effet, ils ne réussissent pas à trouver la bonne politique en regard de la règle de correspondance cachée entre récompenses et dimensions du contexte puisqu'une partie pertinente de ce contexte a été tronquée. En revanche, même si *LinUCB* ne diversifie pas autant que dans le cas où nous lui fournissons un vecteur plus complet, *SW-LinUCB* quant à lui offre dans ces mêmes conditions une meilleure diversification que l'algorithme original. Enfin, comme prévu, *UCB* agit comme un algorithme glouton à bornes supérieures de confiance : il trouve le bras optimal et continue de le tirer tout au long des itérations ce qui résulte en une valeur de diversité proche de 0.

Précision Globale VS Individuelle : Comme attendu, *LinUCB* obtient la meilleure performance globale dans tous les cas et pour tout jeu de données. Sans surprise, la précision globale diminue lorsque nous tronquons le vecteur de contexte, mais il est important de noter que même avec le niveau d'éparsité choisi dans notre expérience, les algorithmes de CMABs restent encore meilleurs que l'algorithme de MAB représenté par *UCB*. Cependant, nous observons dans tous les cas (sauf quand $x = x^*$), que *LinUCB* crée un écart de précision individuelle très important entre les utilisateurs. D'autre part, sur l'ensemble des jeux de données et dans tous les cas (sauf quand $x = x^*$), *SW-LinUCB* perd en précision globale par rapport à *LinUCB* mais en revanche trouve, grâce à son mécanisme de diversification, un meilleur compromis en ce qui concerne la distribution de la précision individuelle. Enfin, pour tous les jeux de données, nous observons que plus l'incomplétude du vecteur de contexte est importante, plus un fossé se crée entre les différentes précisions individuelles d'où résultent distinctement une classe de précisions que l'on peut catégoriser de hautes et une classe de précisions dites basses. De la même manière, on remarque que plus x tend vers x^*, plus la distribution de la précision individuelle est uniformément répartie parmi les utilisateurs.

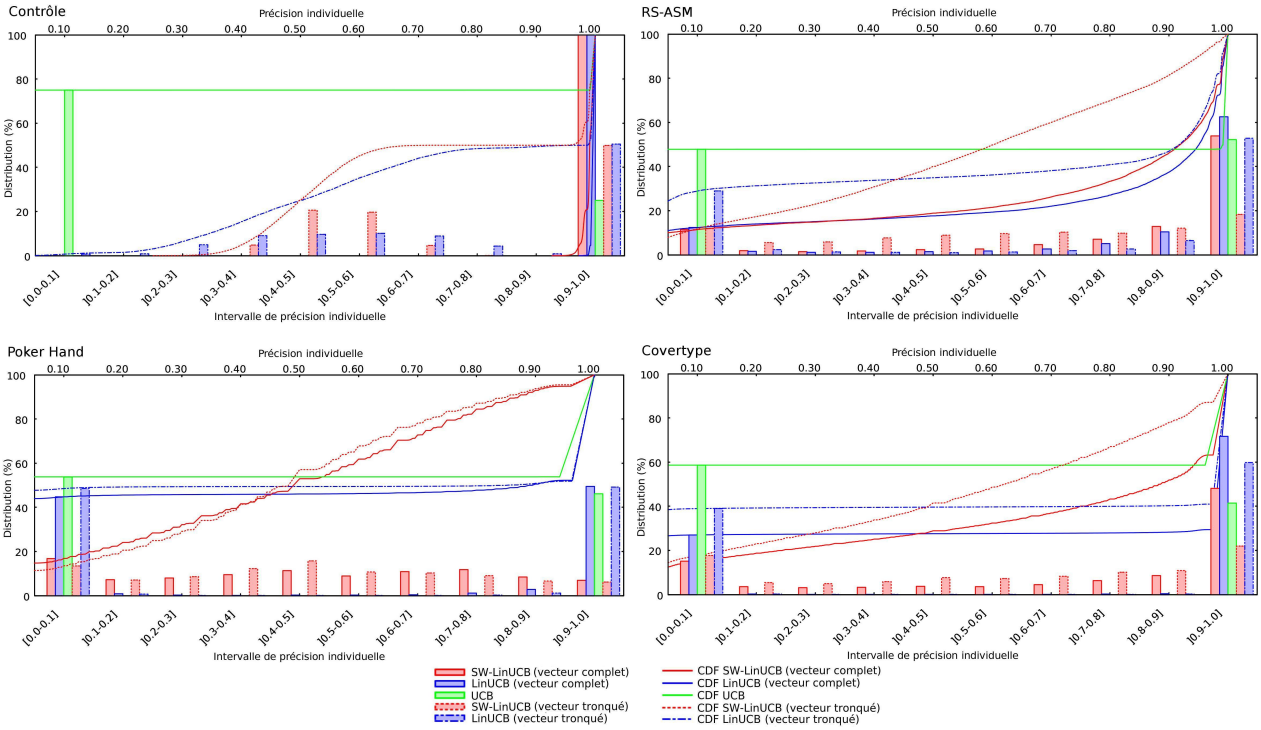


FIGURE 1 – Distribution de la précision individuelle pour chaque algorithme

5.2 Analyse Spécifique sur Covertype

Diversité : On observe que *SW-LinUCB* diversifie plus (*vc* : $Div = 0.47$, *vt* : $Div = 0.41$) que *LinUCB* (*vc* : $Div = 0.44$, *vt* : $Div = 0.35$) alors que l'algorithme *UCB* continue de tirer le même bras tout au long des itérations ($Div = 10^{-3}$). De plus, nous remarquons que lorsque x tend vers x^*, les caractéristiques fournies en tant que dimension du vecteur de contexte permettent à *LinUCB* et *SW-LinUCB* de diversifier d'avantage.

Précision Globale VS Individuelle : On observe que *LinUCB* conserve une meilleure précision globale (*vc* : $Acc = 0.72$, *vt* : $Acc = 0.60$) que *SW-LinUCB* (*vc* : $Acc = 0.69$, *vt* : $Acc = 0.56$). De plus, il est important d'observer que le niveau d'incomplétude du vecteur de contexte n'est pas encore assez important pour permettre à notre algorithme de MAB *UCB* d'être plus précis (0.41). En outre, on observe Figure 1 et Tableau 2, que *SW-LinUCB* reste le meilleur en termes de distribution de la précision individuelle (*vc* : $Q_1 = 0.40$, *vt* : $Q_1 = 0.25$) que *LinUCB* (*vc* : $Q_1 = 0.00$, *vt* : $Q_1 = 0.00$). Ces derniers résultats montrent que notre mécanisme de diversification permet d'augmenter la précision individuelle de la classe dite basse avec la méthode d'origine. Enfin, la comparaison entre les résultats *vc* et *vt* montre que, pour les deux méthodes de CMAB, les précisions globales et individuelles sont toutes deux proportionnelles au niveau d'information et de complétude du vecteur contexte.

6 Conclusion et Perspectives

Dans cet article, nous proposons une nouvelle mesure pour les algorithmes de décision séquentielle visant à évaluer la distribution de la précision individuelle. Nous soutenons que dans certains cas pratiques, la mesure de précision globale n'est pas suffisante pour évaluer les algorithmes de CMAB et que la mesure de précision individuelle doit également être prise en compte. De plus, nous proposons une nouvelle approche adaptée de l'algorithme original *LinUCB* visant à la fois à améliorer la précision individuelle et à maintenir une bonne précision globale. Nous montrons qu'en privilégiant la diversité, notre algorithme *SW-LinUCB* offre un compromis entre précision globale et individuelle que nous pensons mieux adapté à un certain nombre d'applications du monde réel comme les systèmes de recommandations ou les essais cliniques.

Ainsi en perspectives, il semble pertinent de considérer les deux opportunités suivantes : 1) Mettre en place des techniques permettant la construction d'un vecteur plus précis notamment par l'observation approfondie du contexte de l'application concrète qui en découle afin de déterminer les dimensions manquantes et pertinentes ; 2) Concevoir un algorithme pour résoudre le problème multi-objectifs de la maximisation des trois critères de précision globale, de précision individuelle et de diversité. Nous pensons qu'une approche portfolio c'est à dire tirant parti des avantages de plusieurs algorithmes (notamment *LinUCB* et *SW-LinUCB*) pourrait être envisagée.

©AfIA 2018

Références

[1] L. Li, W. Chu, J. Langford, and R. E. Schapire, "A contextual-bandit approach to personalized news article recommendation," in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 661–670.

[2] S. S. Villar, J. Bowden, and J. Wason, "Multi-armed bandit models for the optimal design of clinical trials : benefits and challenges," *Statistical science : a review journal of the Institute of Mathematical Statistics*, vol. 30, no. 2, p. 199, 2015.

[3] S. Agrawal and N. Goyal, "Thompson sampling for contextual bandits with linear payoffs," in *International Conference on Machine Learning*, 2013, pp. 127–135.

[4] D. Bouneffouf and R. Feraud, "Multi-armed bandit problem with known trend," *Neurocomputing*, vol. 205, pp. 16–21, 2016.

[5] S. M. McNee, J. Riedl, and J. A. Konstan, "Being accurate is not enough : how accuracy metrics have hurt recommender systems," in *CHI'06 extended abstracts on Human factors in computing systems*. ACM, 2006, pp. 1097–1101.

[6] H. Robbins, "Some aspects of the sequential design of experiments," *Bulletin of the American Mathematical Society*, pp. 527–535, 1952.

[7] T. L. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules," *Advances in applied mathematics*, vol. 6, no. 1, pp. 4–22, 1985.

[8] P. Auer, "Using confidence bounds for exploitation-exploration trade-offs," *Journal of Machine Learning Research*, vol. 3, no. Nov, pp. 397–422, 2002.

[9] S. Agrawal and N. Goyal, "Analysis of Thompson sampling for the multi-armed bandit problem," in *Conference on Learning Theory*, 2012, pp. 39–1.

[10] J. Langford and T. Zhang, "The epoch-greedy algorithm for multi-armed bandits with side information," in *Advances in neural information processing systems*, 2008, pp. 817–824.

[11] L. Zhou and E. Brunskill, "Latent contextual bandits and their application to personalized recommendations for new users," *In International Joint Conferences on Artificial Intelligence (IJCAI)*, 2016.

[12] D. Bouneffouf, I. Rish, G. A. Cecchi, and R. Feraud, "Context attentive bandits : Contextual bandit with restricted context," *International Joint Conferences on Artificial Intelligence (IJCAI)*, 2017.

[13] S. Craw, B. Horsburgh, and S. Massie, "Music recommenders : user evaluation without real users ?" *In International Joint Conferences on Artificial Intelligence (IJCAI)*, 2015.

[14] A. Lacerda, "Contextual bandits for multi-objective recommender systems," in *Intelligent Systems (BRACIS), 2015 Brazilian Conference on*. IEEE, 2015, pp. 68–73.

[15] M. Ge, C. Delgado-Battenfeld, and D. Jannach, "Beyond accuracy : evaluating recommender systems by coverage and serendipity," in *Proceedings of the fourth ACM conference on Recommender systems*. ACM, 2010, pp. 257–260.

[16] X. Wang, Y. Guo, and C. Xu, "Recommendation algorithms for optimizing hit rate, user satisfaction and website revenue." in *In International Joint Conferences on Artificial Intelligence (IJCAI)*, 2015, pp. 1820–1826.

[17] A. Ashkan, B. Kveton, S. Berkovsky, and Z. Wen, "Optimal greedy diversity for recommendation." *In International Joint Conferences on Artificial Intelligence (IJCAI)*, pp. 1742–1748, 2015.

[18] L. Hu, L. Cao, S. Wang, G. Xu, J. Cao, and Z. Gu, "Diversifying personalized recommendation with user-session context," *In International Joint Conferences on Artificial Intelligence (IJCAI)*, 2017.

[19] R. Allesiardo, R. Féraud, and D. Bouneffouf, "A neural networks committee for the contextual bandit problem," in *International Conference on Neural Information Processing*. Springer, 2014, pp. 374–381.

[20] T. J. Walsh, I. Szita, C. Diuk, and M. L. Littman, "Exploring compact reinforcement-learning representations with linear regression," in *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*. AUAI Press, 2009, pp. 591–598.

[21] A. Garivier and E. Moulines, "On upper-confidence bound policies for switching bandit problems," in *International Conference on Algorithmic Learning Theory*. Springer, 2011, pp. 174–188.

[22] P. Whittle, "Restless bandits : Activity allocation in a changing world," *Journal of applied probability*, vol. 25, no. A, pp. 287–298, 1988.

[23] A. Goëffon, F. Lardeux, and F. Saubion, "Simulating non-stationary operators in search algorithms," *Applied Soft Computing*, vol. 38, pp. 257–268, 2016.

[24] S. Li, A. Karatzoglou, and C. Gentile, "Collaborative filtering bandits," in *The 39th International ACM SIGIR Conference on Information Retrieval (SIGIR)*, 2016.

[25] J. Katsnelson and S. Kotz, "On the upper limits of some measures of variability," *Theoretical and Applied Climatology*, vol. 8, no. 1, pp. 103–107, 1957.

Une alternative à la distinction 'continuant' vs 'occurrent'

G. Kassel

Laboratoire MIS, Université de Picardie Jules Verne
33 Rue Saint-Leu, 80039 Amiens Cedex 1

Gilles.kassel@u-picardie.fr

Résumé
Nous posons les bases d'une ontologie de haut niveau de particuliers dont les principes de structuration sont radicalement différents de l'opposition 'continuant' vs 'occurrent' classiquement retenue en Ontologie Appliquée. Ces principes découlent d'une analyse nouvelle de la nature des entités dites « survenantes » ou « occurrentes ». L'analyse intègre d'une part des travaux récents en ontologie des processus, rapprochant ces derniers des objets physiques dans leur manière d'endurer (Stout, Galton). Elle propose par ailleurs de distinguer *processus* et *événements*, identifiant ces derniers à des objets abstraits de pensée côtoyant les *propositions*. Enfin, nous ouvrons notre inventaire ontologique aux *propriétés* et aux *faits* (Fine, Armstrong). Ce cadre permet de rendre compte du monde physique dans ses aspects à la fois statiques et dynamiques et, également, de la façon dont des agents conçoivent son histoire. Les primitives jouent les rôles suivants : les *faits* rendent compte de la vie des substances – *objets* et *processus physiques* – tandis que les *événements* rendent compte, pour des sujets cognitifs, de l'histoire de la vie de ces substances.

Mots Clef
Ontologie Formelle, Objet, Processus, Événement, Propriété, Fait, Proposition, Continuant, Occurrent

Abstract
We lay the foundations of a high-level ontology of individuals whose structuring principles are radically different from the 'continuing' vs. 'occurrent' opposition conventionally adopted in Applied Ontology. These principles are derived from a new analysis of the nature of "happening" or "occurring" entities. The analysis integrates, on the one hand, recent work in ontology of processes, bringing them closer to physical objects in their way of enduring (Stout, Galton). It also proposes to distinguish *processes* and *events*, identifying the latter with abstract objects of thought alongside *propositions*. Finally, we open our ontological inventory to *properties* and *facts* (Fine, Armstrong). This framework makes it possible to account for the physical world in its both static and dynamic aspects, as well as the way agents conceive its history. Primitives play the following roles: *facts* account for the life of substances – *physical objects* and *processes* - while *events* account, for cognitive subjects, for the life history of these substances.

Keywords
Formal Ontology, Object, Process, Event, Property, Fact, Proposition, Continuant, Occurrent

1 Introduction
Les catégories de *processus* et d'*événement* paraissent avoir un destin intimement lié dans les théories métaphysiques courantes. Toutes deux incarnent la dynamicité du monde en correspondant à des entités qui « occurrent » ou « surviennent ». Elles s'opposent aux objets et à la *matière* dont ceux-ci sont constitués, ces entités incarnant la stabilité du monde en « endurant » ou « continuant ». Une doctrine solidement ancrée considère la relation entre processus et événement comme analogue à celle entre matière et objet : les processus sont la « matière » des événements, ce qui revient à dire que les événements sont « constitués » de processus [31][16][5][38]. Selon cette conception, processus et événements sont des entités concrètes habitant une même région spatio-temporelle du monde physique. Cette doctrine ne semble toutefois pas être gravée dans le marbre. De nouvelles propositions dans plusieurs domaines de la métaphysique la remettent même en cause.

Quelques travaux récents sur la métaphysique des processus, notamment, assimilent ces derniers plutôt à des entités portant temporairement des propriétés et capables de changer à la manière d'endurants [40][41][13]. Bref, il s'agirait de continuants occurrents « survenant » [42], à moins qu'il ne s'agisse d'occurrents continuants, c'est-à-dire d'entités qui, tout en étant étendues dans le temps et ayant des parties temporelles, auraient temporairement des propriétés et seraient susceptibles de changer [38][39]. Un enjeu, lié à la caractérisation de la nature des processus, est ainsi la question de l'ontologie du temps et de l'occupation du temps [5]. Un enjeu corrélé est de préciser la notion de « survenue » par rapport à celle d'« existence » lorsque ces notions se rapportent à des occurrents.

©AfIA 2018

Parallèlement, du côté de la métaphysique des événements, même si la conception de Davidson [7][8] d'événements en tant que des particuliers concrets continue de tenir le haut du pavé, la littérature n'en finit pas d'exprimer des interrogations. Certains auteurs ont très tôt exprimé leur scepticisme quant à l'existence d'événements [21][19], en tout cas d'événements tels que définis par Davidson. Ces doutes ont été relayés récemment par des auteurs travaillant sur l'ontologie de l'action (les actions étant réputées être une espèce d'événements), du fait de difficultés à articuler la conception Davidsonnienne des événements avec notre connaissance de la phénoménologie des actions [22][37]. Par ailleurs, des auteurs admettant la co-existence de processus et d'événements sont tentés d'identifier ces derniers à des entités abstraites [17], reprenant une proposition historiquement faite par Chisholm [4] ou Wilson [46]. De fait, si l'existence de deux catégories ontologiques distinctes - *processus* et *événements* – ne semble pas remise en cause et s'avère même utile [39], en tout cas des doutes existent sur le bienfondé de la doctrine de la constitution d'événements par des processus.

Notre objectif dans cet article est justement de proposer un cadre ontologique qui soit suffisamment cohérent pour qu'il représente une alternative crédible à cette thèse de la constitution. Notre proposition repose tout autant sur une démarche que sur un cadre conceptuel organisant différemment les oppositions entre catégories ontologiques. Elle doit son origine à une proposition d'Antony Galton [13][14] de substituer à la distinction 'continuant' vs 'occurrent' la distinction EXP vs HIST (Fig. 1) entre le monde tel qu'il se déroule et son histoire [14] :

> *[…] processes differ markedly from events in their relation to change. Whereas events are fixed items of history which cannot be described as undergoing change, processes are more like ordinary objects in that they can be directly present at one time and can undergo change as time proceeds. This leads to a fundamental ontological distinction between EXP, the dynamic experiential world of objects and processes as they exist at one time, and HIST, the static historical overview populated by events that are generated by the ongoing process in EXP.*

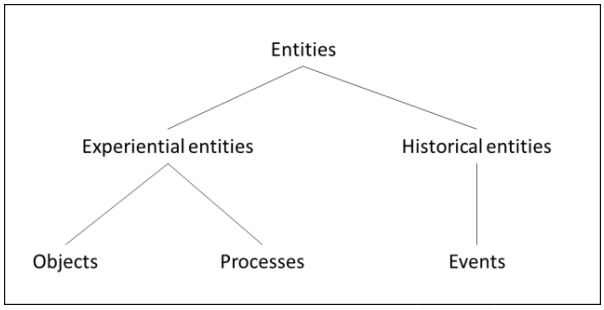


FIGURE 1 – La distinction entre entités expérentielles et historiques (tiré de [14])

Dans cet article, nous reprenons à notre compte cette distinction entre EXP et HIST, toutefois, là où Galton se contente d'y voir une différence de perspectives de description du monde, nous proposons de l'identifier à une distinction métaphysique radicale. La thèse que nous soutenons est que les objets et processus physiques existent dans le monde et ont une vie indépendamment de notre façon de les penser tandis que les événements sont des objets abstraits de pensée, des construits, dont la fonction cognitive essentielle est de disposer d'histoires du monde. Pour résumer notre point de vue, en redéfinissant les catégories introduites par Galton : EXP est le monde physique en mouvement tandis que HIST en est son histoire (passée, présente et à venir) construite par des sujets humains.

Cette thèse est tout à la fois le fruit d'une démarche de recherche métaphysique spécifique et une réponse à un questionnement exprimé par plusieurs auteurs (dont Galton lui-même) quant aux modes d'existence contrastés entre processus et événements, quand ce questionnement ne remet pas en cause l'existence même des événements.

Notre démarche suit le travail méthodologique d'enquête auquel nous convie Maurizio Ferraris [10] pour traquer la frontière entre faits et interprétations (pp. 72-74) :

> *Le point n'est donc pas d'affirmer qu'il y a une discontinuité entre les faits et les interprétations, mais plutôt de comprendre quels sont les objets construits et quels sont ceux qui ne le sont pas […] Ce travail consiste à faire une distinction méticuleuse entre l'existence des choses qui n'existent que pour nous – des choses qui n'existent que s'il y a une humanité – et des choses qui existent même en l'absence de l'humanité.*

Ce travail d'enquête nous conduit à positionner les processus physiques du côté des choses existant même en l'absence de l'humanité et les événements du côté des construits. Plus spécifiquement, le cadre ontologique auquel nous parvenons est le suivant :

- Le monde physique est peuplé de substances – objets et processus – qui, en endurant, assurent sa stabilité tout autant que sa dynamicité ;
- Ces substances, portant temporairement des propriétés et entretenant des relations avec d'autres substances, ont une vie, laquelle consiste en des faits existant dans le monde physique ;
- Des sujets cognitifs, plongés dans le monde physique, se représentent au moyen d'événements l'histoire passée, présente et future du monde pour interagir avec lui.

Dans l'article, nous commençons par présenter en Section 2 notre cadre ontologique. La présentation reprend des éléments publiés dans [25] en élargissant le cadre à la catégorie des *faits*. Ceci permet de distinguer clairement, d'une part, la vie des substances et, d'autre part, l'histoire de cette vie. Enfin, en Section 3, nous discutons l'opposition entre 'continuants' et 'occurrents' et montrons qu'elle devient caduque pour organiser le haut niveau d'une ontologie de particuliers.

2 Notre cadre ontologique

Avant de mettre en scène nos primitives ontologiques, nous précisons le type d'ontologie que nous cherchons à établir d'où dérivent des contraintes quant au choix des primitives.

2.1 Des principes de choix de nos primitives ontologiques

Comme point de départ, indiquons que nous suivons le projet de Peter Strawson [43] d'établir une métaphysique descriptive visant à décrire « *the actual structure of our thought about the world* ». L'objectif est ainsi d'établir des catégories et notions rendant compte de la façon dont nous concevons le monde et, adoptant une perspective contemporaine de l'ontologie, nous retenons deux modes principaux de structuration du monde. D'une part, nous retenons que notre appréhension du monde nous conduit à distinguer trois types de réalité : physique, mentale et sociale. Ce découpage de la réalité, adopté par les « nouveaux réalistes » comme Ferraris, relève de la reconnaissance de *modes d'existence* distincts. Un second découpage s'appuie sur la notion de *niveau d'abstraction* : si nos conceptions de sens commun nous révèlent des entités à un niveau mésoscopique, des conceptions savantes instruites par des sciences nous révèlent des entités situées à un niveau plus élevé (ex : astronomique) ou plus bas (ex : microscopique). Les liens entre ces niveaux tiennent davantage de la relation de *constitution* que de la relation *tout/parties* [28]. L'important, pour la caractérisation de nos catégories et notions, est de respecter ces niveaux d'abstraction. Ainsi, si une table ou tout autre objet matériel nous apparaît au niveau mésoscopique comme un objet, le contenu de la même région spatio-temporelle peut être décrit au niveau nano comme un tourbillon de particules dans du vide. La description au niveau nano révèle d'autres entités et ce n'est pas pour autant qu'une table ou tout autre objet matériel doit être identifié à une « masse dynamique » (nous nous démarquons ici de métaphysiciens soutenant des ontologies mono-catégorielles priorisant la primitive du processus, comme on peut le voir chez Johanna Seibt [34]).

Pour le choix de nos primitives ontologiques, une contrainte que nous posons est leur utilité à rendre compte du plus grand nombre possible de niveaux. Ce critère qualifie en priorité l'objet physique dont la présence est attestée à pratiquement tous les niveaux (excepté sans doute au niveau nano) aussi bien dans les théories cognitives de sens commun qu'en physique et dans la théorie de la mesure. Suivant la conception que nous allons donner du processus physique, en proposant dans la section suivante une relation duelle entre processus et objets physiques, cette primitive bénéficiera de la même robustesse.

2.2 Les substances

Pour débuter notre inventaire, nous adoptons une conception classique de l'objet physique. Un objet physique matériel est quelque chose :

- o_i) existant à des instants ;
- o_ii) ayant des propriétés à des instants (ex : couleur, odeur, masse, volume) ;
- o_iii) pouvant changer dans le temps.

Ces propriétés traduisent la conception d'une entité 3D existant à des instants et endurant dans le temps en ayant des propriétés pouvant varier dans le temps. Les propriétés (o_ii) et (o_iii) caractérisent la vie de l'objet physique : nous y revenons en §2.3. En extension, des exemples sont des objets maximalement connectés, qu'ils soient inertes (ex : une pierre, une pomme détachée d'un arbre, une molécule d'eau, une planète), animés-vivants (ex : un être humain, une fleur, un arbre) ou artefactuels (ex : une chaise, un presse-papier, une télévision). Dans le cas des artefacts, nous verrons en §2.3 qu'il s'agit d'objets physiques « simpliciter » dotés d'une vie sociale.

La caractérisation que nous donnons maintenant des processus physiques est largement fondée sur l'idée de continuants dynamiques développée principalement par Rowland Stout [40][41], Antony Galton [13][14], Galton et Riichiro Mizoguchi [16] Un processus physique est quelque chose :

- p_i) existant à des instants ;
- p_ii) ayant des propriétés à des instants (ex : direction, vitesse d'exécution, niveau sonore, amplitude spatiale),
- p_iii) pouvant changer dans le temps.

Nous retrouvons la même caractérisation que pour l'objet physique, un ensemble de trois propriétés définissant ce que nous considérons être une *substance* dans cet article. En extension, des exemples de processus physiques sont : le mouvement d'un objet physique (conduisant au déplacement ou à la rotation de l'objet sur lui-même) ; la croissance en taille d'un corps physique ; le processus de vie d'une personne ; le mûrissement d'un fruit ; l'oxydation d'un objet métallique ferreux ; la fonte d'un glacier.

Reprenons les propriétés attribuées aux processus physiques une à une pour les éclairer. On trouve la première (p_i) exprimée chez Stout [40] :

> *The phrase, 'What is happening now', is naturally taken to denote a whole process ; and we do want to claim that what is happening now is literally identical with what is happening at some other time – the very same process.*

Cette thèse repose sur un premier engagement fort quant à la nature des processus physiques : la cause fait partie du processus, lequel processus constitue le moteur de changements. Dans le cas d'un changement de localisation d'un objet, ce changement (ou déplacement) résulte d'un mouvement de l'objet – le processus. Le fait d'associer la cause au processus, toujours selon Stout, est nécessaire pour tordre le cou à la conception Russellienne d'un mouvement comme série d'états successifs [41][1] :

> *[The] motion should not be understood in Russell's way as the arrow being in one state and then in another and in the meantime being in all the intervening states. The arrow's motion is what gets it*

1 L'idée d'encapsuler la cause dans le processus n'est en fait pas nouvelle. On la retrouve en effet dans la notion de comportement de Fred Drestke [9]. Selon Drestke, un « comportement » ou « processus causal » consiste dans le fait qu'une cause provoque un mouvement. Le processus physique que nous caractérisons ici correspond ainsi au comportement ou processus causal de Drestke.

©AfIA 2018

through this continuous series of states - it effects the transition.

La conception Russellienne est en fait héritière de la philosophie de l'école d'Élée qui, en concevant précisément des déplacements comme une série d'états, donc d'immobilités, a tourné le dos à une compréhension du mouvement comme un fluant « indivisible », pour reprendre la terminologie de Bergson [2][2].

Venons-en à la possibilité pour un processus de pouvoir porter des propriétés pouvant varier dans le temps, les propriétés (p_ii) et (p_iii). Il s'agit, selon Galton [13], d'une caractéristique rapprochant les processus des objets :

Like objects, processes can change: the walking can get faster, or change direction, or become limping. All around us processes undergo changes: the rattling in the car becomes louder, or change rhythm, or may stop, only to start again later. The flow of the river becomes turbulent; the wind veers to the north-west.

On notera que le fait d'avoir identifié le processus au moteur du changement et non au changement lui-même est déterminant dans la caractérisation de continuant du processus. Les adeptes d'une conception du processus comme un changement occupant du temps avancent qu'un changement ne peut lui-même changer et que, dès lors, la seule substance existante est celle portant le processus (l'argument a été réitéré récemment par exemple par Thomas Crowther [6]). Ainsi, si la vitesse du mouvement d'un corps vient à changer entre deux instants T_1 et T_2, ce changement est à attribuer non au mouvement mais à ce corps se mouvant. Par exemple, Paul possède la propriété de 'Marcher à la vitesse de 5 km/h' à T_1 et la propriété de 'Marcher à la vitesse de 6 km/h' à T_2. Au contraire, la conception que nous adoptons nous conduit à considérer qu'un processus de marche de Paul a une 'vitesse de 5 km/h' à T_1 et une 'vitesse de 6 km/h' à T_2.

Pour compléter la caractérisation de nos processus physiques, évoquons maintenant un engagement ontologique supplémentaire, à savoir le fait qu'un processus ne soit pas un continuant flottant dans l'air, mais soit ancré dans un *objet support* : il s'agit du mouvement d'une *flèche*, du mûrissement d'un *fruit*, de la fonte d'un *glacier*, etc. Pour rendre compte de ce lien fort constitutif, nous reprenons à notre compte la relation d'*énaction* introduite par Galton et Mizoguchi [16]. Pour ces auteurs, dire qu'un objet « énacte » un processus revient à dire qu'un objet porte un processus « externe » ou « comportement » (p. 94) :

The key notion is that an object, considered from a particular point of view, is characterized in terms of the processes it enacts. These are what we call the external processes or behavior of the object. This behavior arises as a result of various internal processes which causally contribute to it.

Cette caractérisation de la relation d'énaction par Galton et Mizoguchi repose sur une conception de l'objet comme interface entre des processus internes et externes. Elle revient à introduire une relation duelle de dépendances qu'expriment les propriétés suivantes :

o_iv) l'intégrité (existence) de tout objet physique est maintenue par des processus physiques ;

p_iv) tout processus physique est énacté par un objet physique.

De ces propriétés, nous pouvons déduire que l'existence d'un objet physique dépend de celle d'autres objets situés à des niveaux (ontologique ou d'abstraction) inférieurs et, de façon duale, que l'existence d'un processus physique dépend de celle d'autres processus situés à des niveaux inférieurs. Ainsi, un processus de marche d'une personne n'est possible que si, notamment et à un premier niveau, des processus énactés par ses jambes existent concomitamment, et que si, notamment et à un niveau inférieur, des processus physiologiques énactés par ses organes existent concomitamment. Signe de cette relation duelle, l'objet et le processus physique sont habituellement crédités d'une vie continue. Par exemple, rien dans la notion de processus de marche n'évoque une fin possible : celle-ci peut se produire contingentement par exemple suite à une fatigue de la personne ou suite à une décision de la personne.

2.3 La vie des substances

Notre inventaire ontologique est à ce stade composé de substances, des *objets* et des *processus* physiques. S'agissant d'entités persistant dans le temps tout en conservant leur identité, elles existent à différents instants et ceci nous conduit à parler de leur 'vie'. La 'vie' de ces entités est limitée dans le temps entre le moment où elles acquièrent une existence et le moment où elles cessent d'exister. Certaines ont une vie brève, à l'échelle humaine, comme ces particules élémentaires étudiées par les physiciens des hautes énergies dont la durée de vie (et celle de leurs mouvements) n'est qu'une infime fraction de seconde. D'autres ont une vie longue voire semblant éternelle, toujours à l'échelle humaine, à l'instar des astres et galaxies et de leurs mouvements.

Dans cette section, nous donnons corps au concept de *vie* des objets et des processus physiques en lui conférant une certaine légitimité ontologique. Sans pour autant en faire une catégorie ontologique à part entière, nous précisons cette notion en la rattachant à deux primitives ontologiques ayant déjà acquis leurs lettres de noblesse en ontologie formelle (même si plusieurs théories existent à leur sujet), celle de *propriété* et celle de *fait*.

Pour débuter par les *propriétés*, une clarification terminologique s'impose : le terme « propriété » est communément utilisé pour dénoter, d'une part, des universaux (types) et des tropes (instances) caractérisant les substances [1] et, d'autre part, des concepts ou catégories qui structurent nos représentations et théories du monde [27]. Ces entités possèdent des modes d'existence et d'attribution différents. Comme nous allons le montrer par la suite, la cohérence de notre cadre ontologique nécessite de les considérer toutes les

2 Bergson attribue à cette école le fait que le processus ait été relégué au rang d'accident [2] : « *La métaphysique est née, en effet, des arguments de Zénon d'Élée relatifs au changement et au mouvement. C'est Zénon qui, en attirant l'attention sur l'absurdité de ce qu'il appelait mouvement et changement, amena les philosophes – Platon tout le premier – à chercher la réalité cohérente et vraie dans ce qui ne change pas* ».

deux. Notamment, pour la notion de *vie* qui nous intéresse dans ce paragraphe, nous allons voir qu'il est pertinent de distinguer entre vie « physique » et vie « sociale » des substances. Pour les distinguer, nous y référerons en utilisant respectivement les termes « propriété physique » et « propriété conceptuelle ».

Le lecteur trouvera dans [1] un large traitement des propriétés *physiques* dont l'existence est supposée. Ces propriétés correspondent à des « manières d'être » d'entités substantielles. Différentes théories ont été proposées pour rendre compte de leur nature. Pour certains théoriciens, les propriétés sont des universaux, autrement dit des entités se répétant à l'identique dans des substances ; d'autres théoriciens considèrent que les propriétés sont des tropes, autrement dit des particuliers inhérents à leur porteur. Dans cet article, nous ne prendrons pas parti. Précisons que des conceptions hybrides combinant universaux et tropes ont été proposées dans la littérature. En Ontologie Appliquée, les propriétés faisant l'objet d'une expérience perceptuelle de la part de sujets (ex : couleur, forme, odeur, masse) sont souvent conçues comme des tropes. On trouve un tel traitement dans BFO [18] et DOLCE [29].

Concernant les propriétés *conceptuelles*, celles-ci participent de la construction d'une partie du monde appelée « réalité sociale » [32][33]. Pour notre propos, nous nous contenterons de considérer, en empruntant la terminologie de Amie Thomasson [44], que certaines propriétés conceptuelles portent sur des objets physiques (typiquement, des propriétés fonctionnelles comme 'Être un presse-papier') en en faisant des objets « sociaux concrets », tandis que d'autres portent sur des objets non physiques (ex : 'Être une monnaie', 'Être une loi', 'Être une nation'), en en faisant des objets « sociaux abstraits ». Ces derniers correspondent à la catégorie des 'Non-Physical objects' introduite dans DOLCE pour rendre compte des objets n'existant que parce que des agents les conçoivent et communiquent à leur propos. On notera que cette distinction *physique* vs *non physique* tient également pour les processus. Parmi les processus non physiques nous pouvons distinguer des changements de propriétés conceptuelles pour des objets physiques (ex : 'devenir veuf pour une personne', 'perdre de sa valeur faciale pour un billet de banque') ou pour des objets non physiques (ex : 'se déprécier pour une monnaie', 'tomber en désuétude pour une loi'). Nous en profitons pour noter que, dans cet article, nous ne traiterons pas de processus non-physiques.

La catégorie de *propriété* étant introduite, nous continuons à étendre notre inventaire en admettant cette fois celle de *fait* pour rendre compte d'entités comme 'Paul est embarrassé' ou 'Paul est à côté de Marie'. La thèse de l'existence de telles entités, avancée par de nombreux philosophes (ex : Kit Fine [11], Donald Armstrong [1]), est une thèse compagnon de la théorie réelle des propriétés physiques : l'existence simultanée de la substance 'Paul' et de la propriété 'Être embarrassé' ne signifie pas pour autant que la substance 'Paul' exemplifie à un instant donné la propriété 'Être embarrassé', le même constat tenant pour la substance 'Paul' et la relation 'Être à côté de Marie'. Le fait, ou la « circonstance » (pour reprendre le terme de Fine), correspond à ce lien interne unissant substance et propriété/relation en une entité à part entière. L'argument principal de l'existence des faits est qu'ils constituent un *vérifacteur* (truth maker), autrement dit ce qui rend vrai dans le monde des propositions comme 'Paul est embarrassé' ou 'Paul est à côté de Marie' [1].

Le lecteur aura noté l'embarras notationnel dans lequel on se trouve quand on se pose la question *Que sont les faits ?* ou, comme Andrea Iacona [23], *Que sont les propositions ?*. On a vite fait d'élaborer un discours du type « il est vrai que le ciel est bleu car le ciel est bleu », voulant signifier que « la *proposition* selon laquelle 'le ciel est bleu' est vraie car le *fait* 'le ciel est bleu' existe ». Pour résoudre cet embarras, dans cet article nous adoptons une notation commune (pour les philosophes étudiant les faits) faisant apparaître la structure du fait en ses constituants : <Ciel, Être bleu>. Une même syntaxe a été proposée également dans la littérature pour les propositions et les événements, revenant à expliciter leurs constituants. Par ailleurs, pour lever toute ambiguïté, lorsque nous parlons concomitamment de faits et de propositions, nous explicitons la catégorie d'entité en postfixant la notation d'un indice : $<>_F$, $<>_P$ (ex : $<\text{Paul, Être embarrassé}>_F$ tient pour le *fait*, tandis que $<\text{Paul, Être embarrassé}>_P$ tient pour la *proposition*).

De même que nous avons distingué supra entre propriétés physiques et concetuelles, il est important de distinguer plusieurs types de faits. Tels que nous les avons introduits, les faits/circonstances concernent des substances (objets et processus physiques) et il est communément admis que le lien attachant la substance à une propriété correspond à une instanciation dénotée par « est » : Paul *est* 'Être embarrassé' ; Paul *est* 'Être à côté de Marie'. Pour ces faits physiques, une conception courante est par ailleurs de considérer le temps comme constituant : $<\text{Paul, Être embarrassé, À l'instant}>_F$, $<\text{Paul, Être à côté de Marie, À l'instant}>_F$. La raison en est que les substances ont leurs propriétés *à* des instants (cf. les propriétés (o_ii) et (p_ii)). Par contre, toujours pour ce qui concerne les substances, il est d'autres faits pour lesquels l'association d'une propriété à une substance correspond à une stipulation humaine pouvant reposer sur une convention sociale (pour reprendre l'analyse de Searle [32]) : ce morceau de papier *compte pour* 'Être un billet de banque de 10 euros' (pour un agent ou une communauté d'agents, et dans certaines circonstances) ; Ce galet *compte pour* 'Être un presse-papier'[3]. Les substances communément considérées en

3 Précisons que plusieurs théories ontologiques ont été proposées dans la littérature pour rendre compte de telles stipulations consistant à attribuer une fonction à un objet. Suivant la théorie des artefacts développée par Borgo *et al.* [3], attribuer une fonction (ex : 'Être un presse-papier') à un objet physique (ex : un galet) revient à créer un second objet physique – l'artefact galet-presse-papier – constitué du premier objet physique. Au contraire, et en cohérence avec notre choix d'établir une ontologie descriptive, l'interprétation que nous donnons à l'acte cognitif est celle de l'ajout d'une propriété à ce même objet physique [24].

©AfIA 2018

exemples, aussi bien par Searle et par Thomasson, sont des objets physiques. Maintenant, le fait d'admettre les processus physiques comme substances nous procure d'autres exemples : tel mouvement des lèvres *compte pour* un sourire ; telle élévation du bras *compte pour* une indication de tourner à gauche. La différence avec les faits précédents est que ces derniers reposent sur l'attribution de propriétés par des agents à des substances. Ce ne sont plus des « faits bruts » (pour reprendre la terminologie de Searle) mais des « faits sociaux » construits. Les conditions d'existence des faits dès lors varient, suivant qu'ils correspondent à une instanciation indépendante de toute pensée humaine, pour les faits bruts, ou à une stipulation humaine, pour les faits sociaux.

Les exemples de faits que nous venons de prendre illustrent ce que nous entendons par « vie » d'une substance : il s'agit d'une collection de faits temporels se rapportant à la substance (pour lesquels la substance apparaît comme constituant). Les exemples que nous avons considérés concernent plutôt des objets. À ce propos, on notera que la propriété (p_iv) caractérisant les processus – le fait qu'un processus soit énacté par un objet à un temps – inscrit les processus dans la vie des objets (les objets ont une vie processuelle). Mais, intéressons-nous maintenant à la vie des processus. Une catégorie importante de faits participant de la vie des processus correspond à la *perpétuation* des processus. De tels phénomènes interviennent lorsque, par exemple, le mouvement d'une masse d'air « perpétue » le mouvement d'une feuille, le déplacement de la masse d'air « provoquant » le déplacement de la feuille ; ou bien, lorsque le mouvement d'un bras « perpétue » le mouvement d'une montre portée au poignet, une élévation du bras « provoquant » une élévation de la montre. Dans les descriptions que nous venons de donner, en évoquant, d'une part, des mouvements-processus « perpétuant » d'autres mouvements-processus et, d'autre part, des déplacements-événements « provoqués » par d'autres déplacements-événements, nous avons fait attention à distinguer les processus des événements. Les relations intitulées « perpétue » et « provoque » font partie d'un catalogue de relations causales entre états, processus et événements, tel que décrit par Galton [15]. Pour en rester aux processus (sachant que nous ne donnons notre définition des événements qu'en §2.4), nous retenons du catalogue de Galton la relation *perpetuatesAt(p,p',t)* signifiant que « le processus *p* perpétue (ou entretient) le processus *p'* au temps *t*). Cette relation tient entre deux processus déjà existants et on peut la comprendre comme une propagation de causalité : un processus énacté par un objet entretient (en participant de sa cause) un autre processus énacté, soit par un autre objet (cf. nos exemples), soit par le même objet (ex : le mouvement d'un corps entretient l'échauffement de ce corps). Il est intéressant de noter que les faits de perpétuation représentent une partie de la dynamique du monde. Cette remarque met en avant un argument supplémentaire en faveur de l'existence des faits bruts ou physiques (venant compléter l'argument du vérifacteur).

2.4 L'histoire de la vie des substances

Venons-en à évoquer un construit humain, à savoir la façon dont les humains conceptualisent l'histoire de la vie des substances – processus et objets physiques – peuplant le monde. La principale catégorie ontologique que nous allons ajouter à notre inventaire est celle d'événement. Pour commencer à fixer les idées, précisons d'emblée que la notion d'événement que nous visons est proche de celles théorisées historiquement par Roderick Chisholm [4] ou Neil Wilson [46], faisant des événements des objets abstraits à côté des propositions. Par ailleurs, pour continuer à fixer les idées, considérons un premier exemple : 'la marche que Paul a faite jusqu'à la gare ce matin'. Cet événement concerne la vie de multiples processus, dont des processus de marche de Paul (en utilisant le pluriel, nous imaginons que Paul a pu s'arrêter puis reprendre sa marche, autrement dit que plusieurs processus de marche ont successivement existé), et la vie de multiples objets, dont Paul lui-même. Plus généralement, nous considérons qu'un événement est quelque chose :

e_i) existant pour un sujet à des instants ;
e_ii) jouissant d'une méréologie dérivée et robuste
e_iii) pouvant survenir, mais sans être répété

Un événement est avant tout un objet de pensée, une entité psychologique, servant à des sujets à se représenter l'histoire d'un monde (physique ou non). Le terme « histoire » évoque généralement une histoire passée. Je pense ainsi à 'ma dernière chute de vélo' et à d'autres événements survenus dans le passé et faisant partie de notre imaginaire collectif : 'l'assassinat de César par Brutus', 'le naufrage du Titanic'[4]. Mais l'histoire peut être présente : je pense à l'événement 'l'écriture de cet article' au moment même où j'écris l'article. Elle peut également être future : je pense au 'déplacement en voiture que je dois effectuer demain'.

Comme le montrent ces exemples, les événements existent dans des espaces-temps distincts de ceux dans lesquels existent les substances dont ils relatent l'histoire. Ce sont des entités abstraites endurant dans l'espace cognitif d'un sujet (e_i). Cette caractérisation les oppose aux événements concrets de Davidson [7][8] et entraîne de nouvelles propriétés.

Les deux espaces-temps étant découplés, il convient de noter que les événements jouissent d'une méréologie dérivée (e_ii). Ainsi, dire que l'événement 'l'enfance de Paul' « fait partie » de l'événement 'la vie de Paul' ne signifie pas que les deux événements partagent une même région spatio-temporelle dans l'esprit d'un sujet les pensant. Il faut plutôt entendre que les faits concernés par le premier événement font partie (dans un sens ensembliste) des faits concernés par le second événement. À ce propos, on peut se demander dans quelle mesure les faits concernés par un événement sont déterminés ?

Sur ce point, on peut noter que des événements « simples » correspondent à des épisodes de vie d'un seul objet ou d'un seul processus bornés dans un

4 Dans cet article, pour des raisons de place et compte tenu de notre propos, nous ne développerons pas la dimension sociale des événements et nous contenterons donc de considérer des événements pensés par un sujet unique, qui plus est humain.

espace-temps déterminé, comme dans le cas de ‘la chute de Paul à l’instant’. Cet exemple illustre la différence conceptuelle existant entre le processus et l’événement, ce dernier étant obtenu en fixant des limites temporelles à un processus et en le considérer dans la durée, comme le soutiennent Galton et Mizoguchi [16] :

> *We maintain, on the contrary, that so far from being a mark of short duration, boundedness is a precondition for the assignment of any definite duration: processes endure, but only once we have assigned bounds to them can we speak of duration, and the act of assigning bounds means that we have switched our attention from the process to an event.*

Par contre, comme nous l’avons déjà fait remarquer, certains événements tel ‘la marche que Paul à faite ce matin jusqu’à la gare’ sont plus complexes en s’accommodant de faits très différents. Cet événement ne tient pas compte de l’itinéraire suivi par Paul pour se rendre à la gare. On peut parler ici de robustesse méréologique des événements, par opposition à l’essentialisme méréologique qui leur est souvent attribué (à savoir, le fait qu’un événement soit essentiellement déterminé par ses parties). Comme le note Achille Varzi [45], en considérant des événements correspondant à des épisodes de vie de larges systèmes économiques et politiques tels ‘la révolution industrielle’ ou ‘la seconde guerre mondiale’, les événements comportent une indétermination intrinsèque.

Poursuivons notre caractérisation des événements en évoquant une propriété qui les qualifie ordinairement d’« occurrents », à savoir le fait qu’ils puissent survenir (e_iii). Le même qualificatif est également ordinairement attribué aux processus mais, dans leur cas, nous nous sommes contentés d’évoquer leur existence. Pour les événements, les choses se présentent différemment et ceci est dû au fait de les identifier à des objets de pensée. Ceci conduit en effet à distinguer leur *existence* de leur *survenue*. La définition générale que nous retenons de cette propriété est la suivante :

> Soit *e* un événement existant pour le sujet *s* au temps *t* ; l’événement *e* ‘survient’ à un temps *t’* ssi les faits dont *e* relate l’histoire existent au temps *t’*.

La propriété de survenue des événements peut être considérée comme analogue à la propriété de véracité des propositions : l’existence de faits conditionne la survenue de l’événement, comme elle conditionne la vérité d’une proposition. L’existence de faits de chute de Paul à des instants consécutifs T_1 et T_2 - $<Paul,Tombe,T_1>_F$ et $<Paul,Tombe,T_2>_F$ – correspond à la condition de survenue de l’événement $<Paul,Chute,À\ l'instant>_E$ (pour peu que l’événement soit pensé et que l’instant pensé ‘À l’instant’ corresponde à l’intervalle $[T_1,T_2]$). La relation d’ordre entre t et t’ dans la définition détermine le fait que l’histoire soit passée, présente ou future. Dans le cas où l’événement survient au moment même où il est pensé, cela donne la possibilité au sujet de jouer le rôle d’agent en agissant sur les faits réalisant l’événement[5].

5 Dans [25], nous détaillons de telles situations en les caractérisant en termes de couplages temporels et causaux entre processus et événements.

Tout dernier point, que nous mentionnerons sans le justifier. Nous considérons qu’il existe des événements particuliers non répétés (ex : ‘la chute de Paul de ce matin’) et des types d’événements répétables (ex : ‘la marche matinale de Paul jusqu’à la gare’). Intuitivement, la singularité est liée aux faits concernés : un épisode de vie d’un processus vs une classe d’épisodes de vie d’une classe de processus (ce critère serait à généraliser à des événements ne se référant pas à des processus).

2.5 En résumé

En conclusion de cette section, résumons avec la Figure 2 le chemin parcouru depuis la distinction entre entités expérientielles et historiques de Galton [14] qui nous a servi de point de départ.

Le monde tel qu’il se présente à nous est peuplé en premier lieu de substances – des *objets* et des *processus physiques*. Ces substances, dont on peut constater l’existence à tout niveau d’abstraction, dépendent mutuellement l’une de l’autre. Une raison toutefois de les distinguer tient au fait qu’elles portent des propriétés différentes [20] : un processus physique n’a pas plus de couleur ou de volume qu’un objet n’a de rapidité ou d’amplitude. Toujours dans le monde physique, les substances, tout en portant temporairement des propriétés, jouissent d’une vie correspondant à des faits physiques. Parmi ces faits physiques figurent les perpétuations de processus jouant un rôle important dans la dynamique du monde physique.

Pour basculer cette fois dans le réel construit, des sujets cognitifs pensent le monde et, parmi ces objets de pensée, figurent les *événements*. Positionnés à côté des *propositions*, nous considérons que ces entités, représentant le monde, jouent des rôles cognitifs distincts : si la notion de *vérité* caractérise les propositions, celle de *survenue* caractérise les événements. Par ailleurs, du côté des objets de pensée figurent les faits sociaux rendant compte notamment de la vie sociale des substances.

Dans cette section, nous avons présenté un cadre ontologique général en nous attachant à souligner sa cohérence globale. En guise de conclusion, nous discutons la question figurant dans le titre, à savoir : les distinctions ontologiques que nous venons de voir offrent-elles une alternative à l’opposition ‘continuant’ vs ‘occurrent’ ?

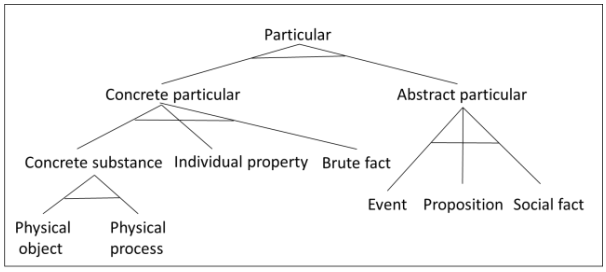


FIGURE 2 – Vue d’ensemble de notre cadre ontologique

©AfIA 2018

3 Discussion

Le cadre ontologique que nous venons d'esquisser remet en cause l'opposition 'continuants' vs 'occurrents' adoptée classiquement en Ontologie Appliquée, notamment par BFO [18] et DOLCE [29]. Cette opposition, rappelons-le, se fonde sur deux doctrines philosophiques caractérisant deux modes distincts d'existence et de persistance pour les entités peuplant le monde : l'*endurantisme* (le fait pour une entité d'exister dans son entièreté à tout moment de son existence et de pouvoir changer dans le temps) et le *perdurantisme* (le fait pour une entité de n'exister que par ses parties à tout moment et de ne pas pouvoir changer dans le temps). Fondamentalement, la remise en cause porte sur deux points :

1) le fait d'admettre dans la classe des continuants une nouvelle catégorie d'entités rendant compte de la dynamique du monde – les processus physiques – avec pour conséquence éventuelle de revoir la notion de continuant ;

2) le fait de devoir réviser radicalement la classe des occurrents pour accueillir les événements conçus comme des objets de pensée, mais également de devoir faire une place aux faits de perpétuation de processus.

3.1 De la notion de *continuant*

La figure type de continuant est celle de l'objet physique matériel, habituellement conçu comme une entité 3D. Cette conception, dont nous rendons compte avec les propriétés (o_i)(o_ii)(o_iii), suppose de distinguer deux notions [12] : l'*existence* de l'objet (souvent dénotée également par le terme « présence ») et son *extension* (ou « localisation ») dans l'espace. La première possède un caractère de tout ou rien – un objet existe, ou non – tandis que la seconde possède un caractère de partialité : si l'objet occupe *totalement* une région spatiale, il n'occupe que *partiellement* chaque partie de cette région spatiale. En ce sens, nous pouvons dire que l'objet *existe* (ou qu'il est présent) « dans sa totalité » à des temps, mais qu'il n'existe (ou qu'il n'est présent) « qu'en partie » dans certaines régions spatiales à des temps. En revanche, l'objet physique ne possède pas d'extension temporelle et cette conception est à opposer à celle de l'entité 4D proposée par certains métaphysiciens (ex : [35]) considérant qu'un objet physique est étendu dans le temps de la même façon qu'il est étendu dans l'espace. L'argument que nous mettons en avant pour défendre la conception 3D de l'objet physique est qu'il s'agit d'une conception intuitive de sens commun alors que la conception 4D est largement contre-intuitive et n'est étayée par aucune donnée de psychologie ou quelconque matériau linguistique. Une conception 4D ne peut donc fonder une ontologie descriptive telle que visée.

Venons-en maintenant aux processus. En les caractérisant au moyen des propriétés (p_i)(p_ii)(p_iii), nous les avons assimilés à des continuants « occurrents » ou « dynamiques », pour reprendre les termes et la conception de Stout [42]. Une telle thèse a été critiquée par Steward [39], arguant du fait que les processus demeurent des occurrents possédant une extension temporelle, ce qui contredit (p_i) :

> *Because processes have temporal parts and temporal extension, they do not exist in their entirety at each moment of their existence, in the way that substances do. A process is an essentially unfolding entity, which is what secures its right to be thought of as an occurrent.*

En conséquence, toujours selon Steward, il est faux de considérer que les processus aient leurs propriétés à des instants (p_ii). Il convient plutôt de considérer qu'ils ont leurs propriétés entre des instants [39] :

> *[Processes] have their properties primarily between times, and this implies that they share some of the characteristics continuants are generally thought to have […] To say that the dripping of a tap was persistent at t1 is a way of saying that t1 was a moment which falls within a period of time over which the dripping was persistent.*

Un élément important du débat nous paraît être la définition en extension de la classe des processus. On notera à ce propos que l'exemple analysé par Steward d'un robinet qui fuit – le seul exemple de processus, du reste, traité dans son [39] – ne correspond pas à notre conception du processus physique. Sémantiquement, le terme 'fuite' évoque un flux d'eau s'échappant anormalement d'un robinet. Supposons que la fuite soit intermittente : chaque flux correspond, selon notre conception, à un processus énacté par une goutte d'eau (ou une quantité d'eau plus importante) ; on peut ajouter que chaque échappement de flux correspond à un processus énacté, cette fois par le robinet ; par contre, la fuite en elle-même, en tant qu'épisode correspondant à l'existence de ces différents processus, est au mieux un événement correspondant à l'histoire du robinet et de ces processus (le terme « au mieux » soulignant la condition qu'un sujet doive penser cette histoire pour qu'elle existe) ; par ailleurs, l'existence d'un événement n'implique pas qu'un processus le constituant existe tout au long de l'événement (surtout si l'événement est qualifié d'« intermittent »).

Mais l'élément essentiel des discussions porte sur la notion même d'*existence*. Nous avons vu que la locution « existe dans sa totalité », lorsqu'elle se rapporte aux objets physiques, signifie « l'objet existe et il est étendu totalement sur une région spatiale ». La question reste donc celle du sens à donner à la propriété (o_i) « l'objet existe à des instants ». Précisons que nous ne considérons pas l'existence comme une propriété physique. Rappelons que nous avons adopté la conception Armstrongienne des propriétés comme des manières d'être d'objets et non l'être même. Inutile donc de chercher du côté des faits pour des vérifacteurs de telles propositions existentielles. Suivant Peter Simons [36], nous considérons qu'un objet dépend existentiellement de processus. Pour un être vivant, la biologie nous renseigne sur la nature de ces processus [36] :

> *For a human being or other animal the relevant processes are those which are vital to it, which are a (probably not exactly delimited) collection of occurrents in its life, involving respiration, blood*

transport, nutrient breakdown and the chemical reactions within the cells.

Pour des objets inertes comme une pierre, ce sont la physique et la chimie qui nous renseignent sur la nature de ces processus [36] :

What keeps the rock in existence? If we look more closely into the physics and chemistry of it we find it much less boring. The widespread cohesion of the crystals in the rock which hold it together as a mechanically unified mass depend on chemical bonds among atoms, and these depend on sharing and exchange of electrons and on the continued existence of the many particles in the rock […] The rock is in fact teeming with occurrents, vast numbers of them, and they are vital to it: if they stopped, it would cease to exist.

Nous retrouvons notre conception d'une dépendance mutuelle entre objets et processus telle qu'exprimée par les propriétés (o_iv) et (p_iv). Notamment, pour les processus, leur existence dépend d'objets qui, à leur tour, dépendent existentiellement de processus, etc. Nous en concluons une proximité d'existence entre processus et objets, suffisante pour classer les processus dans la catégorie des continuants.

3.2 De la notion d'*occurrent*

Pour revenir sur la classe des entités occurrentes, c'est-à-dire des entités dont on dit qu'elles « surviennent », ayant rangé les processus du côté des continuants concrets, nous avons identifié deux autres catégories d'occurrents : les perpétuations de processus (ex : la poussée que j'exerce sur la porte perpétue son ouverture), dont nous avons fait des particuliers concrets, et les événements, positionnés du côté des entités abstraites et classiquement considérés dans la littérature comme des perdurants.

Les faits de perpétuation de processus ne figurent ordinairement pas dans les inventaires d'entités occurrentes. La raison que nous avançons est que la conception prédominante des processus, suivant les travaux de Mourelatos [31], est d'admettre des « macro » processus comme 'ouvrir la porte', 'écrire une lettre', voire même 'construire une maison'. Ces macro-processus se retrouvent dans la doctrine prédominante de la constitution des événements par des processus, conduisant à considérer que des événements comme 'l'ouverture d'une porte' ou 'la construction d'une maison' sont constitués a minima d'UN processus présent toute la durée de l'événement. Notre conception du processus physique met à mal cette doctrine et nous conduit au contraire à identifier plusieurs processus physiques se propageant les uns les autres : une ouverture manuelle d'une porte suppose ainsi l'existence d'un mouvement de la main perpétuant un mouvement de la porte. Sans ces faits de perpétuations, nous serions en peine de rendre compte de ces phénomènes dynamiques.

Côté événements, rompant avec la doctrine de la constitution des événements par les processus, nous leur avons en effet accordé un statut d'entités abstraites, reprenant une proposition faite par Kathleen Gill en conclusion de son [17], seule possibilité selon elle pour préserver les deux catégories « processus » et « événements » [17] :

Just as physical objects apparently form a metaphysical subcategory of objects distinct from, e.g., numbers and spiritual objects, so a more appropriate starting point for developing a metaphysical subcategorization of occurrences would be to distinguish physical occurrences from, e.g., negative, merely possible or perhaps spiritual occurrences.

Comme en témoigne le terme « occurrence spirituelle », le seul fait de ranger les événements du côté des entités abstraites les apparente aussitôt, pour la plupart des métaphysiciens, à des citoyens de « seconde classe ». Nous le voyons ainsi dans le projet de Joseph Mélia [30] de proposer une ontologie reposant uniquement sur des continuants : selon Mélia, si des discours courants portent sur des événements (ex : une fête, un match de football) semblant indiquer une existence, ces derniers n'existent au mieux qu'en tant qu'entités linguistiques mais n'intéressent pas vraiment les métaphysiciens. Ceci explique que l'ontologie des événements reste largement à établir.

Dans cet article, nous avons avancé des pistes en montrant que le perdurantisme ne s'appliquait pas aux événements. La raison essentielle est que les événements sont caractérisés par une méréologie indirecte : s'ils sont étendus dans le cerveau du sujet qui les pense, ce qui compte pour des sujets est qu'ils relatent des vies occupant des régions spatiotemporelles. Du fait qu'ils relèvent de l'histoire de la vie des substances, leur existence (pour un sujet) ne dépend pas de celle des vies en question. Ce sont des continuants psychologiques (voire sociaux) permettant à des sujets d'embrasser des épisodes de vie pouvant survenir, cette survenue dépendant de l'accumulation de faits (on retrouve ici l'intuition du perdurantisme selon laquelle un événement comme un match de football ne peut survenir qu'à la condition qu'un ensemble de faits existent consécutivement).

Cette caractérisation interroge sur le (ou les) rôle(s) cognitif(s) joué(s) par les événements, notamment par rapport aux propositions. Dans [26] nous avons proposé d'identifier notamment les événements à des contenus d'intentions, en leur faisant jouer un rôle dans les mécanismes de spécification et de contrôle d'action. De quoi relativiser leur statut de « citoyen de seconde classe ».

Bibliographie

[1] D.M. Armstrong, *A world of states of Affairs*, Cambridge University Press, 1997.

[2] H. Bergson, *La pensée et le mouvement ; essais et conferences*, Félix Alcan, Paris, 1934.

[3] S. Borgo, M. Franssen, P. Garbacz, Y. Kitamura, R. Mizoguchi & P.E. Vermaas, Technical Artifacts: An integrated perspective, *Applied Ontology*, Vol. 9, N° 3-4, pp. 217-35, 2014.

[4] R. Chisholm, Events and Propositions, *Noûs*, Vol. 4, N° 1, pp. 15-24, 1970.

[5] T. Crowther, The Matter of Events, *The Review of Metaphysics*, Vol. 65, N° 1, pp. 3-39, 2011.

©AfIA 2018

[6] T. Crowther, Processes as Continuants and Process as Stuff, In R Stout (ed.), *Process, Action, and Experience*, Oxford University Press, pp. 58-81, 2018.

[7] D. Davidson, The Individuation of Events. In N. Rescher (ed.), *Essays in Honor of Carl G. Hempel*, Dordrecht: D. Reidel, pp. 216-34, 1969.

[8] D. Davidson, Events as Particulars, *Noûs*, Vol. 4, N° 1, pp. 25-32, 1970.

[9] F. Dretske, *Explaining behavior: Reasons in a world of causes*, Cambridge, MA: MIT Press, 1988.

[10] M. Ferraris, *Manifeste du nouveau réalisme. Hermann*, 2014 : Paris ; traduction française de M. Flusin et A. Robert de l'ouvrage : *Manifesto del nuovo realism*, 2012.

[11] K. Fine, First-Order Modal Theories III – Facts. *Synthèse*, Vol. 53, pp. 43-122, 1982.

[12] K. Fine, In Defense of Three-Dimensionalism. *The Journal of Philosophy*, Vol. 3, N° 12, pp. 699-714, 2006.

[13] A. Galton, On What Goes On: The ontology of processes and events. In R. Ferrario & W. Kuhn (eds.), *Proceedings of the Fourth International Conference on Formal Ontology in Information Systems (FOIS2006)*, pp. 4-11, 2006.

[14] A. Galton, Experience and History: Processes and their Relation to Events, *Journal of Logic and Computation*, Vol. 18, N° 3, pp. 323-40, 2008.

[15] A. Galton, States, Processes and Events, and the Ontology of Causal Relations, In M. Donnelly & G. Guizzardi (eds.), *Proceedings of the 7th international conference on Formal Ontology in Information Systems*, IOS Press, pp. 279-92, 2012.

[16] A. Galton & R. Mizoguchi, The water falls but the waterfall does not fall: New perspectives on objects, proceses and events, *Applied Ontology*, Vol. 4, N° 2, pp. 71-107, 2009.

[17] K. Gill, On the Metaphysical Distinction Between Processes and Events, *Canadian Journal of Philosophy*, Vol. 23, N° 3, pp. 365-84, 1993.

[18] P. Grenon & B. Smith, SNAP and SPAN: Towards dynamic spatial ontology, *Spatial Cognition and Computation*, Vol. 4, N° 1, pp. 69-104, 2004.

[19] P.M.S Hacker, Events, Ontology and Grammar, *Philosophy*, Vol. 57, pp. 477-86, 1982.

[20] P.M.S. Hacker, Events and Object in Space and Time, *Mind*, Vol. 91, pp. 1-19, 1982.

[21] T. Horgan, The Case Against Events, *Philosophical Review*, Vol. 87, N° 1, pp. 28-47, 1978.

[22] J. Hornsby, Actions and activity, *Philosophical Issues*, Vol. 22, N° 1, pp. 233-45, 2012.

[23] A. Iacona, Are there propositions?, *Erkenntnis*, Vol. 58, N° 3, pp. 325-51, 2003.

[24] G. Kassel, A formal ontology of artefacts, *Applied Ontology*, Vol. 5, N° 3-4, pp. 223-46, 2010.

[25] G. Kassel, Processus, événements et couplages temporels et causaux, *Revue d'Intelligence Artificielle*, Vol. 31, N° 6, pp. 649-679, 2017.

[26] G. Kassel, Ontologie de l'action et formes logiques des phrases d'action : de nouvelles perspectives, In T. De Lima et S. Doutre (eds.), *Actes des 12èmes Journées d'Intelligence Artificielle Fondamentale*, Amiens, 13-15 juin, 2018.

[27] E. Margolis & S. Laurence (eds.), *Concepts: Core readings*, MIT Press, 1999.

[28] C. Masolo, Understanding Ontological Levels. In F Lin & U Sattler (eds.), *Proceedings of the 12th International Conference on the Principle of Knowledge Representation and Reasoning (KR 2010)*, pp. 258-268, 2010.

[29] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, A. Oltramari & L. Schneider, The WonderWeb Library of Foundational Ontologies and the DOLCE ontology, *WonderWeb Deliverable D18*, Final Report, vr. 1.0, 2003.

[30] J. Mélia, Continuants and Occurrents, *Proceedings of the Aristotelian Society*, Supplementary Volumes, Vol. 4, pp. 77-92, 2000.

[31] A.P.D. Mourelatos, Events, Processes, and States, *Linguistics and Philosophy*, Vol. 2, N° 3, pp. 415-434, 1978.

[32] J.R. Searle, *The Construction of Social Reality*, NewYork: Free Press, 1995.

[33] J.R. Searle, *Making The Social World: The Structure of Human Civilization*, Oxford University Press, 2010.

[34] J. Seibt, Beyond Endurance and Perdurance: Recurrent Dynamics, In C. Kanzian (ed.), *Persistence*, Frankfurt: Ontos Verlag, pp. 133-164, 2008.

[35] T. Sider, *Four-dimensionalism: An Ontology of Persistence and Time*, Oxford: Oxford University Press, 2001.

[36] P. Simons, Continuants and Occurrents, *Proceedings of the Aristotelian Society*, Supplementary Volumes, Vol. 4, pp. 59-75, 2000.

[37] H. Steward, Actions as Processes, *Philosophical Perspectives*, Vol. 26, N° 1, pp. 373-88, 2012.

[38] H. Steward, Processes, Continuants, and Individuals, *Mind*, Vol. 122, N° 487, pp. 781-812, 2013.

[39] H. Steward, What is a continuant?, In *Proceedings of the Aristotelian Society*, Supplementary Volume: LXXXIX, pp. 109-123, 2015.

[40] R. Stout, Processes, *Philosophy*, Vol. 72, N° 279, pp. 19-27, 1997.

[41] R. Stout, The life of a process, In G. Debrock (ed.), *Process Pragmatism: Essays on a Quiet Philosophical Revolution*, Rodopi, pp. 145-57, 2003.

[42] R. Stout, The category of occurrent continuants, *Mind*, Vol. 125, N° 497, pp. 41-62, 2016.

[43] P. Strawson, *Individuals. An Essay in Descriptive Metaphysics*. Methuen, London, 1959.

[44] A.L. Thomasson, Foundations for a social ontology, *Protosociology*, Vol. 18, pp. 269-90, 2003.

[45] C.A. Varzi, Events, Truth and Indeterminacy, *The Dialogue*, Vol. 2, pp. 241-264, 2002.

[46] N. Wilson, Facts, Events, and Their Conditions, *Philosophical Studies*, Vol. XXV, pp. 303-321, 1974.

Sur la Compilation de Jeux de Prédictions Combinatoires

Frédéric Koriche

CRIL, CNRS UMR 8188, Université d'Artois

frederic.koriche@cril.fr

Résumé

En optimisation en-ligne, le but est de choisir séquentiellement des solutions de manière à minimiser le coût moyen au cours du temps. Dès lors que l'espace de solutions faisables est décrit par des contraintes combinatoires, le problème est généralement NP-difficile. Dans cet article, nous investiguons l'idée de compiler un ensemble de contraintes combinatoires un circuit arithmétique de type dDNNF*, pour lequel les opérations d'optimisation linéaire et d'échantillonnage de solutions prennent un temps linéaire. Dans ce cadre de compilation, nous présentons des algorithmes efficaces pour diverses stratégies de prédiction en-ligne, en portant une attention particulière sur les approches de type « descente miroir ». Ces stratégies sont expérimentalement comparées sur plusieurs benchmarks du monde réel, pour lesquels les contraintes on été préalablement compilées en un circuit* dDNNF*.*

Mots Clef

Apprentissage en-ligne, compilation de connaissances, jeux de prédictions combinatoires.

Abstract

*In online optimization, the goal is to iteratively choose solutions from a decision space, so as to minimize the average cost over time. As long as this decision space is described by combinatorial constraints, the problem is generally intractable. In this paper, we consider the paradigm of compiling the set of combinatorial constraints into a deterministic and Decomposable Negation Normal Form (*dDNNF*) circuit, for which the tasks of linear optimization and solution sampling take linear time. Based on this framework, we provide efficient characterizations of existing combinatorial prediction strategies, with a particular attention to mirror descent techniques. These strategies are compared on several real-world benchmarks for which the set of Boolean constraints is preliminarily compiled into a* dDNNF *circuit.*

Keywords

Online learning, knowledge compilation, combinatorial prediction games.

1 Introduction

L'optimisation combinatoire est un domaine important de l'informatique et des mathématiques discrètes, avec un large spectre d'applications s'étendant depuis l'allocation de ressources et l'ordonnancement, jusqu'à la planification d'actions et les logiciels de configuration. Un des problèmes récurrents de ce domaine est de minimiser une fonction de perte linéaire $\boldsymbol{\ell}$ sur un espace combinatoire $\mathcal{S} \subseteq \{0,1\}^d$ de solutions faisables, représenté de manière compacte par un ensemble de contraintes. Dans la version *hors-ligne* de ce problème, toute l'information nécessaire à la spécification de la tâche d'optimisation est fournie à l'avance, et le but est de construire des algorithmes qui sont meilleurs que l'énumération naïve de toutes les solutions possibles. Dans la version *en-ligne* de ce problème, la fonction objectif $\boldsymbol{\ell}$ est susceptible de changer au cours du temps. Le défi est ici encore plus difficile, puisque l'algorithme d'optimisation doit choisir de manière séquentielle des solutions dans $\mathcal{S}$, afin de minimiser le coût moyen de ses choix.

De manière conceptuelle, le problème d'optimisation combinatoire en-ligne peut être vu comme un jeu répétitif entre un algorithme d'apprentissage et son environnement [1, 2]. Durant chaque tour t du jeu, l'apprenant choisit une solution faisable $\boldsymbol{s}_t$ dans l'espace $\mathcal{S}$ et, simultanément, l'environnement sélectionne un vecteur de perte $\boldsymbol{\ell}_t \in [0,1]^d$. A la fin du tour, l'apprenant subit la perte $\langle \boldsymbol{\ell}_t, \boldsymbol{s}_t \rangle = \sum_{i=1}^d \ell_t(i) s_t(i)$ et, selon le feedback fourni par son environnement, met à jour sa stratégie afin d'améliorer la chance de trouver de meilleures solutions durant les prochains tours.

Diverses classes de jeux de prédictions combinatoires peuvent être déclinées selon le type d'espace de décision et le type de feedback observé. Dans cet article, nous étudions les jeux à *information totale*, dans lesquels il est supposé que le feedback émis par l'environnement au tour t est le vecteur $\boldsymbol{\ell}_t$. En d'autres termes, l'apprenant connaît parfaitement la décision prise par l'environnement à la fin de chaque tour. En revanche, nous supposons que $\mathcal{S}$ peut être décrit par n'importe quelle formule SAT, c'est-à-dire un ensemble arbitraire de contraintes représentées par des clauses booléennes. Comme les codages SAT sont fréquemment utilisés dans les applications académiques et industrielles [5], notre cadre conceptuel couvre une classe importante de jeux de prédictions combinatoires.

©AfIA 2018

En apprentissage en-ligne, la performance d'un algorithme de prédiction est caractérisée par deux mesures. La première est le *regret* : c'est la différence de perte cumulative entre l'algorithme de prédiction et la meilleure solution possible, qui aurait été choisie avec le recul sur l'historique du jeu. Dans cette étude, nous ne faisons aucune hypothèse sur la séquence des vecteurs de perte ; en particulier, $\boldsymbol{\ell}_t$ peut dépendre des décisions antérieures $\boldsymbol{s}_1, \cdots, \boldsymbol{s}_{t-1}$ choisies par l'apprenant. Pour ces environnements "sans oubli", pouvant être vus comme des adversaires, l'apprenant est autorisé à choisir ses décisions de manière aléatoire, et sa performance prédictive est mesurée par le *regret espéré* :

$$R_T = \mathbb{E}\left[\sum_{t=1}^{T} \langle \boldsymbol{\ell}_t, \boldsymbol{s}_t \rangle\right] - \min_{\boldsymbol{s} \in \mathcal{S}} \sum_{t=1}^{T} \langle \boldsymbol{\ell}_t, \boldsymbol{s} \rangle$$

La seconde mesure de performance est la *complexité calculatoire*, c'est-à-dire la quantité de ressources nécessaires pour trouver $\boldsymbol{s}_t$ à chaque tour t, étant donné la séquence de feedbacks observés jusqu'à présent.

1.1 Travaux Relatifs

Dans la littérature des jeux de prédictions combinatoires, trois principales stratégies ont été proposées pour atteindre un regret espéré qui est sous-linéaire en l'horizon de jeu T, et polynomial en la dimension d de l'espace de décision. La première, et sans doute la plus simple des stratégies, s'appelle *Follow the Perturbed Leader* (FPL) : à chaque tour t, l'apprenant choisit aléatoirement un vecteur de perturbation $\boldsymbol{z}_t \in \mathbb{R}^d$, et sélectionne un minimiseur de $\eta \boldsymbol{L}_t + \boldsymbol{z}_t$, où $\eta \in (0, 1]$ est un paramètre d'apprentissage et $\boldsymbol{L}_t$ est la perte cumulative : $\boldsymbol{L}_t = \boldsymbol{\ell}_1 + \cdots + \boldsymbol{\ell}_{t-1}$. Introduit par Hannan [15], et amélioré dans [18, 19], l'algorithme FPL atteint un regret espéré de $\mathcal{O}(d^{\frac{3}{2}}\sqrt{T})$.

La seconde stratégie est fondée sur le prédicteur "exponentially weighted average" bien connu dans le cadre de la prédiction avec avis d'experts [6]. L'idée générale est de maintenir un poids pour chaque solution faisable $\boldsymbol{s} \in \mathcal{S}$, qui décroit exponentiellement avec la perte cumulative de $\boldsymbol{s}$. Spécifiquement, à chaque tour t, l'apprenant choisit une solution $\boldsymbol{s}_t \in \mathcal{S}$ aléaoirement selon la distribution exponentielle $p_t(\boldsymbol{s}) \sim \exp(-\eta \langle \boldsymbol{L}_t, \boldsymbol{s} \rangle)$. Cette stratégie, appelée *Expanded Hedge* (EH) dans [20], atteint un regret espéré de $\mathcal{O}(d^{\frac{3}{2}}\sqrt{T})$, similaire à celui de FPL.

Enfin, la troisième stratégie appelée *Follow the Regularized Leader*, est le paradigme principalement utilisé en optimisation convexe en-ligne [16]. L'apprenant utilise ici l'enveloppe convexe de $\mathcal{S}$, notée $\mathrm{conv}(\mathcal{S})$. A chaque tour t, l'apprenant démarre en choisissant un point $\boldsymbol{p}_t \in \mathrm{conv}(\mathcal{S})$ qui minimize $\eta \langle \hat{\boldsymbol{L}}_t, \boldsymbol{p} \rangle + F(\boldsymbol{p})$, où F est une fonction de régularisation. Ensuite, $\boldsymbol{p}_t$ est décomposé en une combinaison convexe de solutions faisables dans $\mathcal{S}$, à partir de laquelle une décision $\boldsymbol{s}_t$ est choisie aléatoirement. Pour les fonctions de perte linéaire, cette stratégie est équivalent à l'algorithme *Online Stochastic Mirror Descent* (OSMD) [2, 30] qui, à chaque tour de jeu, accomplit une descente de gradient dans l'espace dual de $\mathrm{conv}(\mathcal{S})$ selon F, et projette ensuite cette solution sur l'espace primal, selon la divergence de Bregman définie sur F. En particulier, lorsque F est le régularisateur euclidien, OSMD coïncide avec la descente de gradient stochastique (SGD) [31]. Lorsque F est le régularisateur entropique, OSMD correspond à l'algorithme *Component Hedge* (CH) [20], qui atteint un regret espéré *optimal* de $\mathcal{O}(d\sqrt{T})$.

Si l'on se focalise sur la mesure de regret, l'état de l'art nous indique que peu d'innovations restent à faire dans les jeux à information totale. Cependant, le constat est bien différent dès lors que l'on tient compte des aspects calculatoires : les trois stratégies mentionnées plus haut font intervenir de puissants oracles pour prendre des décisions dans des espaces $\mathcal{S}$ représentés par des contraintes combinatoires. En particulier, la stratégie EH nécessite, à chaque itération, d'échantillonner une solution selon une famille exponentielle sur $\mathcal{S}$, ce qui est un problème #P-difficile [12]. De manière analogue, la stratégie FPL doit répétitivement résoudre un problème d'optimisation linéaire sur $\mathcal{S}$, ce qui est généralement NP-difficile [8]. Pour l'algorithme OSMD et ses spécialisations SGD et CH, la difficulté calculatoire est exacerbée par le fait que, même si l'apprenant a accès à un oracle d'optimisation linéaire, il doit accomplir à chaque tour de jeu une projection de Bregman pour laquelle les meilleurs algorithmes ont une complexité temporelle en $\mathcal{O}(d^6)$ [33] ou $\mathcal{O}(d^4)$ [24].

Bien que les jeux de prédictions combinatoires soient généralement NP-difficiles, il est dans certains cas possible d'obtenir des implémentations efficaces pour les oracles d'échantillonnage et d'optimisation. Par exemple, lorsque l'espace de décision $\mathcal{S}$ coïncide avec les bases d'un matroïde binaire, ou les appariements parfaits d'un graphe biparti, l'optimisation linéaire peut être résolue en temps polynomial, et donc des implémentations efficaces de FPL et OSMD peuvent être dérivées [17, 20, 34, 30]. D'autre part, lorsque les solutions faisables de $\mathcal{S}$ correspondent aux chemins d'un graphe orienté sans circuit (DAG), l'oracle d'échantillonnage peut être simulé par la technique de *weight pushing* [25], qui évalue récursivement la constante de partition d'une famille exponentielle sur les arêtes du DAG. A partir de cette technique, des implémentations efficaces de EH peuvent être dérivées [35, 29].

1.2 Nos Résultats

Envisager les solutions faisables comme des chemins d'un DAG n'est qu'une des nombreuses abstractions possibles qui ont été proposées dans la littérature en complexité de circuits, pour représenter des espaces combinatoires. En compilation de connaissances [11], diverses classes de circuits booléens ont été identifiées, chacune associée avec une collection de tâches d'inférence qui peuvent être accomplies en temps polynomial. Ces résultats théoriques nous incitent naturellement à nous poser la question suivante : est-il possible de *compiler* un ensemble de contraintes représentant un espace combinatoire $\mathcal{S}$ en un circuit booléen compact, pour lequel l'échantillonnage de solution et l'optimisation

linéaire sont traitables ? En regardant le processus de compilation comme une étape de « prétraitement », nous pouvons obtenir des implémentations efficaces des oracles d'échantillonnage et d'optimisation, dès lors que la taille du circuit compilé n'est pas trop grande.

Cet article vise à résoudre des jeux de prédictions combinatoires en compilant les espaces de décision en circuit dDNNF (*deterministic Decomposable Negation Normal Form*) [9]. Pour cette classe de circuits, il existe des compilateurs génériques prenant en entrée une formule SAT représentant un espace de décision $\mathcal{S}$, et retournant en sortie un circuit C qui code $\mathcal{S}$ [10, 23]. Bien que la taille de C puisse croître de manière exponentielle en la "treewidth" de la formule d'entrée, elle est souvent bien plus petite en pratique ; les compilateurs actuels sont capables de compresser des espaces combinatoires définis sur des milliers de variables et de contraintes.

Avec ces outils de compilation en main, nos contributions sont les suivantes : (i) nous montrons que pour les circuits dDNNF, l'oracle d'échantillonnage pour EH, et l'oracle d'optimisation linéaire pour FPL, prennent tous deux un temps linéaire en utilisant de simples variantes de la technique weight-pushing ; (ii) pour les stratégies SGD and CH, nous développons une méthode de projection-décomposition de Bregman qui utilise seulement $\mathcal{O}(d^2 \ln(dT))$ appels à l'oracle d'optimisation linéaire ; (iii) nous montrons expérimentalement sur des tâches de configuration séquentielle et de planification en-ligne que EH et FPL sont rapides, mais nos variants de SGD et CH sont plus efficaces pour minimiser le regret empirique.

2 Inférence Efficace

Pour les jeux de prédictions combinatoires considérés dans ce papier, nous supposons que l'espace de décision $\mathcal{S}$ est défini à partir d'un ensemble fini d'attributs binaires, et nous utilisons $X = \{x_1, \cdots, x_d\}$ pour caractériser l'ensemble de toutes les paires "attribut-valeur", appelées littéraux ou *indicateurs*. Une *solution* est un vecteur $\boldsymbol{s} \in \{0,1\}^d$ tel que $s(i) + s(j) = 1$ pour chque paire d'indicateurs distincts $x_i, x_j \in X$ définis sur le même attribut. Ainsi, $\|\boldsymbol{s}\|_1 = \frac{d}{2}$ pour toute solution $\boldsymbol{s} \in \mathcal{S}$.

Un circuit NNF sur X est un DAG enraciné, dont les nœuds internes sont étiquetés par $\vee$ (or-node) ou $\wedge$ (and-node), et dont les feuilles sont étiquetées par un indicateur dans X, ou bien une constante dans $\{0,1\}$. La taille de C, notée $|C|$, est donnée par le nombre d'arêtes dans C. L'ensemble des attributs apparaissant dans le sous-graphe de C, enraciné au nœud c, est noté $att(c)$.

Pour des raisons de clarté, nous supposerons que tout circuit NNF C satisfait deux propriétés : (i) tout nœud interne c dans C a exactement deux enfants, notés c_l et c_r, et (ii) $att(c_l) = att(c_r) \neq \varnothing$ pour tout or-node c de C. Un circuit NNF satisfaisant ces deux conditions est dit *lisse*. Comme il est montré dans [9], tout circuit booléen C peut être transformé en un circuit NNF lisse équivalent, dont la taille est linéaire en $|C|$.

Q	R	$\oplus$	$\otimes$	$\top$	$\bot$
maxmin	$\{0,1\}$	max	min	1	0
minsum	$\mathbb{R} \cup \{+\infty\}$	min	$+$	0	$+\infty$
sumprod	$\mathbb{R}$	$+$	$*$	1	0

TABLE 1 – Semi-anneaux commutatifs

En voyant les indicateurs comme des "portes d'entrées" et les nœuds comme des "portes de sortie", nous pouvons définir diverses tâches d'inférence sur les circuits booléens, qui dépendent du type d'entrée et de la sémantique des nœuds. Comme le suggèrent Friesen & Domingos pour les fonctions sommes-produits [13], les tâches d'inférence peuvent être caractérisées par des opérations de semi-anneaux. Rappelons ici qu'un *semi-anneaux commutatif* est un tuple $(R, \oplus, \otimes, \bot, \top)$ tel que R est un ensemble contenant les éléments $\bot$ et $\top$, $\oplus$ est un opérateur binaire associatif et commutatif sur R avec l'élément neutre $\bot$, $\otimes$ est un opérateur binaire associatif sur R avec l'élément neutre $\top$ et l'élément absorbant $\bot$, et l'opérateur $\otimes$ se distribue (à droite et à gauche) sur l'opérateur $\oplus$.

Les tâches d'inférence sur un circuit NNF C sont spécifiées par le choix d'un semi-anneau commutatif $Q = (R, \oplus, \otimes, \bot, \top)$ et d'un vecteur d'entrée $\boldsymbol{w} \in R^d$. La *sortie* d'un nœud c dans C pour Q étant donné $\boldsymbol{w}$ est notée $Q(c \,|\, \boldsymbol{w})$, et elle est récursivement donnée par

$$Q(c \,|\, w) = \begin{cases} w(i) & \text{si } c \text{ est un indicateur } x_i, \\ \top & \text{si } c \text{ est la constante } 1, \\ \bot & \text{si } c \text{ est la constante } 0, \\ Q(c_l \,|\, \boldsymbol{w}) \oplus Q(c_r \,|\, \boldsymbol{w}) & \text{si } c \text{ est le nœud } \vee, \\ Q(c_l \,|\, \boldsymbol{w}) \otimes Q(c_r \,|\, \boldsymbol{w}) & \text{si } c \text{ est le nœud } \wedge \end{cases}$$

Nous notons $Q(C \,|\, \boldsymbol{w})$ la sortie de la racine de C pour Q selon $\boldsymbol{w}$. Les semi-anneaux commutatifs jouant un rôle majeur dans cette étude sont décrits dans le tableau 1 ; maxmin, minsum, and sumprod, sont utilisés pour capturer les tâches d'inférence reliées respectivement à l'évaluation de modèle, l'optimisation linéaire, et l'échantillonnage de modèle.

2.1 Évaluation de Modèle

Étant donné un circuit NNF C sur X, l'évaluation de modèle (*model checking*) est de décider si une solution possible $\boldsymbol{s} \in \{0,1\}^d$ est vraie dans C selon la sémantique propositionnelle des nœuds. Il est clair que $\boldsymbol{s}$ est un modèle de C ssi $\mathrm{maxmin}(C \mid \boldsymbol{s}) = 1$, ce qui peut être déterminé en temps $\mathcal{O}(|C|)$. Un circuit NNF C est appelé *représentation* d'un ensemble de solutions faisables $\mathcal{S} \subseteq \{0,1\}^d$ si $sol(C) = \mathcal{S}$, où $sol(C)$ est l'ensemble des modèles de C.

Si l'on excepte l'évaluation de modèle, la totalité des tâches d'inférence dans les circuits NNF sont NP-difficiles. En effet, le langage NNF couvre la classe des formules SAT. Ainsi, nous avons besoin de contraindre cette

©AfIA 2018

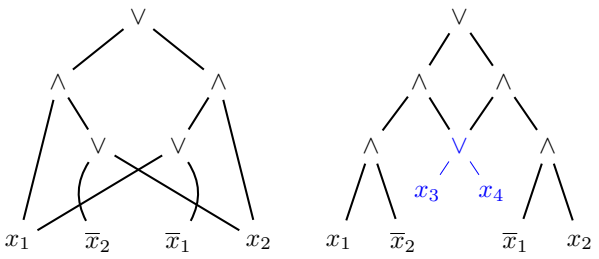


FIGURE 1 – A gauche, un circuit DNNF et à droite un circuit dDNNF.

Algorithm	Regret	Runtime
EH	$\mathcal{O}(d^{\frac{3}{2}}\sqrt{T})$	$\mathcal{O}(\lvert C\rvert)$
FPL	$\mathcal{O}(d^{\frac{3}{2}}\sqrt{T})$	$\mathcal{O}(\lvert C\rvert)$
SGD+PCG	$\mathcal{O}(d(\sqrt{T}+\ln T))$	$\mathcal{O}(d^2\lvert C\rvert\ln T)$
δ-CH+PCG	$\mathcal{O}(d(\sqrt{T}+\ln T))$	$\mathcal{O}\left(\frac{d^2\lvert C\rvert}{\delta}\ln\frac{T}{\delta}\right)$

TABLE 2 – Regret espéré et complexité temporelle (par tour) de chaque stratégie implémentée sur un circuit dDNNF C.

classe pour obtenir des formes efficaces d'optimisation et d'échantillonnage.

2.2 Décomposabilité et Optimisation

Un circuit booléen C est *décomposable* si pour tout and-node c de C, nous avons $att(c_l) \cap att(c_r) = \varnothing$. La classe des circuits NNF décomposables est notée DNNF. Pour de tels circuits, qui sont similaires aux réseaux sommes-produits [28], nous pouvons obtenir une implémentation efficace de l'oracle d'optimisation linéaire.

Proposition 1. Soit $\mathcal{S} \subseteq \{0,1\}^d$ un espace de décision non vide représenté par un circuit DNNF C, et soit $\boldsymbol{w} \in \mathbb{R}^d$ une fonction objectif linéaire. Trouver un minimiseur de $\boldsymbol{w}$ dans $\mathcal{S}$ peut être réalisé en temps $\mathcal{O}(|C|)$.

2.3 Déterminisme et Échantillonnage

Comme le problème de comptage du nombre de modèles d'un circuit DNNF est #P-difficile [11], nous avons besoin de spécialiser cette classe afin d'obtenir une implémentation efficace de l'oracle d'échantillonnage. Dans cette perspective, un circuit NNF C est dit *déterministe* si $\mathrm{minmax}(c_l \mid \boldsymbol{s}) + \mathrm{minmax}(c_r \mid \boldsymbol{s}) \leq 1$ pour tout or-node $c \in C$ et toute solution faisable $\boldsymbol{s}$. La classe des circuits DNNF déterministes est notée dDNNF.

Proposition 2. Soit $\mathcal{S} \subseteq \{0,1\}^d$ un espace de décision représenté par un circuit dDNNF circuit C. Pour un vecteur $\boldsymbol{w} \in \mathbb{R}^d$, soit $\mathbb{P}_{\boldsymbol{w}}$ la famille exponentielle sur $\mathcal{S}$ donnée par :

$$\mathbb{P}_{\boldsymbol{w}}(\boldsymbol{s}) = \frac{\exp\langle \boldsymbol{w}, \boldsymbol{s}\rangle}{\sum_{\boldsymbol{s}' \in \mathcal{S}} \exp\langle \boldsymbol{w}, \boldsymbol{s}'\rangle}$$

L'échantillonnage $\boldsymbol{s} \sim \mathbb{P}_{\boldsymbol{w}}$ peut être fait en temps $\mathcal{O}(|C|)$.

3 Prédiction Efficace

Après une excursion dans les langages de compilation, nous sommes maintenant prêts à examiner des caractérisations efficaces de stratégies de prédictions combinatoires. Nos résultats sont résumés dans le tableau 2. En utilisant le fait que $\|\boldsymbol{s}\|_1 = d/2$, les bornes de regret pour EH and FPL peuvent facilement être dérivées de [1] and [18]. Les deux stratégies sont aussi faciles à implémenter sur des dDNNF. En effet, rappelons que EH tire, à chaque tour t, une solution faisable $\boldsymbol{s}_t \in \mathcal{S}$ aléatoirement selon la distribution $\mathbb{P}_{-\eta \boldsymbol{L}_t}$, où $\boldsymbol{L}_t = \boldsymbol{\ell}_1 + \cdots + \boldsymbol{\ell}_{t-1}$. Ainsi, par application directe de la proposition 2, cette stratégie s'exécute en temps $\mathcal{O}(|C|)$ par tour, dès lors que l'espace $\mathcal{S}$ a été compilé au préalable en un circuit dDNNF C. Pour la stratégie FPL, chaque tour t est accompli en choisissant un minimiseur $\boldsymbol{s}_t \in \mathcal{S}$ de la fonction objectif $\eta \boldsymbol{L}_t - \boldsymbol{z}_t$, où $\boldsymbol{z}_t \in \mathbb{R}^d$ est un vecteur de perturbation dont les composants sont des variables aléatoires exponentielles indépendantes. Par application de la proposition 1, FPL s'exécute en temps $\mathcal{O}(|C|)$ par tour, en utilisant un codage dDNNF C de $\mathcal{S}$, et le fait que $|C|$ est en $\Omega(d)$.

Cependant, la stratégie OSMD, et ses spécialisations SGD et CH, requièrent plus d'attention du fait qu'elles utilisent une coûteuse étape de projection-décomposition à chaque tour de jeu.

3.1 Descente Miroir Stochastique

L'idée générale de l'algorithme Online Mirror Descent (OMD) est de « suivre le leader régularisé »par le biais d'une approche primale-duale [26, 4]. Soit $\mathcal{K}$ un ensemble convexe, et notons $int(\mathcal{K})$ son intérieur. Étant donné une fonction de régularisation F définie sur $\mathcal{K}$, OMD réalise de manière itérative une descente de gradient dans l'intérieur de l'espace dual $\mathcal{K}^*$, et projette le point dual sur l'espace primal $\mathcal{K}$. La connexion entre $\mathcal{K}$ et $\mathcal{K}^*$ est assurée par les gradients ∇F et ∇F^*, où F^* est le conjugué convexe de F défini sur $\mathcal{K}^*$. L'étape de projection est capturée par la divergence de Bregman de F, qui est une fonction de la forme $B_F \colon \mathcal{K} \times \mathrm{int}(\mathcal{K}) \to \mathbb{R}$ donnée par :

$$B_F(\boldsymbol{p}, \boldsymbol{q}) = F(\boldsymbol{p}) - F(\boldsymbol{q}) - \langle \nabla F(\boldsymbol{q}), \boldsymbol{p} - \boldsymbol{q}\rangle$$

Dans le variante stochastique de OMD, introduite par Audibert et. al. [1, 2], et spécifiée dans l'algorithme 1, chaque étape de projection est réalisée sur le sous-ensemble $\mathrm{conv}(\mathcal{S})$ of $\mathcal{K}$, et le point de projection $\boldsymbol{p}_t$ est décomposé en une combinaison convexe de solutions faisables dans $\mathcal{S}$, à partir de laquelle une solution est tirée aléatoirement pour l'étape de prédiction.

Pour les régularisateurs habituels, le gradient $\nabla F(\boldsymbol{p}_t)$ et son dual $\nabla F^*(\boldsymbol{u}_t)$ sont facilement calculables, et nous supposerons ici que le temps dépensé pour leur construction est négligeable par rapport au temps d'exécution de l'oracle

Algorithme 1 OSMD

Entrée : espace $\mathcal{S} \subseteq \{0,1\}^d$, horizon $T \in \mathbb{Z}_+$
Paramètres : régularisateur F sur $\mathcal{K} \supseteq \mathrm{conv}(\mathcal{S})$, scalaire $\eta \in (0,1]$

soit $\boldsymbol{u}_1 = \mathbf{0}$
pour $t = 1$ **to** T **faire**
  soit $\boldsymbol{p}_t \in \mathrm{Argmin}_{\boldsymbol{p} \in \mathrm{conv}(\mathcal{S})} B_F(\boldsymbol{p}, \nabla F^*(\boldsymbol{u}_t))$
  jouer $\boldsymbol{s}_t \sim \boldsymbol{p}_t$ et observer $\boldsymbol{\ell}_t$
  choisir $\boldsymbol{u}_{t+1} = \nabla F(\boldsymbol{p}_t) - \eta \boldsymbol{\ell}_t$
end pour

Algorithme 2 PCG

Entrée : $\mathcal{S} \subseteq \{0,1\}^d$, $f\colon \mathcal{K} \to \mathbb{R}$, $n \in \mathbb{Z}_+$
Paramètres : scalaires $\{\eta_j\}_{j=1}^n$

soit $\boldsymbol{p}_1$ un point arbitraire de $\mathcal{S}$
pour $j = 1$ **to** n **faire**
  soit $\sum_{i=1}^j \alpha_i \boldsymbol{s}_i$ une décomposition convexe de $\boldsymbol{p}_j$
  soit $\boldsymbol{s}_j^+ \in \mathrm{Argmin}_{\boldsymbol{p} \in \mathrm{conv}(\mathcal{S})} \langle \nabla f(\boldsymbol{p}_j), \boldsymbol{p} \rangle$
  soit $\boldsymbol{s}_j^- \in \mathrm{Argmin}_{\boldsymbol{s} \in \{s_1, \cdots, s_j\}} \langle -\nabla f(\boldsymbol{p}_j), \boldsymbol{s} \rangle$
  choisir $\boldsymbol{p}_{j+1} = \boldsymbol{p}_j + \eta_j(\boldsymbol{s}_j^+ - \boldsymbol{s}_j^-)$
end pour

d'optimisation. En fait, le goulot d'étranglement de OSMD est de trouver un minimiseur $\boldsymbol{p}_t$ de $B_F(\boldsymbol{p}, \nabla F^*(\boldsymbol{u}_t))$ dans l'enveloppe convexe de $\mathcal{S}$, et de décomposer $\boldsymbol{p}_t$ en une combinaison convexe de solutions dans $\mathcal{S}$. Heureusement, selon des hypothèses raisonnables sur la courbure de B_F, cette étape de projection-décomposition peut être calculée efficacement, en s'inspirant de résultats récents sur les algorithmes d'optimisation convexe *projection-free*.

De manière plus formelle, nous disons qu'une divergence de Bregman B_F possède le *nombre conditionnel* β/α si, en son premier argument, B_F est α-fortement convexe et β-lisse [1] selon la norme euclidienne $\|\cdot\|_2$. Pour de tels régularisateurs, le résultat suivant établit que l'étape de projection-décomposition peut être approximée efficacement, en exploitant la méthode *Pairwise Conditional Gradient* (PCG), qui est une variante de l'algorithme Frank-Wolfe dont la vitesse de convergence a été analysée dans [22, 14, 3].

Lemme 1. Soit $\mathcal{S} \subseteq \{0,1\}^d$ un ensemble de décision représenté par un circuit dDNNF C, et F un régularisateur sur $\mathcal{K} \supseteq \mathrm{conv}(\mathcal{S})$ tel que B_F possède le nombre conditionnel β/α. Alors, pour tout $\boldsymbol{q} \in \mathrm{int}(\mathcal{K})$ et $\epsilon \in (0,1)$, on peut trouver en temps $\mathcal{O}(\frac{\beta}{\alpha} \ln \frac{\beta d}{\epsilon} d^2 |C|)$ une décomposition convexe de $\boldsymbol{p} \in \mathrm{conv}(\mathcal{S})$ telle que

$$B_F(\boldsymbol{p}, \boldsymbol{q}) - \min_{\boldsymbol{p}' \in \mathrm{conv}(\mathcal{C})} B_F(\boldsymbol{p}', \boldsymbol{q}) \leq \epsilon$$

Notons OSMD+PCG la version modifiée de l'algorithme OSMD qui utilise PCG à chaque tour t, afin d'approximer l'étape de projection-d"composition. En plus du régularisateur F et du paramètre η, OSMD+PCG prend en entrée une séquence $\{\epsilon_t\}_{t=1}^T$ telle que

$$B_F(\boldsymbol{p}_t, \boldsymbol{q}_t) - B_F(\boldsymbol{p}_t^*, \boldsymbol{q}_t) \leq \epsilon_t$$

où $\boldsymbol{p}_t$ est le point retourné par PCG, $\boldsymbol{q}_t = \nabla F_\phi^*(\boldsymbol{u}_t)$, et $\boldsymbol{p}_t^*$ est le minimiseur de $B_F(\boldsymbol{p}, \boldsymbol{q}_t)$ sur $\mathrm{conv}(\mathcal{S})$.

Théorème 1. Supposons que OSMD+PCG prenne en entrée un encodage dDNNF C d'un espace de décision $\mathcal{S} \subseteq \{0,1\}^d$, ainsi qu'un horizon T, et utilise un régularisateur F sur $\mathcal{K} \supseteq \mathrm{conv}(\mathcal{S})$ tel que B_F possède le nombre conditionnel β/α, avec le paramètre d'apprentissage $\eta \in (0,1]$ et une séquence de valeurs de précision $\{\epsilon_t\}_{t=1}^T$ telle que $\epsilon_t = \gamma/t^2$ pour $\gamma > 0$. Alors OSMD+PCG atteint le regret espéré :

$$\begin{aligned} R_T \leq \sqrt{\frac{2\gamma d}{\alpha}}(\ln T + 1) + \frac{1}{\eta} \max_{\boldsymbol{s} \in \mathcal{S}} B_F(\boldsymbol{s}, \boldsymbol{p}_1^*) \\ + \frac{1}{\eta} \sum_{t=1}^T B_{F^*}(\nabla F(\boldsymbol{p}_t^*) - \eta \boldsymbol{\ell}_t, \nabla F(\boldsymbol{p}_t)) \end{aligned} \tag{1}$$

avec une complexité temporelle en $\mathcal{O}\left(\frac{\beta}{\alpha} \ln \frac{\beta d T}{\gamma} d^2 |C|\right)$.

3.2 Descente de Gradient Stochastique

L'algorithme (en-ligne) SGD est dérivé de OSMD en utilisant le régularisateur euclidien $F(\boldsymbol{p}) = \frac{1}{2}\|\boldsymbol{p}\|_2^2$. Dans ce cadre simple, l'espace primal et l'espace dual coïncident avec $\mathbb{R}^d$, et donc, $F^*(\boldsymbol{u}) = \boldsymbol{u}$, $\nabla F(\boldsymbol{p}) = \boldsymbol{p}$, et $\nabla F^*(\boldsymbol{u}) = \boldsymbol{u}$. De plus, B_F a le nombre conditionnel $1/1$, puisque $B_F(\boldsymbol{p}, \boldsymbol{q}) = \frac{1}{2}\|\boldsymbol{p} - \boldsymbol{q}\|_2^2$. Nous notons ici SGD+PCG l'instance de OSMD+PCG définie sur le régularisateur euclidien.

Proposition 3. L'algorithme SGD+PCG atteint un regret espéré de $d(\sqrt{T} + \ln T + 1)$ avec une complexité temporelle par tour de $\mathcal{O}(d^2|C| \ln T)$ en utilisant $\eta = 1/\sqrt{T}$ et $\gamma = d/2$.

3.3 Component Hedge

L'algorithme CH est dérivé de OSMD en utilisant le régularisateur entropique $F(\boldsymbol{p}) = \sum_{i=1}^d p(i)(\ln p(i) - 1)$, pour lequel le conjugué est $F^*(\boldsymbol{u}) = \sum_{i=1}^d \exp u(i)$. Ici, on ne peut pas trouver un nombre conditionnel fini pour la divergence associée $B_F(\boldsymbol{p}, \boldsymbol{q}) = \sum_{i=1}^d p(i) \ln \frac{p(i)}{q(i)} - (p(i) - q(i))$, puisque son gradient n'est pas borné. Cette difficulté peut cependant être contournée en utilisant une astuce suggérée dans [21]. L'idée consiste à remplacer le régularisateur entropique par la fonction $F_\delta(\boldsymbol{p}) = F(\boldsymbol{p} + \boldsymbol{\delta})$, où $\delta \in (0,1)$ et $\boldsymbol{\delta} = (\delta, \cdots, \delta)$. Pour cette fonction l'espace primal est $(-\delta, +\infty)$, et puisque $F_\delta^*(\boldsymbol{u}) = F^*(\boldsymbol{u}) - \langle \boldsymbol{u}, \boldsymbol{\delta} \rangle$, l'espace dual est $\mathbb{R}^d$. Il est facile de montrer que

$$\frac{\partial F_\delta(\boldsymbol{p})}{\partial p(i)} = \ln(p(i) + \delta) \qquad \frac{\partial F_\delta^*(\boldsymbol{u})}{\partial u(i)} = e^{u(i)} - \delta$$

$$B_{F_\delta}(\boldsymbol{p}, \boldsymbol{q}) = B_F(\boldsymbol{p} + \boldsymbol{\delta}, \boldsymbol{q} + \boldsymbol{\delta}) \quad B_{F_\delta}^*(\boldsymbol{u}, \boldsymbol{v}) = B_F^*(\boldsymbol{u}, \boldsymbol{v})$$

1. Ces propriétés de courbure sont habituelles en optimisation convexe [16].

©AfIA 2018

où $B_F^*(\boldsymbol{u},\boldsymbol{v}) = \sum_{i=1}^d e^{v(i)}(e^{v(i)-u(i)} + v(i) - u(i) - 1)$. Enfin, puisque les dérivées partielles premières et secondes de $B_{F_\delta}^*(\boldsymbol{p},\boldsymbol{q})$ à la coordonnée $p(i)$ sont

$$\frac{\partial B_{F_\delta}(\boldsymbol{p},\boldsymbol{q})}{\partial p(i)} = \ln\frac{p(i)+\delta}{q(i)+\delta} \quad \frac{\partial^2 B_{F_\delta}(\boldsymbol{p},\boldsymbol{q})}{\partial^2 p(i)} = \ln\frac{1}{p(i)+\delta}$$

il suit que B_{F_δ} possède le nombre conditionnel $1+\delta/\delta$. En effet, étant donné un point arbitraire $\boldsymbol{q} \in \mathrm{int}(-\delta,+\infty)$, soit $H_{\boldsymbol{q}}(\boldsymbol{p})$ la matrice hessienne de $B_{F_\delta}(\boldsymbol{p},\boldsymbol{q})$ at $\boldsymbol{p} \in \mathrm{conv}(\mathcal{S})$. Alors, pour tout $\boldsymbol{z} \in \mathbb{R}^d$, les entrées de la diagonale de $H_{\boldsymbol{q}}(\boldsymbol{p})$ satisfont

$$\frac{1}{1+\delta} \leq \frac{\partial^2 B_F(\boldsymbol{p},\boldsymbol{q})}{\partial^2 p(i)} z(i)^2 \leq \frac{1}{\delta}$$

en utilisant le fait que $p(i) \in [0,1]$. Ainsi, $\alpha\boldsymbol{I} \preccurlyeq H_{\boldsymbol{q}}(\boldsymbol{p}) \preccurlyeq \beta\boldsymbol{I}$ pour $\alpha = 1/1+\delta$ et $\beta = 1/\delta$. Dans la suite, l'instance de OSMD+PCG qui utilise F_δ comme régularisateur est notée δ-CH+PCG.

Proposition 4. L'algorithme δ-CH+PCG atteint un regret espéré de $d(1+2\delta)(\sqrt{T}+\ln T+1)$ avec une complexité temporelle par tour de $\mathcal{O}\left(d^2|C|/\delta \ln T/\delta\right)$, en utilisant $\eta = 1/\sqrt{T}$ et $\gamma = 2d(1/2+\delta)/(1+\delta)$.

4 Expérimentations

Afin d'évaluer la performance des différentes stratégies examinées dans la section précédente, nous avons considéré 16 instances de la SAT Library [2], décrites dans le tableau 3. Les six premières rangées sont des problèmes de configuration (automobile) et les rangées suivantes sont des tâches de planification. Les quatre premières colonnes du tableau indiquent le nom de l'instance SAT, le nombre d'attributs ($d/2$), le nombre de contraintes ($|$SAT$|$), et le nombre de solutions faisables $|\mathcal{S}|$. Nous avons utilisé le compilateur D4 [3] [23] pour transformer les instances SAT en circuits dDNNF. La taille $|C|$ du circuit compilé est reportée dans la cinquième colonne.

Afin de simuler des jeux de prédictions combinatoires, nous avons utilisé le protocole suivant. Supposons que l'ensemble $X = \{x_1, \cdots, x_d\}$ des indicateurs soit trié par ordre lexicographique, de telle manière que pour tout entier impair i, le couple (x_i, x_{i+1}) code les deux configurations du même attribut. Nous construisons d'abord un vecteur $\boldsymbol{\mu}_0$ de $d/2$ variables de Bernoulli indépendantes. Pour chaque tour t du jeu, $\boldsymbol{\mu}_t$ est fixé à $\boldsymbol{\mu}_{t-1}$ avec probabilité 0.9, où tiré aléatoirement et uniformément sur $[0,1]^{d/2}$ avec probabilité 0.1. Le feedback fourni à l'apprenant est un vecteur de perte $\boldsymbol{\ell}_t \in \{0, 1/d\}^d$ tel que $\ell_t(i) + \ell_t(i+1) = 1/d$, et $\ell_t(i) = 1/d$ avec probabilité $\mu_t(i+1/2)$ pour chaque entier impair i. Ainsi, $\ell_t(i+1) = 1/d$ avec probabilité $1 - \mu_t(i+1/2)$. Bien que ce protocole soit essentiellement stochastique, l'environnement change (en secret) $\boldsymbol{\mu}_t$ avec probabilité 0.1 à chaque tour, afin de déjouer l'apprenant.

2. www.cs.ubc.ca/~hoos/SATLIB/index-ubc.html
3. www.cril.univ-artois.fr/KC/d4.html

Les quatre stratégies de prédiction ont été implémentées en C++ et testées sur un Intel i7-5930K (6 cœurs) avec 32 GiB RAM [4]. Pour les algorithmes FPL et EH, nous avons utilisé les valeurs de η reportées dans [1] et [18]. Concernant SGD+PCG et δ-CH+PCG, nous avons utilisé les valeurs de η et γ déterminées par notre analyse théorique : les valeurs de $\{\eta_t\}$ pour PCG ont été calculées par recherche binaire, comme il est préconisé dans [14], et la valeur de δ a été fixée à $1/\ln d$ afin de garder une complexité temporelle quadratique pour δ-CH+PCG. Enfin, nous avons fixé l'horizon T à 10^3, et le temps de calcul total pour les T tours de jeu à 24 heures.

Dans nos expérimentations, le regret est mesuré par la différence de perte cumulative entre la stratégie étudiée et la meilleure solution faisable, obtenue par recul sur le jeu à l'horizon T. Cette mesure établie en moyenne sur 10 simulations, et divisée par T pour donner un regret empirique moyen. Les résultats correspondants sont donnés dans les quatre dernières colonnes du tableau 3. Le symbole "$-$" indique que l'apprenant n'a pas réussi à accomplir les T tours en une journée. Concernant le regret, SGD+PCG et δ-CH+PCG sont meilleurs que EH et FPL, ce qui confirme nos résultats théoriques. Nous mentionnons au passage que SGD+PCG et δ-CH+PCG sont remarquablement stables. Par contraste, FPL exhibe une grande variance.

Concernant les temps de calcul, EH et FPL sont sans surprise plus rapides que SGD+PCG et δ-CH+PCG. Pour les très grandes instances c140-fc et c163-fw, nous avons constaté que EH et FPL réalisent chaque tour en moins d'une minute, alors que les algorithmes OSMD+PCG requièrent plusieurs minutes par tour, à cause des ressources dépensées pour approximer la projection de Bregman. Cependant, il est important de souligner que la vitesse de convergence de PCG est, en pratique, bien plus rapide que l'analyse théorique en $\tilde{\mathcal{O}}(d^2|C|)$. En fait, SGD+PCG et δ-CH+PCG peuvent traiter pratiquement toutes les instances de problèmes de planification en moins d'une minute par tour. Nous avons observé que, pour la plupart des instances, SGD+PCG est légèrement plus rapide que δ-CH+PCG ; c'est particulièrement le cas pour les grands domaines où les faibles valeurs de δ ont un impact sur le temps de calcul. En résumé, SGD+PCG offre le meilleur compromis entre précision et temps de calcul ; puisque toutes les solutions sont denses ($\|\boldsymbol{s}\|_1 = d/2$), la différence de précision entre SGD+PCG et δ-CH+PCG n'est pas significative.

5 Perspectives

En lumière des résultats obtenus dans cette étude, une perspective naturelle de recherche est d'étendre notre cadre à d'autres classes de jeux de prédictions combinatoires. Notamment, le mode *semi-bandit* semble à portée de main. L'extension de EH en mode-semi-bandit, appelée EXP2 [2], utilise un ajustement des poids pour estimer le vecteur de perte à chaque itération. Par simple application de la proposition 2, cet ajustement peut être calculé en temps linéaire sur des circuits dDNNF. De manière analogue,

4. github.com/frederic-koriche/ccpg.git

Name	$d/2$	$\lvert$SAT$\rvert$	$\lvert\mathcal{S}\rvert$	$\lvert C\rvert$	EH	FPL	SGD+PCG	δ-CH+PCG
c140-fc	1828	4267	$5.74\ 10^{151}$	$6.94\ 10^{5}$	164 ± 25 (22s)	215 ± 48 (21s)	-	-
c163-fw	1815	3580	$2.97\ 10^{140}$	$8.93\ 10^{5}$	98 ± 27 (24s)	112 ± 44 (23s)	-	-
c169-fv	1411	637	$3.22\ 10^{15}$	$7.20\ 10^{2}$	3 ± 1 (<1s)	5 ± 2 (<1s)	1 ± 0.2 (1s)	1 ± 0.2 (1s)
c211-fs	1635	2536	$1.37\ 10^{67}$	$1.75\ 10^{3}$	12 ± 3 (<1s)	12 ± 5 (<1s)	9 ± 2 (1s)	8 ± 2 (2s)
c250-fv	1465	1050	$1.20\ 10^{37}$	$1.82\ 10^{2}$	11 ± 3 (<1s)	16 ± 4 (<1s)	9 ± 2 (1s)	9 ± 1 (1s)
c638-fvk	1761	1893	$8.83\ 10^{121}$	$5.45\ 10^{3}$	104 ± 15 (<1s)	127 ± 28 (<1s)	68 ± 4 (2s)	62 ± 4 (4s)
4-step	165	396	$8.34\ 10^{4}$	$1.29\ 10^{2}$	3 ± 2 (<1s)	5 ± 3 (<1s)	1 ± 0.8 (<1s)	1 ± 0.8 (<1s)
5-step	177	459	$8.13\ 10^{4}$	$5.80\ 10^{1}$	3 ± 1 (<1s)	3 ± 1 (<1s)	1 ± 0.9 (<1s)	1 ± 0.8 (<1s)
log-1	939	3742	$5.64\ 10^{20}$	$9.43\ 10^{2}$	46 ± 5 (<1s)	51 ± 8 (<1s)	7 ± 3 (<1s)	7 ± 1 (<1s)
log-2	1337	24735	$3.23\ 10^{10}$	$1.16\ 10^{4}$	16 ± 3 (1s)	19 ± 6 (1s)	12 ± 3 (16s)	11 ± 3 (22s)
log-3	1413	29445	$2.79\ 10^{11}$	$4.96\ 10^{3}$	19 ± 4 (<1s)	21 ± 7 (<1s)	15 ± 4 (5s)	13 ± 4 (5s)
log-4	2303	20911	$2.34\ 10^{28}$	$9.47\ 10^{4}$	77 ± 11 (2s)	94 ± 27 (2s)	19 ± 4 (72s)	17 ± 4 (81s)
tire-1	352	1022	$8.29\ 10^{36}$	$1.37\ 10^{3}$	3 ± 1 (<1s)	4 ± 2 (<1s)	1 ± 0.8 (<1s)	1 ± 0.8 (< 1s)
tire-2	550	1980	$7.39\ 10^{11}$	$7.26\ 10^{2}$	9 ± 2 (<1s)	12 ± 4 (<1s)	7 ± 2 (<1s)	5 ± 1 (<1s)
tire-3	577	1984	$2.23\ 10^{11}$	$6.31\ 10^{3}$	9 ± 4 (<1s)	14 ± 7 (<1s)	7 ± 2 (1s)	5 ± 2 (2s)
tire-4	812	3197	$1.03\ 10^{14}$	$3.97\ 10^{3}$	15 ± 3 (<1s)	17 ± 5 (<1s)	8 ± 3 (<1s)	7 ± 3 (1s)

TABLE 3 – Résultats expérimentaux pour les différentes stratégies sur des instances SAT codées par des circuits dDNNF.

l'extension semi-bandit de FPL exploite une méthode d'échantillonnage géométrique pour estimer les vecteurs de perte [27]. À nouveau, cette méthode itérative peut être implémentée en temps linéaire (par itération) en utilisant la proposition 1. L'adaptation de OSMD aux semi-bandits est cependant moins immédiate : même si l'extension de CH atteint encore un regret espéré optimal dans ce contexte, son utilisation pratique reste limitée à cause de la projection de Bregman. Une question ouverte intéressante ici est de trouver si une combinaison de CH avec PCG est capable, en mode semi-bandit, d'atteindre un regret quasi-optimal en temps quadratique. Naturellement, le mode *bandit* pur est encore plus difficile. Sur ce point, Sakaue et. al. [32] ont obtenu de premiers résultats en utilisant des OBDDs pour une implémentation efficace de l'algorithme COMBBAND [7]. Mis à part ce résultat très récent, tout reste à faire dans les bandits combinatoires, notamment avec des circuits dDNNF, qui sont plus succincts que les OBDDs.

Références

[1] J-Y. Audibert, S. Bubeck, and G. Lugosi. Minimax policies for combinatorial prediction games. In *Proceedings of the 24th Annual Conference on Learning Theory (COLT 2011)*, pages 107–132, 2011.

[2] J-Y. Audibert, S. Bubeck, and G. Lugosi. Regret in online combinatorial optimization. *Mathematics of Operations Research*, 39(1) :31–45, 2014.

[3] M. A. Bashiri and X. Zhang. Decomposition-invariant conditional gradient for general polytopes with line search. In *Advances in Neural Information Processing Systems 30, (NIPS 2017)*, pages 2687–2697, 2017.

[4] A. Beck and M. Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operational Research Letters*, 31(3) :167–175, 2003.

[5] A. Biere, M. Heule, H. van Maaren, and T. Walsh. *Handbook of Satisfiability*. IOS Press, 2009.

[6] N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

[7] N. Cesa-Bianchi and G. Lugosi. Combinatorial bandits. *Journal of Computer and System Sciences*, 78(5) :1404–1422, 2012.

[8] N. Creignou, S. Khanna, and M. Sudan. *Complexity Classification of Boolean Constraint Satisfaction Problems*. SIAM Monographs on Discrete Mathematics and Applications, 2001.

[9] A. Darwiche. Decomposable negation normal form. *Journal of the ACM*, 48(4) :608–647, 2001.

[10] A. Darwiche. A compiler for deterministic, decomposable negation normal form. In *Proceedings of the 18th National Conference on Artificial Intelligence, (AAAI 2002)*, pages 627–634, 2002.

[11] A. Darwiche and P. Marquis. A knowledge compilation map. *Journal of Artificial Intelligence Research (JAIR)*, 17 :229–264, 2002.

[12] M. E. Dyer, L. A. Goldberg, and M. Jerrum. The complexity of weighted Boolean #CSP. *SIAM Journal of Computing*, 38(5) :1970–1986, 2009.

[13] A. L. Friesen and P. M. Domingos. The sum-product theorem : A foundation for learning tractable models. In *Proceedings of the 33nd International Conference on Machine Learning, (ICML 2016)*, pages 1909–1918, 2016.

[14] D. Garber and O. Meshi. Linear-memory and decomposition-invariant linearly convergent conditional gradient algorithm for structured polytopes. In *Advances in Neural Information Processing Systems 29, (NIPS 2016)*, pages 1001–1009, 2016.

[15] J. Hannan. Approximation to Bayes risk in repeated play. *Contributions to the Theory of Games*, 3 :97–139, 1957.

[16] E. Hazan. Introduction to online convex optimization. *Foundations and Trends in Optimization*, 2(3-4) :157–325, 2016.

©AfIA 2018

[17] D. P. Helmbold and M. K. Warmuth. Learning permutations with exponential weights. *Journal of Machine Learning Research*, 10 :1705–1736, 2009.

[18] M. Hutter and J. Poland. Adaptive online prediction by following the perturbed leader. *Journal of Machine Learning Research*, 6 :639–660, 2005.

[19] A. T. Kalai and S. Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3) :291–307, 2005.

[20] W. M. Koolen, M. K. Warmuth, and J. Kivinen. Hedging structured concepts. In *Proceedings of the 23rd Conference on Learning Theory (COLT 2010)*, pages 93–105, 2010.

[21] W. Krichene, S. Krichene, and A. M. Bayen. Efficient bregman projections onto the simplex. In *Proceedings of the 54th IEEE Conference on Decision and Control, (CDC 2015)*, pages 3291–3298, 2015.

[22] S. Lacoste-Julien and M. Jaggi. On the global linear convergence of frank-wolfe optimization variants. In *Advances in Neural Information Processing Systems 28 (NIPS 2015)*, pages 496–504, 2015.

[23] J-M. Lagniez and P. Marquis. An improved decision-dnnf compiler. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence, (IJCAI 2017)*, pages 667–673, 2017.

[24] Y. T. Lee, A. Sidford, and S. C. Wong. A faster cutting plane method and its implications for combinatorial and convex optimization. In *IEEE 56th Annual Symposium on Foundations of Computer Science, (FOCS 2015)*, pages 1049–1065, 2015.

[25] M. Mohri. General algebraic frameworks and algorithms for shortest-distance problems. Technical Report 981210-10TM, AT & T Labs-Research, 1998.

[26] A. S Nemirovski and D. B. Yudin. *Problem Complexity and Method Efficiency in Optimization*. J. Wiley and Sons, 1983.

[27] G. Neu and G. Bartók. Importance weighting without importance weights : An efficient algorithm for combinatorial semi-bandits. *Journal of Machine Learning Research*, 17 :154 :1–154 :21, 2016.

[28] H. Poon and P. M. Domingos. Sum-product networks : A new deep architecture. In *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence (UAI 2011)*, pages 337–346, 2011.

[29] H. Rahmanian and M. K. Warmuth. Online dynamic programming. In *Advances in Neural Information Processing Systems 30, (NIPS 2017)*, pages 2824–2834, 2017.

[30] A. Rajkumar and S. Agarwal. Online decision-making in general combinatorial spaces. In *Advances in Neural Information Processing Systems 27, (NIPS 2014)*, pages 3482–3490, 2014.

[31] H. Robbins and S. Monro. A stochastic approximation method. *The Annals of Mathematical Statistics*, 22(3) :400–407, 1951.

[32] S. Sakaue, M. Ishihata, and S-I Minato. Efficient bandit combinatorial optimization algorithm with zero-suppressed binary decision diagrams. In *Proceedings of the 21st International Conference on Artificial Intelligence and Statistics*, 2018.

[33] D. Suehiro, K. Hatano, S. Kijima, E. Takimoto, and K. Nagano. Online prediction under submodular constraints. In *Proceedings of the 23rd International Conference on Algorithmic Learning Theory (ALT 2012)*, pages 260–274, 2012.

[34] E. Takimoto and K. Hatano. Efficient algorithms for combinatorial online prediction. In *Proceedings of the- 24th International Conference on Algorithmic Learning Theory (ALT 2013)*, pages 22–32, 2013.

[35] E. Takimoto and M. K. Warmuth. Path kernels and multiplicative updates. *Journal of Machine Learning Research*, 4 :773–818, 2003.

Utilisation des systèmes d'interaction vocale dans l'habitat : quels retours d'expérience utilisateur et quels critères d'interaction à intégrer en conception ?

D. Lahoual[1] M. Fréjus[1]

[1]EDF Recherche & Développement – 7 Boulevard Gaspard Monge 91120 Palaiseau

dounia.lahoual@gmail.com
myriam.frejus@edf.fr

Résumé
L'essor de la commande vocale et des assistants vocaux intelligents transforme progressivement les utilisations, les modes d'interaction, ainsi que les expériences utilisateurs. Combinés à de l'intelligence artificielle, ces systèmes d'interaction vocale proposent un ensemble de services autour du confort, de la sécurité, du divertissement et de la santé en intégrant désormais la sphère domestique. Dans une visée de conception d'interaction vocale efficiente et adaptée aux besoins et aux activités des habitants, nous proposons d'établir une cartographie des retours d'expérience utilisateurs à partir d'une revue de la littérature.

Mots-clés
Interaction vocale, expérience utilisateur, habitat, utilisabilité, appropriabilité, acceptabilité, assistants vocaux, usages.

Abstract
The rise of the voice commands and smart speakers are progressively changing uses, ways of interacting and user experiences. Combined with artificial intelligence, these voice interaction systems are offering a set of services relating to comfort, security, entertainment and health in the domestic sphere. In order to design an efficient voice interaction suited for users'needs and activities, we propose a literature review about users experience with voice commands and smart speakers.

Keywords
Vocal interaction, user experience, housing, usability, appropriability, acceptability, smart speaker, uses.

1. Introduction
Afin de s'affranchir des dispositifs matériels contraignants et simplifier les usages, l'intelligence artificielle propose de nouveaux outils et de nouvelles formes d'interaction visant une fluidité et une instantanéité à l'aide du langage naturel [27, 18]. Grâce à ses grandes capacités d'apprentissage automatique, à la programmation de nombreux algorithmes et l'utilisation de données massives, l'IA vise à augmenter les capacités humaines en s'appuyant sur une coopération entre l'homme et la machine [16, 27]. Ainsi, ces dernières années ont été marquées par de nombreux progrès dans le domaine de l'IA, et cela dans plusieurs secteurs : robotique, véhicules autonomes, traitement de la parole et compréhension du langage naturel, maisons intelligentes, etc. [25, 9, 20]. D'abord limités à l'utilisation du smartphone, les assistants vocaux intelligents intègrent les habitats avec l'arrivée des enceintes intelligentes telles qu'Alexa d'Amazon ou Google Home [41]. Un assistant vocal intelligent représente un logiciel intégré dans un dispositif tangible, disposant de la faculté de dialoguer avec un humain pour réaliser des tâches sur ordre vocal des utilisateurs (répondre à des questions en sélectionnant des informations sur le web, rappeler des actions importantes aux utilisateurs, effectuer certaines tâches sur des objets connectés, etc.). Disposant d'une capacité d'apprentissage automatique, ces intelligences artificielles se développent continuellement sans l'intervention des développeurs, apprennent de leurs erreurs, et cela, tout en intégrant des informations sur les utilisateurs afin de proposer des services contextualisés et personnalisés [29, 14]. Si la question de la bonne intégration de la reconnaissance de la parole dans les interfaces et produits grand public n'est pas nouvelle [44], la maturité technologique des assistants actuels questionne l'utilisation réelle qui peut en être fait tant par ses utilisateurs finaux que par les concepteurs, en lien avec les autres formes d'interaction. Nous considérons que pour développer ces nouveaux systèmes vocaux ou les intégrer à des produits ou services, il importe d'étudier et de comprendre les situations d'utilisation dans lesquelles ils s'insèrent, ainsi que les leviers et les freins à leur appropriation. Proposer des systèmes adaptés à l'activité des utilisateurs en prenant en compte le point de vue de l'humain et de ses activités constitue un des principes fondateurs de la démarche en ergonomie que nous proposons de mobiliser.

Intégrées dans un contexte de dispositifs ambiants, la commande et la reconnaissance vocale constituent ainsi de nouveaux défis en apportant de nouvelles situations et formes d'utilisation. De plus, face à la multiplication récente des services digitaux et des systèmes d'interaction vocale, l'ensemble des attitudes, des besoins, et des utilisations provenant des utilisateurs demeurent encore méconnus [51, 53]. Ces problématiques vocales questionnent ainsi en profondeur les principes d'utilisabilité, d'appropriabilité et d'acceptabilité classiques pour diverses raisons. En premier lieu, ces dispositifs vocaux créent une rupture dans les représentations, les utilisations et les routines des usagers. Ces systèmes diffèrent grandement des interfaces tangibles et traditionnelles. En second lieu, l'interface vocale privilégie l'audition et l'usage de la parole : ces capacités habituellement mobilisées en situation d'interaction humaine sont dorénavant au cœur des interactions entre l'humain et la machine. Ces

 ©AfIA 2018

interactions se trouvent aussi modifiées en ce que le système se trouve doté d'une autonomie et d'une capacité d'adaptation qui rompt avec l'interaction homme machine connue. En dernier lieu, ce qui prévaut comme critères de conception pour les IHM traditionnelles n'est pas assuré de fonctionner avec des interfaces vocales.

Afin de développer des connaissances sur ces nouvelles formes d'utilisation vocale, nous proposons dans cet article de réaliser un état des lieux sur les systèmes d'interaction vocale en situation domestique. Cette première étape de veille constitue un préalable dans notre projet de recherche orienté vers la conception de services et de dispositifs vocaux dans l'habitat. Ces éléments récoltés nous permettront de répondre à un ensemble de questions :

- *Quels sont les expériences d'utilisation réussies et empêchées avec ces technologies vocales ?*
- *Comment améliorer et optimiser les interactions entre les systèmes d'interaction vocale et les utilisateurs à partir de la connaissance du contexte et de l'activité de ces derniers et en lien avec les autres formes d'interaction présentes ?*

Tout d'abord, nous présenterons un ensemble de critères d'interaction entre l'humain et les systèmes d'IA au cœur des démarches de conception. Ensuite, nous établirons un état des lieux sur les systèmes d'interaction vocale en nous focalisant sur les enjeux de la parole, les utilisations appréciées et les difficultés rencontrées par les utilisateurs. Puis, nous proposerons des principes d'interaction à intégrer lors de la conception. Enfin, nous conclurons et discuterons des résultats en précisant nos perspectives de recherche.

2. Repères théoriques sur les critères d'interaction Homme-IA

En mobilisant une démarche ergonomique, nous nous intéressons aux multiples dimensions de l'activité, à savoir ; ses aspects situés, finalisés, systémiques, diachroniques, instrumentés, individuels et collectifs. L'identification et la compréhension de ces dimensions contribuent à développer des connaissances sur l'activité réelle des utilisateurs. À terme, il importe de les intégrer dans la conception des dispositifs à base d'IA pour garantir l'utilité, l'utilisabilité, l'appropriabilité et l'acceptabilité de ces systèmes par les utilisateurs.

2.1 Une interaction efficace et opérante sur le temps : utilité, utilisabilité, appropriabilité et acceptabilité

Pour atteindre une situation durable intégrant des systèmes d'IA, il demeure indispensable de se pencher vers les critères d'utilité, d'utilisabilité, d'acceptabilité et d'appropriabilité qui sont au cœur de la démarche de conception et qui conditionnent l'adoption des systèmes. L'**utilité** renvoie à l'adéquation du dispositif et de ses services aux besoins et aux attentes des utilisateurs [50]. Ce critère tend le plus souvent à s'articuler à celui de l'**utilisabilité** qui vise à « *atteindre des buts définis avec efficacité, efficience et satisfaction, dans un contexte d'utilisation spécifié* » d'après la norme ISO 9241-11 [26]. En d'autres termes, il s'agit d'assurer une simplicité d'utilisation pour des situations et des utilisateurs spécifiques. Quand bien même ces critères seraient intégrés dès la conception, le dispositif ne serait toujours pas assuré d'être accepté, ni approprié [13] car l'activité incorpore un aspect dynamique et longitudinal imprévisible qu'il importe d'investiguer en situation réelle. Cette caractéristique est davantage mise en lumière avec la notion d'**appropriation**, et sa déclinaison en conception, l'appropriabilité [23]. L'appropriation représente l'intégration, partielle ou totale, d'un objet, d'un outil ou d'un dispositif au corps propre ou la culture propre de l'acteur. L'appropriation est souvent accompagnée d'une individuation de son usage et éventuellement, de transformations plus ou moins importantes de cet objet, de cet outil ou de ce dispositif lui-même [52]. L'appropriation s'inscrit notamment sur une perspective temporelle longue intégrant les histoires d'utilisations, les représentations des utilisateurs, et leurs nombreux épisodes de succès et d'échecs colorant ce parcours singulier [40, 45]. Par ailleurs, la dynamique de cette trajectoire d'utilisation implique une confrontation itérative de l'utilisateur avec le dispositif dans laquelle les univers de l'utilisateur et du concepteur se rencontrent [5, 40] et peuvent être susceptibles non seulement de susciter des besoins inédits, mais aussi de provoquer l'évolution des schèmes d'usage et la transformation du dispositif par l'utilisateur [30, 46]. Cette dimension de l'appropriabilité renvoie ici à la capacité d'un dispositif à s'adapter et à réagir de façon pertinente selon le contexte, les préférences et le développement des utilisateurs sur le long terme [13]. Et c'est pourquoi les nouvelles perspectives d'autonomie des systèmes à base d'IA re-questionnent l'appropriabilité. Quant au critère d'**acceptabilité**, bien que déterminant pour comprendre les problématiques d'adoption ou non d'un dispositif, ce critère reste encore vivement débattu auprès des chercheurs, notamment dans ses liens avec l'appropriabilité. La technologie doit être perçue en adéquation avec les pratiques, les ressources, les objectifs, le système de valeurs et les normes éthiques des utilisateurs pour être acceptable [13, 36].

2.2 Une déclinaison des critères de conception pour favoriser l'appropriation

Dans un contexte d'innovation technologique en lien avec l'émergence de l'IA, nous avons pu identifier un ensemble de dimensions déterminantes à intégrer dès la conception afin de garantir une meilleure appropriation des dispositifs. Résultant des critères de conception précédemment évoqués, ces dimensions visent l'augmentation des capacités humaines à partir d'une coopération entre l'homme et la machine pour une interaction efficace et opérante en situation.

Parmi ces dimensions, la **complémentarité** évoque tout d'abord la possibilité pour l'utilisateur de s'appuyer sur des dispositifs complémentaires, fonctionnellement, pour pallier aux limites de chaque système, mais aussi en termes de modalités d'interactions. L'interaction

multi-modale permet à l'utilisateur de privilégier une interaction vocale, tactile, télécommandée ou via le clavier selon ses préférences, ses objectifs et les activités en cours [15].

La multi-modalité permet aussi de répondre à un autre objectif : l'**adaptabilité** du système à l'activité des utilisateurs. Celle-ci s'appuie sur une compréhension des contextes pour s'adapter aux situations et agir efficacement, et cela sans perturber le flux de l'activité des utilisateurs : cela peut être illustré par l'allumage des lumières provoqué par des capteurs de mouvement lors de l'entrée d'un usager. Ce système peut notamment être orienté vers l'anticipation des besoins par exemple, en déclenchant le chauffage de certaines pièces avant l'arrivée d'un habitant (op. cit.). Il s'agit donc d'un système « *flexible* » pouvant réagir de façon implicite aux besoins et aux préférences des utilisateurs, tout en prenant en compte leurs expériences [3].

Ces différentes formes d'interaction couplées à leur bonne intégration aux activités en cours doivent permettre de répondre à une recherche **d'invisibilité**, preuve de leur appropriabilité. Un dispositif devient transparent dans l'usage grâce à un processus « *d'in-corporation* ». Ce dernier conduit à intégrer un objet au « *corps propre* » de l'utilisateur comme un instrument pour penser, percevoir et agir. Cette in-corporation peut être illustrée par un conducteur qui fait corps avec sa voiture, de telle façon qu'il pourra percevoir les aspérités de la chaussée et adapter sa conduite [52].

Le développement de systèmes intelligents doit aussi intégrer des critères favorisant le dialogue et la coopération afin de garantir à l'utilisateur un niveau de contrôle, de reprise en main et de visibilité sur les actions du système. Pour ces raisons, l'**intelligibilité** du dispositif [4] demeure cruciale pour informer l'utilisateur des situations en cours, rendre visible la logique de fonctionnement et les capacités du système, ainsi que les conséquences des actions passées et à venir des usagers [16].

Enfin, le critère d'**évolutivité** permet d'intégrer les dimensions temporelles et développementales inhérentes à l'homme. Ce critère « *renvoie à la possibilité du système de proposer des fonctionnalités en contexte et renouvelées en réponse à l'évolution des besoins et capacités* » [17], tout en s'adaptant au contexte d'utilisation des usagers [10]. Cette évolutivité du système peut consister par exemple à proposer des fonctionnalités selon le développement de l'utilisateur ; à prendre en compte les systèmes de ressources développées par les utilisateurs au fil de ses expériences (technologiques, humaines, organisationnelles) ; ou encore, l'évolutivité peut intégrer l'évolution des routines individuelles et collectives sur diverses temporalités et espaces [17].

2.3 L'interaction vocale : quels enjeux ?

La parole représente une modalité fondamentale dans la communication humaine. Dans toutes les cultures, les individus persuadent, informent et construisent des relations à travers la parole principalement. Ainsi, la sensibilité à la parole et à la voix joue ainsi un rôle critique dans les interactions au sein du collectif. De plus, les humains ont développé les instruments de parole et d'écoute exprimant une large variété de signaux sociaux appropriés auxquels le cerveau humain est apte à analyser et à répondre. Parmi les paramètres de la parole, le rythme, le volume, la cadence transmettent des caractéristiques humaines subtiles telles que la personnalité, l'émotion, ou encore la ville d'origine d'une personne. Les humains sont si accoutumés aux caractéristiques vocales qu'ils peuvent rapidement et précisément distinguer une voix d'une autre [34]. Ainsi, dans le contexte de l'affective computing, la voix devient un facteur déterminant car il constitue un indicateur fiable pour détecter l'émotion des utilisateurs, tout comme le choix des mots peut également signaler une humeur particulière [22].

Pour ces raisons, la parole demeure une forme d'interaction intuitive, invisible et rapide pour réaliser une diversité de tâches à partir des commandes vocales. Non seulement les interfaces vocales libèrent les utilisateurs des contraintes « *WIMP* » (Fenêtre, icônes, menus et curseur), mais elles visent également à accompagner ceux et celles dont les yeux et les mains sont affairés à d'autres tâches [34, 22]. Elles permettent aussi de surmonter certaines contraintes individuelles et situationnelles, telles que les personnes portant des handicaps et des déficiences visuelles ou encore les enfants, qui n'ont pas la dextérité ou une taille de main suffisante pour les claviers et les pointeurs. Et tous les environnements ne sont pas tous disposés à proposer des commandes via une manipulation directe [19, 22].

Afin d'encourager une coopération et une meilleure acceptabilité de ces formes d'interaction, certains travaux démontrent l'importance de concevoir une interaction qui puisse s'accorder de façon appropriée à l'état et au timbre émotionnel de l'utilisateur afin de communiquer de l'empathie, du soutien et de la sincérité dans l'échange [34]. Dans cette visée, ces auteurs invitent notamment à créer un sentiment de similarité entre l'utilisateur et l'interface à travers l'adéquation entre le genre, la personnalité et l'accent qui encourageraient des sentiments de sympathie et de confiance pour assurer le succès de ces technologies. Ces recommandations pour la conception s'appuient sur un résultat majeur sur la façon dont l'humain interagit vocalement avec une machine : le cerveau humain ferait rarement la distinction entre le fait de parler à une machine (même si elle possède une faible compréhension de la parole et une qualité médiocre de production langagière) et parler à une personne. Ainsi, les humains mobiliseraient les mêmes parties du cerveau pour interagir avec les machines et les humains. Les machines ne seraient donc pas exemptées des attentes sociales activées lors des échanges et de l'écoute (op. cit.).

Ces éléments soulignent ainsi l'essor d'une forme d'interaction qui pourrait s'intégrer de façon harmonieuse dans l'activité réelle des utilisateurs. De plus, les caractéristiques évolutives et adaptatives de l'IA sous-tendant ces systèmes laissent présager une meilleure adéquation pour leur appropriabilité par un

©AfIA 2018

meilleur respect des dimensions que nous avons évoquées. Toutefois, la combinaison souhaitable avec d'autres modalités ainsi que la question de l'intelligibilité des systèmes à base d'IA, souvent présentés comme une boite noire, restent ouvertes. Ces questions et premières suppositions devront être mises en perspective avec l'activité réelle et les points de vue des utilisateurs.

3. État de l'art sur l'activité réelle des systèmes d'interaction vocale

Nous présenterons dans les prochaines sections un état des lieux sur les systèmes d'interaction vocale. Pour cela, nous nous appuierons sur une diversité de sources telles que les travaux scientifiques, les enquêtes menées par des instituts spécialisés, ainsi qu'une diversité d'articles et de billets récents sur les pages du web. Parmi nos références bibliographiques, notre webographie constitue une source importante compte tenu du développement des technologies vocales et de leur intégration récente dans les foyers. Afin de garantir une validité des données récoltées dans cette revue de l'état de l'art sur les interactions vocales, nous nous appuyons sur la triangulation des données visant une multiplication et une combinaison des matériaux récoltés. Cette triangulation nous permettra notamment d'obtenir une connaissance plus complète et différenciée du même phénomène étudié [15, 42].

3.1 Les technologies vocales : des expériences d'utilisation réussies et appréciées

Parmi les utilisations des technologies vocales en situation domestique et professionnelle, nous retrouvons des expériences réussies et appréciées par les utilisateurs que nous avons catégorisés en trois catégories principales : les qualités des formes d'interaction vocale, les qualités du système vocal et les effets des interactions vocales sur l'activité des utilisateurs.

3.1.1 Les qualités de l'interaction : accélérateur opératoire, multi-modalité, adaptabilité intra-individuelle et courtoisie

Parmi les qualités retrouvées dans les interactions vocales, nous retrouvons tout d'abord **l'instantanéité, la facilité et le confort d'utilisation** de mobiliser la parole comme un moyen d'interaction. Il intègre une rapidité de commande et de contrôle dépassant largement les actions consistant à presser un bouton ou cliquer à l'aide d'une souris « *Alexa m'a gâtée par le confort qu'elle propose, mais elle me fait réellement gagner du temps* » [21]. Les retours vocaux sont notamment rapidement compréhensibles, ils évitent toute ambiguïté que nous pourrions rencontrer via les alarmes sonores et les voyants lumineux [51]. De plus, la commande vocale semble avantageuse pour réaliser des actions complexes nécessitant plusieurs touches et de multiples interfaces. Une étude réalisée par Turunen et al. [53] s'est appuyée sur une évaluation multimodale d'un gestionnaire de dispositifs domestiques (un centre multimédia) auprès de 26 participants. Les résultats ont démontré que la commande vocale avait été très appréciée par les usagers et qu'elle apportait notamment des bénéfices supplémentaires pour réaliser des commandes complexes nécessitant plusieurs touches ou l'utilisation de multiples interfaces (exemple d'une commande dite complexe sans le vocal : « *Enregistrer tous les épisodes de Friends* »).

En définitive, le vocal constitue une façon naturelle, fluide, intuitive et aisée d'interagir avec son environnement : « *Si je travaille tard, en montant les escaliers, je demande : Alexa, allume la chambre à coucher 1. Et quand j'arrive dans la pièce, la lumière est allumée. C'est une expérience plus naturelle que d'utiliser des applications* » [48]. L'interaction vocale agit ainsi en tant qu'accélérateur opératoire en simplifiant ou en supprimant certains modes opératoires [6].

Nous apprenons notamment que la **complémentarité des formes d'interaction** assure une meilleure acceptabilité du vocal lorsqu'il est associé à d'autres formes d'interaction : certains utilisateurs apprécieraient par exemple de combiner le vocal à la manipulation télécommandée [53]. Des travaux ont été réalisés par DeRenzi et al. [12] pour évaluer une interface professionnelle combinant le vocal et l'affichage web auprès d'agents de la santé. Les auteurs révèlent que certains professionnels préfèrent associer les canaux vocaux et visuels pour mieux ajuster les commandes vocales, enrichir les données récoltées, ou encore, pour vérifier la concordance des informations apportées [12]. Dans cette visée, nous observons que les modalités se soutiennent les unes aux autres [51]. De plus, certaines modalités se prêtent mieux à certaines opérations telles que les interactions gestuelles pour exprimer une commande télécommandée fréquente en situation de visionnage de film dans l'obscurité [53].

L'adaptation à l'utilisateur et au contexte ambiant représente également un facteur clé dans l'utilisation du vocal. Diverses caractéristiques sont appréciées en situation : les utilisateurs n'ont guère besoin de connaître la multitude des services et des applications intégrées dans l'assistant vocal pour utiliser ses principales fonctions. Le haut-parleur intelligent s'ajuste généralement à la demande formulée et au vocabulaire de l'utilisateur tout en continuant à cumuler des connaissances fines sur les utilisateurs cibles, de façon à contextualiser chaque proposition de service [29]. Par ailleurs, l'expérience utilisateur s'en trouve améliorée lorsque les interactions vocales continuent d'être efficaces en situation de bruits de fond. Whitenton affirme ainsi que l'assistant vocal Echo d'Alexa privilégie les interactions vocales sur tout le reste : intégrant sept microphones, il met l'accent sur la distinction entre la commande vocale et les bruits de fonds tels que les sources de musique [54].

Enfin, certains utilisateurs semblent apprécier le fait de privilégier certaines **formes d'adressage et d'interaction** avec ces interfaces **intégrant une forme de courtoisie et de respect des genres.** Contrairement à l'assistant vocal Alexa proposant des formes

d'interaction sur le ton de l'injonction (« *Alexa, éteins les lumières* » ; « *Alexa, programme mon réveil pour 8h* »), Google Home invite davantage à converser, tour en proposant une interaction plus respectueuse (« *Ok Google, lance l'actualité* » ; « *Hey Google, programme mon réveil pour 8h* »). Une utilisatrice attire ainsi l'attention sur l'impact émotionnel des technologies sur l'humain et sur le modèle et les valeurs sociales véhiculées par certains assistants vocaux : « *La différence d'expérience vécue ici est énorme. Lorsque j'utilise Alexa après la Google Home, je me sens autoritaire (...) Ce n'est pas isolé de la façon dont les hommes considèrent les femmes, mais cela s'étend aussi sur la manière dont les femmes se considèrent entre elles. Les moindres détails d'expérience utilisateur ont de l'importance (...) Je souhaite remercier Google. Merci d'avoir porté une attention à ces détails, et aux utilisatrices* » [37].

3.1.2 Les qualités du système : efficience, adaptabilité inter-individuelle et évolutivité

La **concision et l'adaptation des feedbacks** aux requêtes facilitent l'acceptabilité des technologies vocales. En effet, la pertinence et la justesse des réponses apportées par ces interfaces contribue grandement à la réussite de l'expérience utilisateur. Cette expérience positive est renforcée si l'utilisateur utilise son propre vocabulaire, il s'agit d'une commande vocale libre [53]. Il a d'ailleurs été observé que certains assistants vocaux, tels qu'Alexa, disposent d'une plus grande précision de réponse face aux requêtes des utilisateurs [54]. Alexa proposerait un traitement sémantique plus performant que Siri : « *Si vous mettez en place un chronomètre et que vous souhaitez savoir combien de temps il reste : Hey Siri, combien de temps il reste ? ==> J'ai trouvé un article sur le Times. Dois-je te le lire ? (...) Alexa, combien de temps reste-t-il ? ==> Environ 6min et 10 sec (...) Alexa interprète le temps comme lié au chronomètre et Siri le relie à une recherche du web, jusqu'à ce que le mot-clé « Timer » soit ajouté* » (op. cit.). Ainsi, éviter les biais d'interprétation en s'appuyant sur la commande vocale libre de l'utilisateur contribue à améliorer grandement l'expérience utilisateur.

Par ailleurs, nous pouvons remarquer que les technologies vocales intègrent naturellement une forme d'interaction pouvant s'adapter à une grande variété d'utilisateurs et d'individus, dont les personnes en situation de handicap ou « déficientes ». La commande et la reconnaissance vocale intègrent ainsi une **adaptation inter-individuelle** consistant à aménager des produits, des services ou des équipements susceptibles d'être utilisés par toute personne, peu importe le sexe, l'âge, la situation ou le handicap. L'interaction devient ainsi utilisable et mobilisable par une diversité de publics telles que les enfants, les illettrés, les personnes âgées, les malvoyants, ou les publics ayant des difficultés à interpréter certains graphiques ou certaines données complexes [43, 33, 12].

L'**évolution continue de l'IA** représente un autre facteur clé non négligeable dans l'intégration des technologies vocales dans les situations quotidiennes des utilisateurs. Les technologies de « *deep learning* » s'appuient sur un apprentissage automatique continu favorisant un cumul de connaissances contextualisées sur les utilisateurs qui sont constamment mis à jour [29]. De plus, les techniques vocales intégrées aux assistants vocaux évoluent sans cesse au rythme du web et non au rythme des cycles de conception, ce qui permet de renouveler de façon fluide de nouvelles applications et de nouveaux services [48].

3.1.3 Les effets favorables sur l'activité : cumul, continuité et personnalisation

La **capacité à cumuler des tâches** est l'une des conséquences et des opportunités permises par l'interaction vocale qui est grandement plébiscitée par les utilisateurs. Lors de situations contraintes où les mains et le regard sont affairés, les utilisateurs sont ainsi en capacité de réaliser des doubles tâches durant les routines matines par exemple : « *Pendant que je m'habille, je lui demande* : « *Alexa, lance le carburant » pour démarrer la machine à café (...) En continuant ma routine dans la salle de bain et ma chambre, je demande à Alexa de lancer n'importe quelle chanson, artiste, genre ou playlist dont j'ai envie* » [21]. Ce cumul de tâche révèle notamment des utilisations qui s'ancrent profondément dans les routines en faisant de la commande vocale un élément incontournable du quotidien : « *Avant même que j'ouvre les yeux, je discute avec Alexa. Éteins l'alarme. Allume la radio. Quel temps fait-il aujourd'hui ?* » ; « *D'habitude le matin, je peux mettre de la musique avant de sortir du lit. Ou bien, lorsque je sors de la douche avec les mains mouillées, plutôt que d'aller vers téléphone avec les mains humides, je peux juste lancer la musique pendant que je me prépare pour aller travailler* » [47]. Il est donc possible de gérer une diversité de tâches à partir des commandes vocales depuis n'importe quelle pièce sans l'utilisation des mains [51, 31, 21].

Dans le cours d'activité des utilisateurs et des habitants, nous observons que les technologies vocales favorisent une **mobilité et une continuité de l'expérience**. Ainsi, la possibilité de gérer des commandes vocales depuis n'importe quelle pièce du foyer sans les mains, ni du regard, favorise non seulement une grande mobilité dans l'espace domestique, mais elle permet notamment de soutenir la dynamique et le passage d'une activité à l'autre dans le cours d'action en situation domestique ou à l'extérieur (automobile, transports, etc.) [21, 47]. Nous pouvons illustrer par un verbatim cette forme de continuité de l'expérience d'un espace à un autre encouragée par les systèmes d'interaction vocale : « *Si je veux continuer mon livre audio pendant que je m'occupe du linge, du tricot ou autre chose qui ne demande un certain niveau d'attention, je dirai : Alexa, lit mon livre, pour reprendre là où j'en étais dans la voiture* ». L'utilisatrice n'a ainsi guère besoin de préciser de quel ouvrage il s'agit, car le dispositif intelligent comprend la demande sous-jacente [21].

La **création d'un écosystème personnalisé** représente une autre expérience réussie. En effet, les interactions vocales intégrées dans les haut-parleurs intelligents

©AfIA 2018

proposent un cadre dans lequel les utilisateurs sont invités à programmer des fonctions et des services adaptés à leurs besoins et leurs activités. Cet écosystème personnalisé s'accompagne par exemple d'une personnalisation des objets connectés environnants lorsque certains utilisateurs les commandent à distance : « *Alexa, allume la salle de contrôle* » (chambre) ; « *Alexa, lance le carburant* » (café) ; « *Alexa, éteint le repaire* » (salon) ==> « *Ça m'amuse, le nom est unique (...) ce n'est pas une mauvaise chose pour démarrer sa journée en ricanant* » [21]. Ce système constitué par les utilisateurs vise à contrôler un ensemble de services et de dispositifs via cet instrument pivot représenté par l'assistant vocal à l'aide des interactions vocales. Ce qui semble faire émerger une forme de centralisation des tâches via cette « tour de contrôle » qui facilite la diversité des commandes : « *Nous ne faisons pas des choses nouvelles. Nous réalisons juste un peu plus des tâches que nous faisions déjà. Et c'est plus facile comme ça. Nous écoutions déjà la radio, mais maintenant c'est plus facile de le faire. Tout est réuni au même endroit* » [47].

Cette structuration en écosystème semble également évoluer vers une **distribution spatiale et inter-individuelle des dispositifs** dans les foyers. Une enquête menée par Edison Research a démontré que 42% des foyers disposant d'un assistant vocal possède au moins deux haut-parleurs intelligents dans leurs foyers [28]. Lorsque l'utilisation est satisfaisante, ces foyers multiplient ces dispositifs afin de partager cette expérience utilisateur et ces services avec d'autres membres du foyer d'une part, et pour favoriser une continuité des activités via la commande vocale dans une diversité de pièces d'autre part : « *Notre premier Echo (Alexa), nous l'avons placé dans le salon pour le rendre accessible à n'importe qui. Nous l'avons tellement apprécié que nous avons acheté un Dot (Alexa) pour la chambre aussi* ». Ainsi, les dispositifs, initialement localisés dans les pièces collectives, sont progressivement intégrés vers les pièces plus intimes telles que les chambres à coucher, les salles de bain, ou encore les toilettes [47].

3.2 Les technologies vocales : difficultés et problèmes éthiques soulevés

3.2.1 Des formes d'interaction imposées

- **Empêchement dans l'action : visibilité et mobilité**

Des formes d'interaction contraintes ont notamment été soulevées telles que **l'absence d'écran visuel** pour soutenir le vocal afin de proposer une plus grande diversité et richesse des données[1] [31, 12, 32, 54]. Une étude réalisée par Luria et al. [31] portant sur des dispositifs domestiques intégrant une diversité d'interaction (un haut-parleur à commande vocale ; un écran tactile mural, une application mobile ; un robot social) a justement démontré l'importance du canal visuel auprès des utilisateurs pour disposer d'informations sur la réussite ou non des commandes réalisées et pour obtenir des informations sur la situation en cours et le contexte de la machine : « *Au final, je ne sais rien avec le vocal. C'était si facile de réaliser les commandes, mais finalement vous ne saviez pas ce qui n'allait pas* ». Ce qui peut contribuer à créer une impression de manque de contrôle (que nous aborderons dans la section suivante). D'autres empêchements opératoires ont été recensés tels que l'impossibilité pour un utilisateur d'annuler des achats via la commande vocale (l'usager doit passer une application), de paramétrer ou de personnaliser des fonctions via l'interaction vocale exclusivement [48]. L'**absence de mobilité** de ces dispositifs vocaux (la plupart des haut-parleurs intelligents doivent rester branchés électriquement, excepté le modèle Tap d'Amazon par exemple) empêche son transport d'une pièce à l'autre [54, 35]. Ce qui peut être susceptible de créer des ruptures dans la continuité des activités lorsque l'habitant se déplace d'une pièce à l'autre. Toutefois, pour pallier à cette contrainte, nous avons vu précédemment que de plus en plus d'habitants installent une multiplicité d'assistants vocaux par pièce du foyer [47].

- **Déstabilisation et malaise des usagers : sentiment de présence**

L'interaction vocale peut provoquer pour certains utilisateurs qui ne sont pas accoutumés, une **forme de déstabilisation** pouvant constituer un frein dans l'acceptation de ce type d'interaction : « *Pour utiliser l'interface vocale, on doit s'arrêter, et réfléchir à ce qu'on doit dire* » [31]. Par ailleurs, nous retrouvons une autre forme de déstabilisation plus récurrente avec ces technologies vocales : il s'agit d'un malaise et d'une crainte provoqués par un **sentiment de présence** liée à cette parole qui s'apparente à la présence d'un humain : « *Je n'aime pas le fait de me parler à moi-même, ça me donne une impression étrange et très bizarre… Je ne sais pas, c'est bizarre d'avoir une voix qui vous répond. C'est un peu effrayant* » ; « *Je ne peux pas vraiment expliquer pourquoi, mais je ne me sens pas à l'aise avec. Parce que c'est une machine et qu'il peut me répondre. Je me sens juste inconfortable* » ; « *Il pourrait devenir une sorte d'amis au bout d'un certain temps. On pourrait se sentir très à l'aise avec, comme si on était plus seul ? Peut-être que c'est un peu effrayant* » [31]. Ces technologies biomimétiques révèlent ainsi le risque de certaines technologies et intelligence artificielle à brouiller les frontières entre l'être vivant et les machines [1].

Ce malaise provoqué par ce sentiment de présence peut être d'autant plus marqué lorsque le dispositif vocal se met à fonctionner indépendamment des commandes ou du libre choix des utilisateurs. Pour illustrer cet aspect, nous pouvons citer un problème technique ayant été cité récemment dans l'actualité en Mars 2018 : un ensemble de haut-parleurs Alexa dans les foyers se mettaient à rire spontanément sans aucune raison apparente, et sans

1. À l'heure où nous écrivons ces lignes, de nouveaux assistants vocaux disposant d'un écran visuel ont été déployés sur le marché. Toutefois, notre étude ne s'appuie pas sur ces dispositifs multimodaux.

aucun lien avec une commande vocale, et cela à n'importe quel moment de la journée. Les utilisateurs ont proposé une diversité de qualificatifs pour décrire ce rire pouvant être « *sardonique* » ; « *effrayant* » ; « *inquiétant* » ; « *glaçant* » ; ou encore « *démoniaque intempestif* ». Voici quelques verbatims des utilisateurs ayant vécu cette défaillance technique : « *J'étais dans mon lit et sur le point de m'endormir quand Alexa a lancé un rire fort et terrifiant. Il y a de grandes chances que je sois assassiné cette nuit* » [7] ; « *Donc, Alexa a décidé de rire aléatoirement pendant que j'étais dans la cuisine. Je pensais qu'un gamin riait derrière moi* » ; « *On était en train de parler de choses assez confidentielles et Alexa s'est mise à rire. Quelqu'un d'autre a déjà eu ça ? On n'avait pas l'impression de l'avoir déclenchée accidentellement, elle s'est juste mise à rire comme ça. C'était très effrayant* » [24]. Vécue et jugée comme une expérience désagréable, voire effrayante par les usagers, ce bug technique renforce également ce sentiment de malaise par les enjeux d'anthropomorphisation des technologies que cela soulève.

3.2.2 Défaillances du système : manque d'intelligibilité, d'efficience et d'adaptabilité inter-individuelle au sein d'un même foyer

Parmi les défaillances du système vocal, nous pouvons citer une forme **d'opacité** sur les potentialités des services vocaux. Ce manque d'information peut nuire à l'utilisabilité si les utilisateurs ne connaissent pas l'étendue, ni les limites fonctionnelles de leurs appareils. Pour illustrer ce type de défaillance, nous citerons la logique de fonctionnement technique du dispositif vocal Alexa. Cet assistant vocal intègre des paramétrages par défaut que bien souvent, les utilisateurs ne connaissent pas d'emblée ou n'en saisissent pas toutes les subtilités et les impacts possibles sur leurs activités. Parmi ces défauts de paramétrage : Alexa propose des services ou des produits que l'on ne cherche pas à acheter ; l'utilisateur ne peut pas demander à Alexa d'annuler un achat, il faut le faire depuis une application ; le paramétrage par défaut n'inclut pas un code pin à 4 chiffres pour éviter des requêtes d'achats qui ne proviennent pas de l'utilisateur principal. Et enfin, Alexa autorise n'importe quelle personne proche du dispositif à commander des appareils et des services depuis le compte Amazon de l'utilisateur principal, et cela inclus les enfants, les voix de la radio ou de la télévision [8]. Ce qui induit une perte de contrôle nuisible à l'utilisabilité et l'acceptabilité de tels dispositifs.

L'échec d'interprétation des commandes et du contexte constitue un autre frein à l'acceptabilité de ces technologies. Parmi cette catégorie nous retrouvons trois principales contraintes. Tout d'abord, la *mauvaise reconnaissance vocale* induit un manque de fiabilité en cas de mauvaise interprétation des commandes tels que les biais sémantiques évoqués précédemment, d'une mauvaise qualité de la reconnaissance vocale pour certaines voix retrouvées plus particulièrement auprès des personnes âgées, les enfants ou les personnes atteintes d'une maladie [43, 51]. Cet échec de reconnaissance vocale ou de détection des phrases d'activation peut aussi intervenir en cas de perturbation selon l'ambiance sonore tels que les fonds musicaux ou les voix multiples [54]. Nous retrouvons ensuite la *dimension incommode* de la commande vocale selon des situations spécifiques. Réaliser une liste d'achats par exemple soulève certaines limites ainsi qu'un défaut d'interprétation dans la commande vocale qui entache l'expérience utilisateur : « *Ajouter un élément dans la liste de course est facile et rapide, mais ajouter plusieurs articles devient vite répétitif et coûteux temporellement. Vous devez attendre la confirmation verbale avant de répéter le procédé (...) Après le troisième article, vous commencez à penser qu'il pourrait sûrement y avoir un moyen plus rapide d'y parvenir. Vous pouvez raccourcir les étapes en ajoutant plusieurs articles en même temps, mais la liste finit par afficher une chaîne de texte comme si cela représentait un seul article* » (op. cit.). Enfin, nous identifions une dernière dimension liée au *collectif*. Si l'assistant vocal ne dispose pas de la capacité à distinguer les différentes voix du foyer, son partage auprès de plusieurs membres peut empêcher la personnalisation du système à l'utilisateur cible et ainsi modifier les futures recommandations et la compréhension du contexte de l'usager [11].

Par ailleurs, certains utilisateurs évoquent une déception et une mauvaise expérience utilisateur lorsqu'ils découvrent les **limites de certains « skills »** d'assistants vocaux ne supportant pas les requêtes complexes impliquant un ensemble de facteurs : « *En lançant l'application « tremblement de terre », on ne peut pas demander le plus gros tremblement de terre qui a eu lieu la semaine passée et l'application envoie une information erronée* » [48].

3.2.3 Renoncement à certains principes éthiques : confidentialité, capacité d'agir, contrôle et diversité servicielle

L'utilisation des interfaces vocales, et plus précisément celui des assistants vocaux, invite à questionner la dimension éthique des données. En effet, cette problématique met en exergue la confidentialité et la sécurisation des données mises à mal par ces technologies qui doivent cumuler une quantité massive de données personnelles pour fonctionner efficacement [11]. Dans un souci d'efficacité du dispositif et de contextualisation des services, les utilisateurs sont encouragés à **renoncer à la confidentialité** totale de leurs données dont on recommande un stockage permanent sur les serveurs afin d'assurer une mémoire des actions et des requêtes [43, 11, 32]. Certains utilisateurs évoquent une intrusion dans l'intimité liée à l'enregistrement continue des activités et des paroles, un inconfort à diffuser certaines requêtes et l'absence d'un bouton mute à part entière sur certains dispositifs : « *Je n'aime pas parce que j'ai la sensation d'être toujours écouté, et je dois rester prudente sur ce que je lui dis* » [31]. Cette question de la confidentialité est également soulevée en situation d'utilisation partagée d'un assistant vocal au sein du foyer : ce qui pourrait

©AfIA 2018

occasionner l'intervention spontanée du dispositif pour proposer un produit ou un service dans un moment inadapté ou en présence d'autres membres ou de personnes extérieures. Ces situations d'utilisation nous amènent ainsi à distinguer deux types de confidentialité à intégrer en conception : il s'agit de la *discrétion* – renvoyant au souhait de ne pas être entendu des autres personnes lors des interactions vocales (échelle locale) – et de la *privauté* – suggérant la protection et la sécurisation des données personnelles stockées par les dispositifs (échelle plus globale) –.

Les publics âgés mettent notamment l'accent sur certains risques inhérents à la commande vocale, celui de développer une **passivité et une oisiveté** dans les tâches quotidiennes qu'ils préfèrent prendre le temps de réaliser manuellement [51, 43]. Il est donc impératif de privilégier des tentatives d'aide à la personne qui ne réduisent ni le sens du contrôle, ni leur propre estime [49]. Les travaux de Portet et al. [43] sur les systèmes vocaux en situation domestique soulignent bien l'importance de préserver auprès des personnes âgées ce sentiment de contrôle et leur capacité à agir en toute autonomie : « *J'aime bien agir plutôt que parler (...) J'aime bien fermer mes volets, etc. (...) Moi en ce moment je préfère faire les choses parce que sinon c'est glisser vers l'inactivité. Faudrait vraiment que je puisse plus le faire, parce que sinon on fait plus rien, on se couche et puis voilà* » (op. cit.).

La **perte de contrôle** représente un autre frein à l'acceptabilité du vocal cité par certains utilisateurs. Ils évoquent une forme d'inconfort à perdre l'initiative et le contrôle de certaines fonctions ou utilisations d'objets domestiques : « *En utilisant le vocal, ça donne l'impression que c'est vous qui décidez mais quelqu'un d'autre décide aussi pour vous. Sur cet écran tactile, vous sentez que vous avez le contrôle à 100%, mais ici avec le vocal, c'est comme si vous attendiez qu'ils acceptent ce que vous demandez. L'inconvénient c'est qu'on ne se sent pas à 100% en situation de contrôle comme les autres interfaces. C'est comme si on sollicitait une chose à la maison et qu'on doit attendre que ce soit approuvé* » [31].

Des utilisateurs évoquent aussi un **manque de diversité de choix** dans les propositions spontanées et contextualisées de certains assistants vocaux [11, 8]. Davis [11] évoque plusieurs paramètres inhérents au dispositif vocal Alexa intégrant des recommandations (dont la logique de fonctionnement reste opaque pour les utilisateurs) sont proposées aux utilisateurs sans choix possibles : l'assistant vocal propose des achats à partir de l'historique de l'usager, du choix d'Amazon (meilleures notations, meilleurs prix, etc.) et des items éligibles lors de la commande « *prime* ». Ce dispositif se retrouve donc dans une position privilégiée et unique pour dicter des choix de produits et des choix de médias (op. cit.). Ce fonctionnement s'appuie sur l'usage d'algorithmes de recommandation pouvant non seulement empêcher l'utilisateur de disposer d'une liberté de choix dont il disposerait en délaissant les interactions vocales, mais ces algorithmes peuvent notamment déployer une bulle filtrante autour des usagers. Ce qui les amènerait à consommer des produits ou des services, à se divertir et à s'informer le plus souvent à partir des mêmes sources sans s'ouvrir à d'autres univers ou d'autres opinions [38].

4. Des principes de conception d'interaction à privilégier

En nous appuyant sur les éléments récoltés concernant la compréhension des formes d'utilisation de technologies vocales en termes de leviers et de freins dans l'utilisabilité et l'acceptabilité, nous proposons ici des lignes directrices pour alimenter la conception de systèmes d'interaction vocale à destination des utilisateurs dans de multiples contextes. Selon la diversité des utilisations appréciées et des difficultés rencontrées, nous recommandons de privilégier dans les interfaces vocales les dix propositions suivantes :

Feedback utilisateur : encourager un temps de réponse relativement rapide (entre la commande et le retour vers l'utilisateur) ; privilégier la présence d'un voyant lumineux ou d'une réponse vocale confirmant la réception de la commande vocale ; intégrer une transparence des sources utilisées pour les demandes d'information et les services de recommandation divers (pour les utilisateurs intéressés).

Mode d'interaction et multimodalité : concevoir l'interaction vocale en complémentarité avec les autres modalités d'interaction existantes pour favoriser une richesse et une diversité de commande et de feedbacks (exemple : vocal, télécommande et écran visuel) ; proposer des interactions vocales privilégiant la parole libre (mode de conversation routinier).

Adaptabilité : assurer une adaptation du système à partir de l'intégration du vocabulaire de l'utilisateur ; proposer un système qui apprend également de ses erreurs ; inclure la possibilité de former le système par l'utilisateur.

Confidentialité et sécurité : intégrer un bouton « mute » à part entière ; assurer une confidentialité des données récoltées et stockées sur un serveur protégé ; privilégier une utilisation multi-utilisateurs (profils, droits d'accès) ; permettre le contrôle, l'exploitation et la modification de ses données personnelles et rendre visible les données recueillies par le système.

Contrôle du système : inclure la possibilité aux utilisateurs de garder ou de reprendre la main sur le système en cas de dysfonctionnement (pouvoir utiliser les luminaires ou l'utilisation de volets électriques) ou selon les préférences des habitants au sein du foyer. Par exemple, il peut s'agir du débrayage, d'interagir de façon spontanée ou encore de prévoir des automatisations selon le niveau de maîtrise du système (via le mobile, le PC ou par commande vocale). Et surtout il s'agit d'informer de façon continue, anticipée ou instantanée les utilisateurs sur les changements d'états du système via un système d'alertes et de notifications pertinentes.

Personnalisation : offrir la possibilité de modifier les sources utilisées pour les recherches d'information et les

services de recommandations ; programmer et personnaliser les services avec l'interaction vocale (en plus des interactions via le mobile) ; proposer une personnalisation originale des objets connectés et des phrases d'activation ; permettre et faciliter l'utilisation de la programmation visant l'automatisation ou le déclenchement d'actions uniques ou regroupées selon des situations spécifiques avec les commandes dite IFTTT (« *If This, Then That* ») ; adapter le système aux contextes utilisateurs en intégrant la compréhension de l'implicite auprès des dispositifs (ex : inférer la position et l'activité des utilisateurs).

Qualité de capture sonore : favoriser une reconnaissance vocale selon la diversité des contextes ambiants via l'intégration de micro embarqués, de capteurs et de dispositifs d'annulation d'écho (pour compenser la position de l'utilisateur, l'acoustique de l'habitat et la présence de bruits de fond).

Ubiquité : favoriser une continuité de l'expérience entre les membres du foyer, leurs diverses préoccupations domestiques, les différentes espaces empruntés, les nombreux dispositifs utilisés et les autres modalités d'interaction existantes en proposant un écosystème se déployant sur une variété de supports.

Multi-utilisateurs : privilégier un système de reconnaissance des voix au sein d'un même foyer (pour éviter le « détournement » du système par des voix extérieures : TV, radio, etc.) ; proposer un paramètre de contrôle pour l'utilisateur principal afin de garder la main sur ses propres données et déléguer s'il le souhaite une partie ou la totalité du contrôle du dispositif à d'autres utilisateurs.

Dimension humanisante : encourager la dimension humanisante lors des interactions pour les utilisateurs intéressés par cette expérience utilisateur. Cela consisterait pour l'utilisateur à choisir entre un « *assistant fonctionnel* » ou un « *assistant de compagnonnage* ». Ce dernier choix proposerait une variété de critères et d'attributs que l'utilisateur pourrait ajouter, ajuster ou enlever selon ses préférences : variation de voix ou de timbre, intégration de marques de politesse et/ou d'humour, sélectionner un type de personnalité pour l'IA, expression et reconnaissance des émotions, etc.

5. Discussion et perspectives

En définitive, les systèmes d'interaction vocale n'ont pas vocation à remplacer les autres systèmes et modalités d'interaction. En effet, nous avons observé que les utilisateurs préfèrent garder le contrôle de leurs dispositifs en ayant recours à diverses modalités selon la situation d'utilisation et les besoins exprimés. Bien que basique dans les formes d'utilisations déployées, l'interaction vocale vient surtout faciliter et renforcer des utilisations préexistantes dans les routines des habitants (écouter de la musique/actualités, allumer/éteindre les lumières ou son alarme, noter sa liste de course, etc.). Par ailleurs, nous remarquons que les interactions vocales sont majoritairement utilisées en situation de multitâches lorsque le regard et les mains sont affairés, et cela afin de poursuivre quatre principales logiques d'utilisation : la recherche d'assistance, le divertissement, le pilotage d'objets connectés, ainsi que l'anticipation et l'organisation de son activité. Mais pour renforcer l'acceptabilité de ces systèmes, il est nécessaire de renforcer une expérience utilisateur intuitive et efficiente dans laquelle ces systèmes s'adaptent au vocabulaire propre à l'usager, comprennent l'implicite de certaines commandes vocales et s'adaptent aux diverses situations d'utilisations. Les technologies vocales telles que les assistants vocaux favorisent notamment une expérience partagée par tous les membres du foyer, il est donc indispensable de penser les interactions et les utilisations vers une diversité de profils (adultes, enfants, âgés) et d'y intégrer une discrétion des données grâce à la différenciation des voix et une hiérarchisation des niveaux de contrôle de l'assistant en cas d'utilisation collective. Enfin, il est indispensable de penser le vocal comme une modalité d'interaction pouvant soutenir la continuité des activités.

La synthèse des utilisations et des retours d'expériences sur les technologies vocales et les assistants vocaux démontre bien l'importance de mettre à jour les activités réelles des utilisateurs, leurs préoccupations, ainsi que leurs logiques intrinsèques. Élaborées à partir d'une première exploration de la littérature sur les systèmes vocaux, les principes de conception d'interaction précédemment évoqués ne constituent qu'un cadre général pour intégrer les logiques d'utilisation et les besoins utilisateurs dans la conception de systèmes utilisant IA et reconnaissance vocale. Nous pointons l'importance de baser la conception et le développement de ces dispositifs sur une approche centrée utilisateurs, itérative et pluridisciplinaire intégrant :

1. L'analyse et la modélisation des activités en situation de référence où le futur dispositif sera/est déployé : celle-ci vise à contextualiser les besoins et les logiques d'utilisation des différents profils utilisateurs ;
2. Une définition des fonctionnalités afin de proposer un ensemble de scénarios d'usages (use-case) pertinents et des principes d'interactions à respecter pour le projet de conception ;
3. Des phases de maquettage et d'évaluation de ces fonctionnalités, formes d'interaction et scénarios. L'évaluation en situation réelle ou réaliste d'utilisation permet de valider les choix de conception et de connaître l'appropriation réellement faite du système.

En application de cette démarche, et en complémentarité avec l'analyse de la littérature ici présentée, nous réalisons actuellement des observations en situation réelle afin de modéliser le développement et les utilisations d'écosystèmes d'objets connectés dont des systèmes à base de reconnaissance vocale. Ces travaux de recherche en ergonomie contribuent à la conception de services et de dispositifs soutenant la maîtrise de l'énergie dans l'habitat dans le secteur énergétique

©AfIA 2018

adaptés au grand public et s'insérant de façon fluide et durable dans les activités et les routines des habitants.

Bibliographie

[1] Alliance des Sciences et Technologies du Numérique (ALLISTENE). *Éthique de la recherche en robotique*. Rapport n°1 de la CERNA, Commission de réflexion sur l'Éthique de la Recherche en sciences et technologies du Numérique d'Allistene, 2014.

[2] J. Barcenilla & J. Bastien. L'acceptabilité des nouvelles technologies : quelles relations avec l'ergonomie, l'utilisabilité et l'expérience utilisateur ? *Le travail humain*, vol. 72, (4), 311-331, 2009.

[3] J. M. C. Bastien, D.L. Scapin & C. Leulier, Une comparaison des critères ergonomiques et des principes de dialogue ISO 9241-10 dans une tâche d'évaluation d'interface. *Revue d'interaction homme-machine*, 1 (1), pp. 33-63, 1998.

[4] V. Belloti & K. Edwards, Intelligibility and accountability: Human consideration in context-aware systems. *Human-Computer Interaction*, 162 (2-4): pp. 193-212, 2001.

[5] M.E. Bobiller-Chaumon, M. Dubois & M. Retour, *E-banking: Nouveaux services, nouveaux usages, nouvelles compétences*. Paris, France, 2003.

[6] E. Brangier, A. Dufresne, S. Hammes-Adelé, Approche symbiotique de la relation humain-technologie : perspectives pour l'ergonomie informatique. *Le travail humain*, 2009/4 (Vol. 72), pp. 333-353, 2009.

[7] A. Charney, Amazon Alexa s'esclaffe à cause d'un problème d'écoute, 2018. *01net.com*, http://www.01net.com/actualites/quand-amazon-alexa-s-esclaffe-et-terrifie-ses-utilisateurs-1391170.html

[8] S. Cobb, Amazon Echo and the Alexa dollhouses: Security tips and takeaways, 2017. *We live security.com*, https://www.welivesecurity.com/2017/01/10/amazon-echo-alexa-dollhouse-security/

[9] Commission Nationale Informatique & Libertés (CNIL), *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*. Synthèse du débat public animé par la CNIL dans le cadre de la mission de réflexion confiée par la loi pour une république numérique, Décembre 2017.

[10] J. Coutaz & L. Nigay, Multimodalité et plasticité des interfaces homme-machine en informatique ambiante : concepts et espaces de conception. Florence Sèdes, Jean-Marc Ogier, Pierre Marquis. *Information Interaction intelligence - Le point sur le i* (3), Cépaduès Editions, pp.179-214, 2012.

[11] B. Davis, The problem with voice user interfaces like Amazon Alexa, 2016. *Econsultancy*, https://econsultancy.com/blog/68499-the-problem-with-voice-user-interfaces-like-amazon-alexa

[12] B. DeRenzi, N. Dell, J. Wacksman, S. Lee & N. Lesh, Supporting Community Health Workers in India through Voice- and Web-Based Feedback. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 2017.

[13] M. Dubois & M. E. Bobillier-Chaumon, L'acceptabilité des technologies : bilans et nouvelles perspectives. *Le travail humain*, 72(4), 305, 2009.

[14] A. Edip, Siri, Alexa, Google Home, M… Les assistants numériques, révolution ou gadget dangereux ?, 2017. Capital.fr, *Polemik*, https://www.capital.fr/polemik/siri-alexa-google-home-m-les-assistants-numeriques-revolution-ou-gadget-dangereux-1230739

[15] M-F. Fortin. *Le processus de la recherche : de la conception à la réalisation*. Montréal : Décarie éditeur, 1996.

[16] M. Fréjus, *Construire la coopération Homme-IA*. Document interne. Saclay : EDF R&D, 2017.

[17] M. Fréjus, Élargissement et renouvellement des questions traitées par l'ergonomie dans le domaine du développement durable : retour sur 12 ans de travaux sur les activités domestiques et la maîtrise des consommations énergétiques. *Psychologie française*, à paraître.

[18] FUTUR.E.S/Cap Digital, Futur.e.s # Humain X Machine : 3 dimensions de l'interaction illustrées par nos démos, 2017. *Blog.futuresfestivals.com*, https://blog.futuresfestivals.com/futur-e-s-humainxmachine-3-dimensions-de-linteraction-illustrees-nos-demos/

[19] D. Gardner-Bonneau, *Human Factors and Voice Interactive Systems*. Kluwer Academic Publishers, Norwell, MA, 1999

[20] H. Guillaud, Vers une éthique pour l'intelligence artificielle ?, 2016. *Internet actu.net*, http://www.internetactu.net/a-lire-ailleurs/vers-une-ethique-pour-lintelligence-artificielle/

[21] A. L. Hamilton, How i use Alexa, 2017, *Love my Echo.com*,https://lovemyecho.com/2017/01/03/how-i-use-alexa/

[22] R. Harris, V*oice interaction Design, Crafting the new conversational speech systems.* San Francisco: Morgan Kaufmann Publishers, 2005.

[23] J.-B., Haué, J.-B. Intégrer les aspects situés de l'activité dans une ingénierie cognitive centrée sur la situation d'utilisation. *Activités*, 2004.

[24] H. Hillard, Amazon. L'assistant personnel Alexa se met à rire tout seul et effraye les utilisateurs, 2018. *Ouest-France.fr*, https://www.ouest-france.fr/economie/entreprises/amazon/amazon-son-assistant-personnel-alexa-se-met-rire-tout-seul-et-effraye-les-utilisateurs-5610264

[25] Institut National de Recherche en Informatique et en Automatique (INRIA), *Intelligence Artificielle. Les défis actuels et l'action d'Inria*. Livre blanc N°01. Coordonné par B. Braunschweig, 2016.

[26] ISO 9241-11, *Exigences ergonomiques pour travail de bureau avec terminaux à écrans de visualisation (TEV)* - Partie 11: Lignes directrices relatives à l'utilisabilité, 1998.

[27] L. Janin, L'intelligence artificielle : mythes et réalités, 2017. *Stratégie.gouv.fr*, http://www.strategie.gouv.fr/point-de-vue/lintelligence-artificielle-mythes-realites

[28] B. Kinsella, 42 Percent of amazon echo owners have 2 or more devices, 2017. *Voicebot.ai*, https://www.voicebot.ai/2017/06/22/42-percent-amazon-echo-owners-2-devices/

[29] F. Lefevre, Interaction vocale & intelligences artificielles : état des lieux et opportunités, 2017. *Medium.com*, https://medium.com/@francklefevre/interaction-vocale-intelligences-artificielles-%C3%A9tat-des-lieux-et-opportunit%C3%A9s-3b45239ee4a7

[30] É. Loup-Escande, J. Burkhardt & S. Richir (2013). Anticiper et évaluer l'utilité dans la conception ergonomique des technologies émergentes : une revue. *Le travail humain*, vol. 76, (1), 27-55, 2013.

[31] M. Luria, G. Hoffman & O. Zuckerman, Comparing Social Robot, Screen and Voice Interfaces for Smart-Home Control. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17). ACM, New York, NY, USA, pp. 580-628, 2017.

[32] O. Metwally & B. Gregory B, Alexa, start living healthy! Workshop on Interactive Systems in Health Care (WISH), 2017, http://wish.gatech.edu/wp-content/uploads/metwally_omar.pdf

[33] C. Munteanu & G. Penn, Speech-based Interaction: Myths, Challenges, and Opportunities. *In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, New York, NY, USA, pp. 992-995, 2016.

[34] C. Nass & S. Brave, Wired for Speech – How Voice Activates and Advances the Human–Computer Relationship. *MIT Press*, 2007.

[35] N. Nguyen, I Finally Bought An Amazon Echo And This Is What Happened, 2016. *Buzzfeed.com*, https://www.buzzfeed.com/nicolenguyen/amazon-echo-review?utm_term=.ftMWKOVwe#.ifDRxV3Wv

[36] J. Nielsen, *Usability engineering*. Boston: Academic Press, 1993.

[37] J. Paolino, Google Home vs Alexa. Two Simple User Experience Design Gestures that Delighted a Female User, 2017. *Medium.com*, https://medium.com/startup-grind/google-home-vs-alexa-56e26f69ac77

[38] Pariser, *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Press, 2011.

[39] H. Pasquier, Définir l'acceptabilité sociale dans les modèles d'usage. Vers l'introduction de la valeur sociale dans la prédiction du comportement d'utilisation. *Thèse de Doctorat en psychologie sociale*, Université de Rennes 2, 2012.

[40] C. Pelletier & É. M.-F. Moreau, L'appropriation des technologies de l'Internet et ses facteurs critiques de succès : un défi de plus pour les PME ? *Revue internationale P.M.E : économie et gestion de la petite et moyenne entreprise, vol. 21, 2, pp. 75-117, 2008.*

[41] P. Pestanes & B. Gautier, Essor des assistants vocaux intelligents : nouveau gadget pour votre salon ou fenêtre d'opportunité pour rebattre les cartes de l'économie du web ?, *Wavestone*, 2017.

[42] R. Pinard, P. Potvin & R. Rousseau. Le choix d'une approche méthodologique mixte de recherche en éducation. *Recherches qualitatives*, Vol. 24, pp.58-82, 2004.

[43] F. Portet, M. Vacher & S. Rossato, Les technologies de la parole et du TALN pour l'assistance à domicile des personnes âgées : un rapide tour d'horizon (Quick tour of NLP and speech technologies for ambient assisted living) [in French]. *JEP-TALN-RECITAL 2012, Atelier ILADI 2012: Interactions Langagières pour personnes Agées Dans les habitats Intelligents*, Jun 2012, Grenoble, France. pp.3-16, 2012.

[44] Portolan, N., Nael, M., Renoullin, JL & Naudin, S. (1999). Will we speak to our TV remote control in the future? In: *Proceedings of the 17th international symposium on human factors in telecommunication*, HFT'99, Copenhagen.

[45] S. Proulx, Trajectoires d'usages des technologies de communication : les formes d'appropriation d'une culture numérique comme enjeu d'une société du savoir. *Annales des télécommunications*, 57(3-4), pp. 180-189, 2002.

[46] P. Rabardel. *Les hommes et les technologies, une approche cognitive des instruments contemporains.* Armand Colin, Paris.

[47] Radiocentre, Getting vocal. How voice-activated devices are increasing radio listening & elevating audio branding, 2017. Full report, http://www.radioplayer.co.uk/sites/default/files/getting_vocal_-_final_report.pdf

[48] D. Rasmus, Hey, Alexa: The Pros and Cons of My Amazon Echo Experience, 2016, *Iphonelife.com*, https://www.iphonelife.com/content/hey-alexa-pros-and-cons-my-amazon-echo-experience

[49] J. Rodin, Aging and health: effects of the sense of control. *Science*, 233(4770): pp.1271- 1276, 1986.

[50] B. Shackel, Human Factors for Informatics Usability. In B. Shackel & B. Richardson (Eds.), *Human Factors* (pp. 21-38). Cambridge: Cambridge University Press, 1991.

[51] H. Soronen, M. Turunen & J. Hakulinen, Voice Commands in Home Environment - a Consumer Survey. In *Proceedings of Interspeech 2008*, pp. 2078-2081, 2008.

[52] J. Theureau. *Appropriation 1, 2, 3 ou Appropriation, Incorporation & « Inculturation »*. Conférence journée Ergo-Idf Appropriation & Ergonomie, 16/06/11, Paris, 2011.

[53] M. Turunen, A. Melto, J. Hella, T. Heimonen, J. Hakulinen, E. Mäkinen, T. Laivo, H. Soronen, User expectations and user experience with different modalities in a mobile phone controlled home entertainment system. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services.* ACM, New York, NY, USA, 2009.

[54] K. Whitenton, Voice Interaction UX: Brave New World...Same Old Story, 2016. *Nielsen Norman Group.com*,https://www.nngroup.com/articles/voice-interaction-ux/

©AfIA 2018

Codage SMT dans un espace de plans (liens causaux) pour la planification temporelle en temps continu

Frédéric Maris Maël Valais Julien Vianey
IRIT, Université de Toulouse

{maris, mael.valais, julien.vianey}@irit.fr

Résumé

L'amélioration considérable des solveurs SAT a permis de les utiliser pour la résolution de divers problèmes d'intelligence artificielle, et en particulier la génération automatique de plans d'actions. Récemment, des solveurs SMT (SAT Modulo Theory) prometteurs basés sur des solveurs SAT efficaces ont été développés. Il est donc intéressant d'utiliser le langage SMT qui nous permet de produire directement des codages de problèmes de planification temporelle sans avoir à discrétiser le temps ou à utiliser un gestionnaire de contraintes temporelles externe (Time Map Manager). Cette approche a déjà été explorée, mais la plupart des codages existants considèrent une modélisation discrète du temps. Nous introduisons dans cet article une nouvelle traduction de problèmes de planification temporelle en temps continu, basée sur un codage dans un espace de plans (liens causaux), en formules propositionnelles modulo QF-RDL (Quantifier Free Rational Difference Logic).

Mots Clef

Planification, codages, temps, SAT, SMT.

1 Introduction

Dans le planificateur Satplan [8], un problème de planification est transformé en une formule propositionnelle dont les modèles, correspondant aux plans solution, peuvent être trouvés en utilisant un solveur SAT. Cette approche SAT recherche un plan solution de longueur fixe k. En cas d'échec à trouver un tel plan, cette longueur est augmentée avant de recommencer la recherche d'une solution. Cette approche bénéficie directement des améliorations des solveurs SAT [1]. L'exemple le plus évident est le planificateur Blackbox [10, 11] (et ses successeurs Satplan'04 [7] et Satplan'06 [12]). Ces planificateurs ont obtenu la première place dans la catégorie planificateur optimal (en termes de nombre d'étapes du plan) des compétitions internationales de planification [2] IPC-2004 et IPC-2006. C'était inattendu car ces planificateurs étaient essentiellement des mises à jour de Blackbox et n'incluaient aucune réelle nouveauté : l'amélioration des performances était principalement due aux progrès du solveur SAT sous-jacent.

De nombreuses améliorations de cette approche originale ont été proposées, notamment via le développement de codages plus compacts et efficaces [9, 3, 13, 14, 18, 19, 20]. Suite à ces travaux, de nombreuses autres techniques similaires pour le codage de problèmes de planification ont été développées : programmation linéaire (LP) [23], problèmes de satisfaction de contraintes (CSP) [2]. En particulier, différents types de codages SMT ont aussi été proposés pour résoudre des problèmes de planification temporelle. Les codages de [22] et [21] se basent sur une représentation discrète du temps, alors que dans [17, 16] nous utilisons des atomes de QF-RDL (Quantifier Free Rational Difference Logic) pour coder un temps continu. Dans la section 2, nous définissons notre modèle de planification temporelle, précédemment introduit dans [1]. Nous décrivons dans la section 3 un nouveau codage SAT basé sur des conditions ouvertes (liens causaux) pour la planification classique. Ensuite, dans la section 4, nous proposons une adaptation SMT de ce codage pour la planification temporelle en temps continu.

2 Planification temporelle

Nous étudions la planification temporelle propositionnelle dans un langage basé sur les aspects temporels de PDDL2.1 [4]. Un fluent est une proposition atomique positive ou négative. Comme dans PDDL2.1, nous considérons que les changements des valeurs des fluents sont instantanés mais que les conditions sur la valeur des fluents peuvent être imposées sur un intervalle. Une action a est un quadruplet $\langle Cond(a), Add(a), Del(a), Constr(a)\rangle$, où l'ensemble des conditions $Cond(a)$ est l'ensemble des fluents qui doivent être vrais pour que a soit exécutée, l'ensemble des ajouts $Add(a)$ est l'ensemble des fluents qui sont établis par a, l'ensemble des retraits $Del(a)$ est l'ensemble des fluents qui sont détruits par a, et l'ensemble de contraintes $Constr(a)$ est un ensemble de

1. http://www.satcompetition.org/
2. http://www.icaps-conference.org/index.php/Main/Competitions

 ©AfIA 2018

contraintes entre les temps relatifs des événements qui se produisent pendant l'exécution de a. Un événement correspond à l'une des quatre possibilités : l'établissement ou la destruction d'un fluent par une action a, ou le début ou la fin d'un intervalle sur lequel un fluent est requis par une action a. Dans PDDL2.1, les événements ne peuvent se produire qu'au début ou à la fin des actions, mais nous relâchons cette hypothèse de sorte que les événements peuvent se produire à tout moment à condition que les contraintes $Constr(a)$ soient satisfaites. Notons que $Add(a) \cap Del(a)$ peut être non vide. En effet, il n'est pas inhabituel pour une action durative d'établir un fluent au début de l'action et de le détruire à sa fin. Nous pouvons également observer que la durée d'une action, le temps entre le premier et le dernier événement de l'action, n'a pas besoin d'être explicitement stockée.

Nous utilisons la notation $a \rightarrow f$ pour désigner l'événement que l'action a établit le fluent f, $a \rightarrow \neg f$ pour désigner l'événement que a détruit f, et $f \mid\!\rightarrow a$ et $f \rightarrow\!\mid a$, respectivement, pour indiquer le début et la fin de l'intervalle sur lequel a requiert la condition f. Si f est déjà vrai (respectivement, faux) lorsque l'événement $a \rightarrow f$ ($a \rightarrow \neg f$) se produit, nous considérons toujours que a établit (détruit) f. Un plan temporel peut contenir plusieurs instances de la même action, mais par simplicité de notation, nous ferons seulement la distinction entre actions et instances d'action si cela est absolument nécessaire. Nous utilisons la notation $\tau(e)$ pour représenter l'instant dans un plan où un événement e se produit. Pour une action donnée (ou une instance d'action) a, $Events(a)$ représente les différents événements qui constituent sa définition, à savoir $a \rightarrow f$ pour tout f dans $Add(a)$, $a \rightarrow \neg f$ pour tout f dans $Del(a)$, $f \mid\!\rightarrow a$ et $f \rightarrow\!\mid a$ pour tout f dans $Cond(a)$. La définition d'une action a inclut les contraintes $Constr(a)$ sur les instants relatifs aux événements dans $Events(a)$. Comme dans PDDL2.1, nous considérons que la durée entre événements dans $Events(a)$ n'est pas nécessairement fixe et que $Constr(a)$ est un ensemble de contraintes d'intervalle sur des paires d'événements, tels que $\tau(f \rightarrow\!\mid a) - \tau(f \mid\!\rightarrow a) \in [\alpha, \beta]$ pour des constantes α, β. Nous utilisons $[\alpha_a(e_1, e_2), \beta_a(e_1, e_2)]$ pour désigner l'intervalle de valeurs possibles pour la distance relative entre les événements e_1, e_2 de l'action a. Une durée fixe entre les événements $e_1, e_2 \in Events(a)$ peut, bien sûr, être modélisée en définissant $\alpha_a(e_1, e_2) = \beta_a(e_1, e_2)$. De même, l'absence de contrainte peut être modélisée par l'intervalle $[-\infty, +\infty]$. Nous introduisons maintenant deux contraintes fondamentales que tous les plans temporels doivent satisfaire :

— les *contraintes inhérentes* à l'ensemble des instances d'action $\mathcal{A}$: pour toute $a \in \mathcal{A}$, a satisfait $Constr(a)$, c'est-à-dire pour toutes les paires d'événements $e_1, e_2 \in Events(a)$ nous avons $\tau(e_1) - \tau(e_2) \in [\alpha_a(e_1, e_2), \beta_a(e_1, e_2)]$;

— les *contraintes d'effets contradictoires* sur l'ensemble des instances d'action $\mathcal{A}$: pour toutes $a_i, a_j \in \mathcal{A}$, pour tout fluent positif $f \in Del(a_i) \cap Add(a_j)$ nous avons $\tau(a_i \rightarrow \neg f) \neq \tau(a_j \rightarrow f)$.

Les contraintes inhérentes définissent la structure interne de chaque instance d'action, alors que les contraintes d'effets contradictoires assurent que la valeur de vérité de chaque fluent ne soit jamais indéfinie lors de l'exécution d'un plan temporel.

Définition 1 *Un problème de planification temporelle $\langle I, \mathcal{A}, G \rangle$ consiste en un ensemble d'actions $\mathcal{A}$, un état initial I et un but G, où I et G sont des ensembles de fluents.*

Notation Si $\mathcal{A}$ est un ensemble d'instances d'actions, alors $Events(\mathcal{A})$ est l'union des ensembles $Events(a)$ (pour toutes les instances d'action $a \in \mathcal{A}$).

Définition 2 *$P = \langle \mathcal{A}, \tau \rangle$, où $\mathcal{A}$ est un ensemble fini d'instances d'actions $\{a_1, \ldots, a_n\}$ et τ est une fonction à valeurs réelles sur $Events(\mathcal{A})$, est un plan temporel pour le problème $\langle I, \mathcal{A}', G \rangle$ si*

(1) $\mathcal{A} \subseteq \mathcal{A}'$, et

(2) P vérifie les contraintes inhérentes et contradictoires sur $\mathcal{A}$;

et lorsque P est exécuté (c'est-à-dire que les fluents sont établis ou détruits aux instants donnés par τ) à partir de l'état initial I :

(3) pour toute $a_i \in \mathcal{A}$, chaque $f \in Cond(a_i)$ est vrai lorsqu'il est requis, et

(4) tous les fluents $g \in G$ sont vrais à la fin de l'exécution de P.

(5) P est robuste sous des changements infinitésimaux dans les temps de démarrage des actions.

Les événements sont instantanés, alors que les actions ont non seulement une durée mais celle-ci peut aussi être de longueur variable. Ainsi, un plan temporel P ne planifie pas directement ses instances d'action mais planifie tous les événements dans ses instances d'actions. La condition (5) de la définition 2 signifie que nous interdisons les plans qui exigent une synchronisation parfaite entre différentes actions. Cette condition peut être imposée dans PDDL2.1 [5]. Nous exigeons que dans tous les plans, les fluents soient établis strictement avant le début de l'intervalle sur lequel ils sont requis. La seule exception à cette règle est lorsqu'un fluent f est établi et requis par la même action a. Nous permettons la possibilité d'une parfaite synchronisation au sein d'une action, ce qui signifie que nous pouvons avoir $\tau(a \rightarrow f) = \tau(f \mid\!\rightarrow a)$. De même, les fluents ne peuvent être détruits qu'après la fin de l'intervalle sur lequel ils sont requis. La seule exception à cette regle est quand un fluent f est requis

et détruit par une action a, auquel cas on peut avoir $\tau(f \rightarrow| a) = \tau(a \rightarrow \neg f)$.

Définition 3 *Un problème de planification temporelle $\langle I, \mathcal{A}, G\rangle$ est positif s'il n'y a pas de fluents négatifs dans les conditions d'actions ni dans le but G.*

Dans cet article, nous ne considèrerons que des problèmes de planification temporelle positifs $\langle I, \mathcal{A}, G\rangle$. Il est bien connu que tout problème de planification peut être transformé en un problème positif équivalent en temps linéaire par l'introduction, pour chaque fluent positif f, d'un nouveau fluent pour remplacer les occurrences de $\neg f$ dans les conditions d'actions [6]. En supposant que tous les problèmes sont positifs, G et $Cond(a)$ (pour toute action a) sont composés de fluents positifs. Par convention, $Add(a)$ et $Del(a)$ sont aussi composés exclusivement de fluents positifs. L'état initial I, cependant, peut contenir des fluents négatifs.

3 Codage SAT pour la planification classique

Dans cette section, nous considérons la planification classique comme restriction de la planification temporelle décrite dans la section précédente. Un problème de planification positif est donc ici un triplet $\langle I, \mathcal{A}, G\rangle$ tel que toutes les actions de $\mathcal{A}$ sont instantanées (c'est-à-dire $\forall a \in \mathcal{A}, \forall e_1, e_2 \in Events(a), \tau(e_1) = \tau(e_2)$). De plus, nous considérons pour tout plan $P = \langle \mathcal{A}, \tau\rangle$, que τ est une fonction à valeurs entières de $Events(\mathcal{A})$ dans $N = \{1, \ldots, n\}$ l'ensemble ordonné des index d'étapes de P.
Les codages SAT dans les espaces de plans existants introduits par [15] se réfèrent tous à trois étapes indexées (pas nécessairement consécutives) du plan. Dans notre nouveau codage, nous allons nous référer seulement à deux étapes consécutives. Pour chaque action $a \in \mathcal{A}$ et chaque étape i, nous créons une variable propositionnelle a_i pour déterminer qu'une instance de a doit être planifiée à l'étape i. Pour chaque fluent $f \in \mathcal{F}$ et chaque étape i, nous créons une variable propositionnelle $open_{f,i}$ pour exprimer que f se maintient à l'étape précédente $i-1$ et doit être protégé au moins jusqu'à l'étape i.

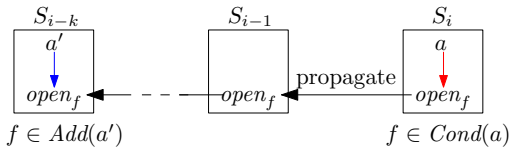


FIGURE 1 – Lien causal : l'action a' produit f à l'étape S_{i-k} pour l'action a qui nécessite f à l'étape S_i.

Dans la figure 1, la variable f est une *condition ouverte* à l'étape S_i, impliquant que $f \in I$ ou une action a' qui ajoute f est exécutée dans une étape précédente S_{i-k}. Les conditions ouvertes doivent être propagées vers l'arrière jusqu'à l'état initial ou une étape dans laquelle elles sont ajoutées par une action.
Dans la suite, nous donnons notre codage pour une longueur de plan fixée *length*. Une borne supérieure pour *length* est le nombre total d'états possibles, soit $2^{|\mathcal{F}|}$.

Conditions ouvertes Si une action a est exécutée dans une étape du plan, alors chaque condition de a doit être une condition ouverte à cette étape (c'est-à-dire qu'un lien causal est requis pour cette condition).

$$\bigwedge_{\mathbf{i} \in [1..length]} \bigwedge_{\mathbf{a} \in \mathcal{A}} \left(\mathbf{a_i} \Rightarrow \bigwedge_{\mathbf{f} \in Cond(\mathbf{a})} open_{\mathbf{f},\mathbf{i}} \right)$$

Dans la dernière étape du plan menant au but tous les fluents du but doivent être des conditions ouvertes ou ajoutées par des actions exécutées dans cette étape.

$$\bigwedge_{\mathbf{f} \in \mathbf{G}} \left(open_{\mathbf{f},length} \vee \bigvee_{\substack{\mathbf{a} \in \mathcal{A} \\ \mathbf{f} \in Add(\mathbf{a})}} \mathbf{a}_{length} \right)$$

Propagation et fermeture Aucune condition ne doit rester ouverte dans la première étape du plan si elle n'est pas fournie dans l'état initial.

$$\bigwedge_{\mathbf{f} \in \mathcal{F} \setminus \mathbf{I}} \neg open_{\mathbf{f},1}$$

Toute condition ouverte dans une étape doit soit rester ouverte soit être ajoutée (fermée) par une action à l'étape précédente.

$$\bigwedge_{\mathbf{i} \in [2..length]} \bigwedge_{\mathbf{f} \in \mathcal{F}} \left(open_{\mathbf{f},\mathbf{i}} \Rightarrow \left(open_{\mathbf{f},\mathbf{i}-1} \vee \bigvee_{\substack{\mathbf{a} \in \mathcal{A} \\ \mathbf{f} \in Add(\mathbf{a})}} \mathbf{a}_{\mathbf{i}-1} \right) \right)$$

Protection des conditions ouvertes Une condition ouverte dans une étape donnée ne peut pas être supprimée à l'étape précédente. Cela garantit de ne rompre aucun lien de causalité dans le plan.

$$\bigwedge_{\mathbf{i} \in [2..length]} \bigwedge_{\mathbf{f} \in \mathcal{F}} \left(open_{\mathbf{f},\mathbf{i}} \Rightarrow \bigwedge_{\substack{\mathbf{a} \in \mathcal{A} \\ \mathbf{f} \in Del(\mathbf{a})}} \neg \mathbf{a}_{\mathbf{i}-1} \right)$$

Prévention des interactions négatives Si une action supprime un fluent qui est nécessaire ou est ajouté par une autre action, alors ces deux actions ne peuvent pas être exécutées toutes les deux dans une même étape.

©AfIA 2018

$$\bigwedge_{\mathbf{i}\in[1..\mathbf{length}]} \bigwedge_{\mathbf{a}\in\mathcal{A}} \bigwedge_{\mathbf{f}\in(Add(\mathbf{a})\cup Cond(\mathbf{a}))} \bigwedge_{\substack{\mathbf{b}\in\mathcal{A}\\ \mathbf{a}\neq\mathbf{b}\wedge\mathbf{f}\in Del(\mathbf{b})}} (\neg\mathbf{a_i} \vee \neg\mathbf{b_i})$$

Dans la section suivante, nous étendons ce codage SAT pour la planification classique en un codage SMT pour la planification temporelle en temps continu.

4 Codage SMT pour la planification temporelle

Nous introduisons une adaptation SMT du codage SAT basé sur les conditions ouvertes que nous avons introduit dans la section précédente pour la planification classique. Ici, les actions ne sont plus instantanées et des contraintes sur les instants $\tau(e)$ auxquels se produisent des événements $e \in Events(\mathcal{A})$ doivent explicitement être ajoutées.

Considérons donc maintenant un problème de planification temporelle positif $\langle I, \mathcal{A}, G\rangle$. Pour chaque instance d'action a et chaque fluent $f \in Cond(a)$, nous introduisons deux variables $\tau_s(open_f)$ et $\tau_e(open_f)$ qui nous permettent de définir un intervalle temporel de protection de lien causal depuis l'état initial ou une étape contenant une action qui produit f vers une étape qui contient l'instance d'action a. Durant cet intervalle temporel, aucune action ne pourra détruire f.

Conditions ouvertes Si une action a est exécutée dans une étape du plan, alors chaque condition de a doit être une condition ouverte à cette étape (c'est-à-dire qu'un lien causal est requis pour cette condition). De plus, le début de l'intervalle sur lequel cette condition est requise se trouve dans l'intervalle de protection du lien causal $[\tau_s(open_f); \tau_e(open_f)]$.

$$\bigwedge_{\mathbf{i}\in[1..\mathbf{length}]} \bigwedge_{\mathbf{a}\in\mathcal{A}} \left(\mathbf{a_i} \Rightarrow \bigwedge_{\mathbf{f}\in Cond(\mathbf{a})} \begin{pmatrix} open_{\mathbf{f},\mathbf{i}} \\ \wedge\,(\tau(\mathbf{f} \mid\!\rightarrow \mathbf{a_i}) \geq \tau_s(open_{\mathbf{f},\mathbf{i}})) \\ \wedge\,(\tau(\mathbf{f} \mid\!\rightarrow \mathbf{a_i}) \leq \tau_e(open_{\mathbf{f},\mathbf{i}})) \end{pmatrix}\right)$$

Dans la dernière étape du plan menant au but tous les fluents du but doivent être des conditions ouvertes ou ajoutées par des actions exécutées dans cette étape. Dans le cas où une action a ajoute une condition ouverte f, le début de l'intervalle de protection du lien causal correspondant est fixé à l'instant où a produit f.

$$\bigwedge_{\mathbf{f}\in\mathbf{G}} \left(\begin{array}{l} open_{\mathbf{f},length} \\ \vee \displaystyle\bigvee_{\substack{\mathbf{a}\in\mathcal{A}\\ \mathbf{f}\in Add(\mathbf{a})}} \begin{pmatrix} \mathbf{a}_{length} \\ \wedge\,(\tau(\mathbf{a}_{length} \rightarrow \mathbf{f}) = \tau_s(open_{\mathbf{f},length})) \\ \wedge\,(\tau_{Goal} = \tau_e(open_{\mathbf{f},length})) \end{pmatrix} \end{array}\right)$$

Propagation et fermeture Aucune condition ne doit rester ouverte dans la première étape du plan si elle n'est pas fournie dans l'état initial.

$$\bigwedge_{\mathbf{f}\in\mathcal{F}\setminus\mathbf{I}} \neg open_{\mathbf{f},1}$$

Si une condition ouverte est fournie par l'état initial, alors le début de l'intervalle de protection du lien causal correspondant est fixé à l'instant initial τ_{Init}.

$$\bigwedge_{\mathbf{f}\in\mathbf{I}} (open_{\mathbf{f},1} \Rightarrow (\tau_{Init} = \tau_s(open_{\mathbf{f},1})))$$

Toute condition ouverte f dans une étape doit à l'étape précédente : (1) soit rester ouverte et dans ce cas l'intervalle de protection du lien causal correspondant reste le même pour ces deux étapes, (2) soit être ajoutée (fermée) par une instance d'action a et dans ce cas le début de l'intervalle de protection du lien causal est fixé à l'instant de production de f par a.

$$\bigwedge_{\mathbf{i}\in[2..length]} \bigwedge_{\mathbf{f}\in\mathcal{F}} \left(\begin{array}{l} open_{\mathbf{f},\mathbf{i}} \Rightarrow \\ \left(\begin{array}{l} \begin{pmatrix} open_{\mathbf{f},\mathbf{i}-1} \\ \wedge\,(\tau_s(open_{\mathbf{f},\mathbf{i}-1}) = \tau_s(open_{\mathbf{f},\mathbf{i}})) \\ \wedge\,(\tau_e(open_{\mathbf{f},\mathbf{i}-1}) = \tau_e(open_{\mathbf{f},\mathbf{i}})) \end{pmatrix} \\ \vee \displaystyle\bigvee_{\substack{\mathbf{a}\in\mathcal{A}\\ \mathbf{f}\in Add(\mathbf{a})}} (\mathbf{a_{i-1}} \wedge (\tau(\mathbf{a_{i-1}} \rightarrow \mathbf{f}) = \tau_s(open_{\mathbf{f},\mathbf{i}}))) \end{array}\right) \end{array}\right)$$

Protection des conditions ouvertes Une condition ouverte dans une étape donnée ne peut pas être supprimée à l'intérieur de l'intervalle de protection du lien causal correspondant $[\tau_s(open_f); \tau_e(open_f)]$. Cela garantit de ne rompre aucun lien de causalité dans le plan.

$$\bigwedge_{\substack{\mathbf{i}\in[1..length]\\ \mathbf{j}\in[1..length]}} \bigwedge_{\mathbf{f}\in\mathcal{F}} \bigwedge_{\substack{\mathbf{a}\in\mathcal{A}\\ \mathbf{f}\in Del(\mathbf{a})}} \left((open_{\mathbf{f},\mathbf{i}} \wedge \mathbf{a_j}) \Rightarrow \begin{pmatrix} (\tau(\mathbf{a_j} \rightarrow \neg\mathbf{f}) < \tau_s(open_{\mathbf{f},\mathbf{i}})) \\ \vee\,(\tau_e(open_{\mathbf{f},\mathbf{i}}) < \tau(\mathbf{a_j} \rightarrow \neg\mathbf{f})) \end{pmatrix}\right)$$

Prévention des interactions négatives Si une action b supprime, à un instant $\tau(b \rightarrow \neg f)$, un fluent f qui est ajouté par une autre action a à un instant $\tau(a \rightarrow f)$, alors ces deux instants sont différents.

$$\bigwedge_{\substack{\mathbf{i}\in[1..length]\\ \mathbf{j}\in[1..length]}} \bigwedge_{\mathbf{a}\in\mathcal{A}} \bigwedge_{\mathbf{f}\in Add(\mathbf{a})} \bigwedge_{\substack{\mathbf{b}\in\mathcal{A}\\ ((\mathbf{i}\neq\mathbf{j})\vee(\mathbf{a}\neq\mathbf{b}))\wedge\mathbf{f}\in Del(\mathbf{b})}} ((\mathbf{a_i} \wedge \mathbf{b_j}) \Rightarrow (\tau(\mathbf{a_i} \rightarrow \mathbf{f}) \neq \tau(\mathbf{b_j} \rightarrow \neg\mathbf{f})))$$

De même, si une action b supprime, à un instant $\tau(b \rightarrow \neg f)$, un fluent f qui est nécessaire à une autre action a sur un intervalle temporel $[\tau(f \mid\rightarrow a), \tau(f \rightarrow\mid a)]$, alors cet instant ne peut être contenu dans cet intervalle.

$$\bigwedge_{\substack{\mathbf{i}\in[1..length]\\ \mathbf{j}\in[1..length]}} \bigwedge_{\mathbf{a}\in\mathcal{A}} \bigwedge_{\mathbf{f}\in Cond(\mathbf{a})} \bigwedge_{\substack{\mathbf{b}\in\mathcal{A}\\ ((\mathbf{i}\neq\mathbf{j})\vee(\mathbf{a}\neq\mathbf{b}))\wedge\mathbf{f}\in Del(\mathbf{b})}} \left((\mathbf{a_i} \wedge \mathbf{b_j}) \Rightarrow \begin{pmatrix} (\tau(\mathbf{f} \rightarrow\mid \mathbf{a_i}) < \tau(\mathbf{b_j} \rightarrow \neg\mathbf{f})) \\ \vee (\tau(\mathbf{b_j} \rightarrow \neg\mathbf{f}) < \tau(\mathbf{f} \mid\rightarrow \mathbf{a_i})) \end{pmatrix} \right)$$

Bornes du plan temporel Enfin, nous devons ajouter une contrainte pour maintenir le plan dans un intervalle de temps borné par l'étape initiale qui produit l'état initial I et l'étape finale qui nécessite tous les fluents du but G.

$$\bigwedge_{\mathbf{i}\in[1..length]} \bigwedge_{\mathbf{a}\in\mathcal{A}} \left(\mathbf{a_i} \Rightarrow \begin{pmatrix} \bigwedge_{\mathbf{f}\in Cond(\mathbf{a})} \begin{pmatrix} (\tau_{Init} \leq \tau(\mathbf{f} \mid\rightarrow \mathbf{a_i})) \\ \wedge (\tau_{Goal} \geq \tau(\mathbf{f} \rightarrow\mid \mathbf{a_i})) \end{pmatrix} \\ \wedge \bigwedge_{\mathbf{f}\in Add(\mathbf{a})} \begin{pmatrix} (\tau_{Init} \leq \tau(\mathbf{a_i} \rightarrow \mathbf{f})) \\ \wedge (\tau_{Goal} \geq \tau(\mathbf{a_i} \rightarrow \mathbf{f})) \end{pmatrix} \\ \wedge \bigwedge_{\mathbf{f}\in Del(\mathbf{a})} \begin{pmatrix} (\tau_{Init} \leq \tau(\mathbf{a_i} \rightarrow \neg\mathbf{f})) \\ \wedge (\tau_{Goal} \geq \tau(\mathbf{a_i} \rightarrow \neg\mathbf{f})) \end{pmatrix} \end{pmatrix} \right)$$

5 Conclusion et perspectives

Après avoir introduit notre modèle de problèmes de planification temporelle basé sur les aspects temporels de PDDL2.1, nous avons présenté un nouveau codage SAT dans un espace de plans basé sur des conditions ouvertes (liens causaux) pour la planification classique. Nous avons proposé une adaptation SMT de ce codage SAT pour la planification temporelle en temps continu en introduisant des atomes de QF-RDL. Il est maintenant nécessaire de tester ce codage sur les problèmes de référence des compétitions internationales de planification (IPC) et de le comparer avec les autres approches SMT existantes.
Une autre voie à explorer serait d'intégrer cette modélisation des conditions ouvertes dans notre planificateur TLP-GP [17, 16] qui couple un algorithme de recherche dans un graphe de planification temporel à la résolution de contraintes SMT.

Références

[1] Martin C. Cooper, Frederic Maris, and Pierre Régnier. Monotone temporal planning : Tractability, extensions and applications. *J. Artif. Intell. Res. (JAIR)*, 50 :447–485, 2014.

[2] Minh Binh Do and Subbarao Kambhampati. Planning as constraint satisfaction : Solving the planning graph by compiling it into CSP. *Artif. Intell.*, 132(2) :151–182, 2001.

[3] M. Ernst, T. Millstein, and D.S. Weld. Automatic SAT-compilation of planning problems. In *Proc. IJCAI-97*, 1997.

[4] Maria Fox and Derek Long. PDDL2.1 : an extension to PDDL for expressing temporal planning domains. *J. Artif. Intell. Res. (JAIR)*, 20 :61–124, 2003.

[5] Maria Fox, Derek Long, and Keith Halsey. An investigation into the expressive power of PDDL2.1. In Ramon López de Mántaras and Lorenza Saitta, editors, *Proceedings of the 16th Eureopean Conference on Artificial Intelligence, ECAI'2004, including Prestigious Applicants of Intelligent Systems, PAIS 2004, Valencia, Spain, August 22-27, 2004*, pages 328–342. IOS Press, 2004.

[6] Malik Ghallab, Dana S. Nau, and Paolo Traverso. *Automated planning - theory and practice.* Elsevier, 2004.

[7] H. Kautz. Satplan'04 : Planning as satisfiability. In *Abstracts of the 4th International Planning Competition, IPC-04*, 2004.

[8] H. Kautz and B. Selman. Planning as satisfiability. In *ECAI-92*, pages 359–363, 1992.

[9] H. Kautz and B. Selman. Pushing the envelope : Planning, propositional logic and stochastic search. In *Proc. AAAI-96*, pages 1194–1201, 1996.

[10] H. Kautz and B. Selman. BLACKBOX : A new approach to the application of theorem proving to problem solving. In *Proc. of AIPS-98 Workshop on Planning as Combinatorial Search*, 1998.

[11] H. Kautz and B. Selman. Unifying SAT-based and Graph-based planning. In *Proc. IJCAI-99*, pages 318–325, 1999.

[12] H. Kautz, B. Selman, and J. Hoffmann. Satplan'06 : Planning as satisfiability. In *Abstracts of the 5th International Planning Competition, IPC-06*, 2006.

[13] A. Mali and S. Kambhampati. Refinement-based planning as satisfiability. In *Proc. Workshop planning as combinatorial search, AIPS-98*, 1998.

[14] A. Mali and S. Kambhampati. On the utility of plan-space (causal) encodings. In *Proc. AAAI-99*, pages 557–563, 1999.

[15] Amol Dattatraya Mali and Subbarao Kambhampati. On the utility of plan-space (causal) encodings. In Jim Hendler and Devika Subramanian, editors, *Proceedings of the Sixteenth National Conference on Artificial Intelligence and Eleventh Conference on Innovative Applications of Artificial Intelligence, July 18-22, 1999, Orlando, Florida, USA.*, pages 557–563. AAAI Press / The MIT Press, 1999.

©AfIA 2018

[16] Frédéric Maris. *Planification SAT et planification temporellement expressive : les systèmes TSP et TLP-GP*. Thèse de doctorat, Université Paul Sabatier, Toulouse, France, 2009.

[17] Frederic Maris and Pierre Régnier. TLP-GP : new results on temporally-expressive planning benchmarks. In *(ICTAI 2008), Vol. 1*, pages 507–514. IEEE Computer Society, 2008.

[18] J. Rintanen. Symmetry reduction for sat representations of transition systems. In *Proc. ICAPS-03*, 2003.

[19] J. Rintanen, K. Heljanko, and Niemelä. Parallel encodings of classical planning as satisfiability. In *Proc. European Conference on Logics in Artificial Intelligence, JELIA-04*, 2004.

[20] J. Rintanen, K. Heljanko, and Niemelä. Planning as satisfiability : parallel plans and algorithms for plan search. *Artificial Intelligence*, 170(1213) :1031–1080, 2006.

[21] Jussi Rintanen. Discretization of temporal models with application to planning with SMT. In Blai Bonet and Sven Koenig, editors, *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA.*, pages 3349–3355. AAAI Press, 2015.

[22] Ji-Ae Shin and Ernest Davis. Processes and continuous change in a sat-based planner. *Artif. Intell.*, 166(1-2) :194–253, 2005.

[23] Steven A. Wolfman and Daniel S. Weld. The LPSAT engine & its application to resource planning. In Thomas Dean, editor, *Proc. 16th International Joint Conference on Artificial Intelligence, IJCAI 99, Stockholm, Sweden, 1999*, pages 310–317. Morgan Kaufmann, 1999.

L'intelligence artificielle au service de la maintenance prévisionnelle

Juliette Mattioli[1] Pierre-Olivier Robic[2] Thomas Redellet[1]

[1] Thales Research & Technology, Campus Polytechnique 1 av. A. Fresnel, 91767 Palaiseau Cedex
[2] Thales Global Services, 19 av. M. Saulnier, 78140 Vélizy Villacoublay.

{juliette.mattioli, pierre-olivier.robic, thomas.reydellet}@thalesgroup.com

Résumé

Avec la transformation numérique du Groupe Thales, l'intelligence artificielle (IA) qu'elle soit orientée par les modèles (IA symbolique) ou basée sur les données (IA connexioniste) devient un élément clé pour la maintenance prévisionnelle. L'objectif de cet article est de dresser un rapide panorama des différentes technologies d'IA mise en oeuvre dans Thales pour outiller les activités de maintenance prévisionnelle.

Mots Clef

Maintenance prévisionnelle, IA symbolique, IA connexioniste, diagnostic, pronostic, planification, ordonnancement.

Abstract

With the Thales Group digital transformation, model-driven AI (Arificial Intelligence) and data-driven AI are becoming a cornerstone to predictive maintenance.The aim of this paper is to give a brief overview of various AI techniques which are use to address predictive maintenance issues.

Keywords

Predictive maintenance, Symbolic AI, Connexionist AI, diagnostic, prognostic, planning, scheduling.

1 Maintenance prévisionnelle : un levier de compétitivité

Selon la définition de l'AFNOR [1], la maintenance couvre *l'ensemble des actions du cycle de vie d'un système, destinées à le maintenir ou à le rétablir dans un état dans lequel il peut accomplir la fonction requise*. Les activités induites sont aujourd'hui de deux types : corrective ou préventive. La maintenance corrective est programmée suite à une défaillance [1]. Cependant, ces activités n'intègrent pas d'aspect préventif. Fondée sur l'adage "*mieux vaut prévenir que guérir*", la maintenance préventive repose sur un plan déterminé a priori. La fréquence des ces opérations s'appuie en général sur des options de garantie, des instructions du fournisseur, la réglementation ou l'expérience.

1. Une défaillance correspond à une cessation de l'aptitude d'une entité à accomplir une ou plusieurs fonctions requises.

Avec la transformation numérique du Groupe Thales, une nouvelle forme de maintenance devient possible grâce aux technologies de l'intelligence artificielle (IA) : la maintenance prévisionnelle [2]. Cette dernière est réalisée suivant les prévisions extrapolées de l'analyse et de l'évaluation de paramètres significatifs de dégradation du système. Ce type de maintenance vise à prédire l'évolution de ces paramètres pour estimer le temps de fonctionnement restant avant défaillance (pronostic). Cette approche est souvent qualifiée de maintenance prédictive, car reposant sur les évolutions de paramètres surveillés dans le futur, et de maintenance proactive qui cherche à déterminer les causes à l'origine de ces dégradations et de ces défaillances.

Thales propose ainsi une offre complète de maintenance pour les systèmes critiques, offre de services outillée allant de la réparation et la fourniture de pièces de rechange à la maintenance prévisionnelle, préventive et corrective. En effet, tout système en fonctionnement qu'il soit critique ou non, manifeste des signes, visibles ou non, de dégradation annoncant une future défaillance. La détection d'anomalies ou de changements préfigurent alors à court terme d'une panne à venir. Ainsi la maintenance prévisionnelle permet par exemple, dans le cadre des services offerts aux exploitants des réseaux de métro, de lutter contre les retards en minimisant les interruptions causées par des défaillances du matériel. Elle permet également de mettre à mal les retards en chaine, à savoir ces retards qui s'accumulent en cascade suite à un incident unique. De tels retards peuvent carrément paralyser une ligne, voire le réseau tout entier.

C'est pourquoi, comme nous le verrons dans la suite de ce papier qui se veut être une synthèse des approches aujourd'hui déployées, Thales propose des solutions innovantes reposant sur des techniques d'IA, pour chacune des étapes de maintenance.

En effet, l'IA que nous définissons comme "*des systèmes d'information et des applications pouvant percevoir un environnement, comprendre et décider en fonction du contexte et des usages, tout en apprenant au fil de l'eau*", qu'elle soit dirigée par les données (IA connexioniste) ou basée sur des modèles et la connaissance (IA symbolique) est devenue pour Thales, incontournable pour la surveillance de l'état de santé du système, la planification et

 ©AfIA 2018

la gestion des opérations de maintenance.

2 Diagnostic

L'idée de la maintenance prévisionnelle est de projeter l'état actuel d'un système dans le futur, afin d'en estimer sa durée de vie restante (i.e., durée de fonctionnement avant défaillance) appelée communément RUL (remaining useful life), ou de calculer sa probabilité pour qu'il fonctionne durant un certain temps ainsi que le risque d'existence ou d'apparition ultérieure d'un ou de plusieurs modes de défaillance. Plusieurs techniques d'IA permettent de diagnostiquer les fautes survenant dans un système qui provoquent des défaillances.

Certaines reposent sur la connaissance disponible du fonctionnement normal (comportement nominal) ou anormal (comportement en présence de faute) du système [3] pour établir des relations causales entre les causes et les effets observés ou les symptômes, les défaillances et les fautes. Ces méthodes exploitent la connaissance des experts sur le système, une connaissance acquise durant les opérations de maintenance ou/et d'un historique des dysfonctionnements du système. Elles peuvent être capturées par le biais par exemple d'**ontologies** [4] et/ou se basent sur l'existence d'un modèle formel du système qui impliquent une connaissance approfondie de son fonctionnement sous la forme de modèle mathématique et/ou logique. Ainsi pour certains systèmes critiques, Thales a mis en oeuvre des méthodes de diagnostic de défaillances [5] modélisations fonctionnelles et matérielles reposant sur des outils industriels comme l'AMDEC (Analyse des Modes de Défaillance, de leur Effet et de leur Criticité) des les arbres de défaillances ou par le biais de modélisation physique consistant à comparer un modèle représentatif du bon fonctionnement du système avec les mesures observées [6]. Tout écart est alors synonyme d'un dysfonctionnement, c'est-à-dire la présence d'un ou plusieurs défauts. A partir de ces études, Thales a développé des **systèmes d'inférence** pour l'aide au diagnostic.

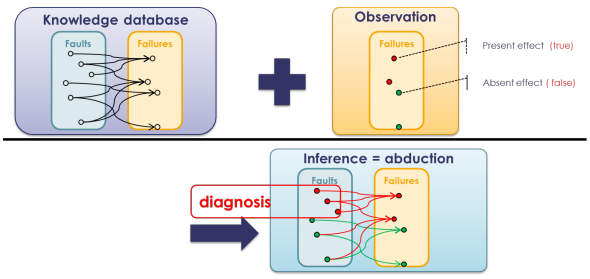


FIGURE 1 – Principe de raisonnenement adbuctif pour l'aide au diagnostic

La propriété principale de cette approche est de pouvoir représenter et restituer les connaissances acquises par les opérateurs de maintenance. Rappelons qu'un tel système comporte une base de connaissances composée d'une base de règles qui modélise la connaissance et d'une base de faits qui contient les informations concernant le système observé et un moteur d'inférence capable de raisonner à partir des informations contenues dans la base de connaissances et de faire des déductions.

Quand cette connaissance n'est pas accessible, d'autres approches consistent à traiter le système comme étant une boite noire. Les mesures (entrées-sorties) constituent souvent la plus forte et la plus sûre source d'informations pour comprendre les phénomènes de dégradation. L'objectif est d'associer ces observations à des états de fonctionnement connus du système. Ces approches à base d'IA basées sur les données, visent à apprendre (par les exemples) le phénomène de dégradation et à capturer les relations subtiles entre données, même si ces relations sont inconnues ou difficiles à décrire. Ainsi, Thales propose une combinaison d'approches [7] combinant des approches statistiques, l'**analyse en composante principale** pour identifier les interactions entre les différents composants du système et réduire la complexité, les **chaînes de Markov cachées** pour retrouver l'ensemble des états les plus probables à partir d'une séquence de sortie, les **réseaux bayésiens** pour permettre la corrélation d'événements et la détection de signaux faibles, les **arbres de décision** ou **réseaux de neurones** comme le **deep learning**. Ces approches exploitent les données de surveillance (issues de capteurs) ou d'interventions de maintenance, données traitées afin d'extraire des caractéristiques reflétant le comportement du système et sa dégradation. Celles-ci sont ensuite utilisées pour apprendre des modèles de prédiction des états du système, et calculer une estimation du RUL.

3 Prognostic

Le diagnostic traite de l'identification et de la quantification des dommages qui se sont produits, alors que le pronostic porte sur la prédiction non seulement des possibles états futurs dans lesquels un système peut évoluer mais aussi des dommages qui n'ont pas encore eu lieu. Sa fonction principale est de renforcer le choix des actions à entreprendre pour proposer des recommandations pour l'ordonnancement des tâches de maintenance. Il permet d'aborder la maintenance prévisionnelle tout en réduisant ses coûts directs. Les méthodes pour le pronostic dépendent comme celles pour le diagnostic, de la connaissance et des données que l'on a sur le système à maintenir. Elles sont soit fondées sur un modèle physique, soit guidées par les données, soit basées sur l'expérience. Cette dernière s'appuie principalement sur des approches fiabilistes, sur des modélisations de processus stochastiques ou sur des systèmes experts.

Les approches à base de modèles, développées dans Thales, consistent à simuler le modèle de comportement du système pour caractériser la dynamique des paramètres reflétant la dégradation. Par exemple, le suivi de la dégradation et l'obtention de sa dynamique peuvent être modélisé par un filtre de Kalman, par un filtrage particulaire ou par

des chaînes de Markov cachées [3].
Les méthodes guidées par les données sont, avec le renouveau de l'IA, en pleines expansions. Elles exploitent des mesures représentant l'état actuel et/ou l'état passé. Elles utilisent des techniques d'**apprentissage** ou des techniques d'**analyse de tendance** pour construire un modèle à partir de ces données pour ensuite extrapoler l'évolution du système. Cependant, elles supposent que la probabilité d'évolution est proche des conditions de fonctionnement utilisées lors de l'apprentissage du modèle. Les **réseaux de neurones adaptatifs**, les **machines à vecteurs de support** sont des approches qui ont révélées leur efficacité pour la prévision de séries temporelles. Leur intérêt réside dans leurs aptitudes à apprendre et à capturer les relations potentielles entre les données même si le comportement est complexe. Ces modèles sont appris par un ensemble d'échantillons (entrées/sorties) directement extrait de la séquence d'observations. Enfin, la **fusion d'informations** [8] permet la corrélation d'un indicateur de santé avec les conditions opératoires permet de déterminer un RUL en considérant à la fois l'état de santé courant du système et son utilisation. Cependant, ces approches sont difficiles à mettre en oeuvre pour les équipements complexes, leur réponse est souvent entachée d'incertitudes de modélisation ce qui induit un pronostic moins robuste. Pour avoir un modèle robuste permettant de prendre en compte les imperfections et les incertitudes, la **logique floue** et les **fonctions de croyances** sont très utilisées [9]. Combinée avec les méthodes citées ci-dessus, elles offrent un compromis raisonnable entre les modèles analytiques rigoureux, les approches reposant sur des connaissances métier et les approches guidées par les données.

4 Planification et Ordonnancement

Une gestion intelligente du matériel permet de faire évoluer la maintenance du réactif vers le prédictif, avec le potentiel de réduire les coûts tout en améliorant la disponibilité du matériel. Dans certains cas, Thales a mis en oeuvre des approches de type **raisonnement par cas**, pour la phase préscriptive et proposer ainsi des actions correctives.
De plus, dans toute opération de maintenance préventive, il s'agit d'établir un planning prévisionnel des tâches sous-jacentes incluant les ressources humaines et matérielles impliquées. Ainsi, les dates de début au plus tôt, les durées et les dépendances des tâches sont connues a priori. Concernant les tâches de maintenance corrective, leur durée est basée sur une estimation et dépend d'un bon diagnostic. Or, un tel planning est généralement figé et peu ouvert aux changements et donc à des actions correctives qui sont par essence non planifiées.

Pour construire un tel ordonnancement mixant maintenance préventive et corrective, Thales utilise une combinaison de la **programmation par contraintes** avec des approches multicritères [10, 11] pour affecter des tâches aux opérateurs sous contrainte de compétence, en minimisant la date de fin de l'ordonnancement et les pénalités de retards [2]. Une approche outillée d'aide à la **décision multicritère** [12] a été mise en place (cf. figure 2), en particulier pour modéliser par le biais d'une fonction multicritère intégrant des dimensions de natures variées comme les arrêts de production, le retard dans l'exécution des tâches, les pénalités et d'autre part, des coûts de maintenance acceptables. De plus, avec la maintenance prévisonnelle, un réordonnancement est à envisager, en tenant compte du RUL. L'évolution du planning devient alors dynamique suite aux dysfonctionnements du système de production [13]. Enfin, la prise en compte de la maintenance dès la conception permet de concevoir des outils d'aide à la décision pour une optimisation de la gestion du système en mode dégradé comme le montre [5].

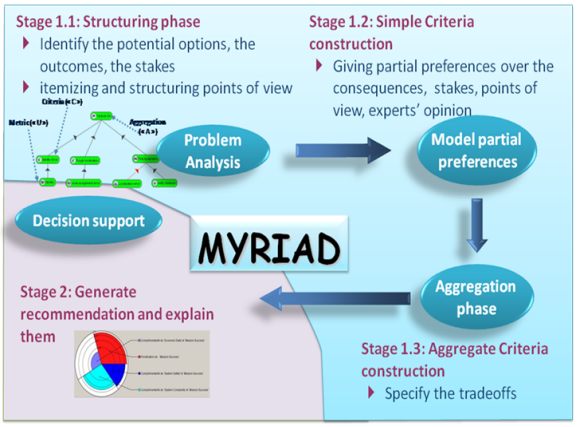


FIGURE 2 – Méthodologie Myriad développée par Thales pour construire une fonction multicritère

5 Conclusion

Une politique de maintenance prévisionnelle efficace doit conduire à une réduction des pertes engendrées par des indisponibilités de production. Son efficacité est fortement liée à sa faculté de caractériser l'état de santé courant (diagnostic) et futur (pronostic) du système. Pour cela, il est nécessaire de pouvoir générer des données permettant de rendre compte de la dynamique de la dégradation du système, de modéliser ou d'apprendre ses modes de dégradation, d'estimer son état de santé et de prédire sa durée de vie résiduelle. Les techniques d'IA qu'elles soient symboliques ou connexionnistes deviennent incontournables pour les activités de maintenance de Thales. Cependant, comme le montre la figure 3, il s'agit d'une combinaison d'approches qui dépendent de la connaissance disponible du fonctionnement nominal et du type de fautes mais aussi des données d'observation de l'état de santé du système.
Les services de maintenance de Thales exploite les données générées par les équipements. Ces données sont surveillées et comparées systématiquement aux profils de performance optimaux.
Lorsque des données envoyées par un équipement, par exemple un moteur d'aiguillage pour le cas du transport

©AfIA 2018

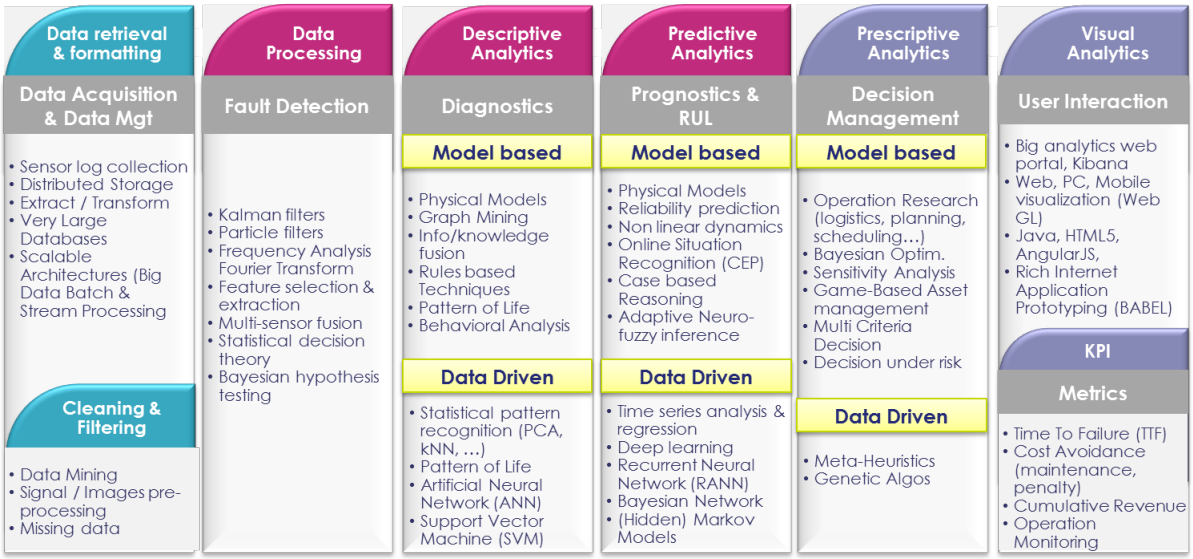


FIGURE 3 – Exemple d'approches mises en oeuvre dans Thales pour la maintenance

urbain, dévient de la norme, des alarmes sont automatiquement déclenchées et des recommandations précalculées sont proposées concernant les actions à prendre. Ce système est capable aussi de comprendre les signes avant-coureurs d'un problème.
L'avantage est que la précision des prédictions augmente au fil des données accumulées, notamment au travers des algorithmes d'apprentissage.

Références

[1] Norme AFNOR. Terminologie de la maintenance. *NF-EN*, 13306 :X60–319, 2001.

[2] Simon Fossier and Pierre-Olivier Robic. Maintenance of complex systems : From preventive to predictive. In *Live Maintenance (ICOLIM), 2017 12th International Conference on*, pages 1–6. IEEE, 2017.

[3] Camille Baysse. *Analyse et optimisation de la fiabilité d'un équipement opto-électrique équipé de HUMS*. PhD thesis, Université Sciences et Technologies-Bordeaux I, 2013.

[4] Luis Palacios, Gaëlle Lortal, Claire Laudy, Christian Sannino, Ludovic Simon, Giuseppe Fusco, Yue Ma, and Chantal Reynaud. Avionics maintenance ontology building for failure diagnosis support. In *Proceedings of the International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, pages 204–209. SCITEPRESS-Science and Technology Publications, Lda, 2016.

[5] J. Sprauel, C. Sannino, and C. Seguin. Techniques d'aide à la décision appliquées à la maintenance d'un avion de type business jet decision aiding techniques applied to the maintenance of a business jet. In *Congrès Lambda Mu 19*, 2014.

[6] Camille Baysse, Didier Bihannic, Anne Gégout-Petit, Michel Prenat, and Jérôme Saracco. Hidden markov model for the detection of a degraded operating mode of optronic equipment. *arXiv preprint arXiv :1212.2358*, 2012.

[7] Michel Giraudeau, Emilie Ronfaut, and D Bihannic. Apport des hums sur les performances sdf et soutien des équipements électroniques embarqués. *7A-Maintenance-diagnostic/pronostic*, 2014.

[8] Claire Laudy and Jean-Gabriel Ganascia. Introducing semantic knowledge in high-level fusion. In *Military Communications Conference, 2009. MILCOM 2009. IEEE*, pages 1–7. IEEE, 2009.

[9] Simon Fossier, Claire Laudy, and Frédéric Pichon. Managing uncertainty in conceptual graph-based soft information fusion. In *Information Fusion (FUSION), 2013 16th International Conference on*, pages 930–937. IEEE, 2013.

[10] Fabien Le Huédé, Michel Grabisch, Christophe Labreuche, and Pierre Savéant. Integration and propagation of a multi-criteria decision making model in constraint programming. *Journal of Heuristics*, 12(4-5) :329–346, 2006.

[11] Fabien Le Huédé, Michel Grabisch, Christophe Labreuche, and Pierre Savéant. Mcs : A new algorithm for multicriteria optimisation in constraint programming. *Annals of Operations Research*, 147(1) :143–174, 2006.

[12] Michel Grabisch and Christophe Labreuche. A decade of application of the choquet and sugeno integrals in multi-criteria decision aid. *Annals of Operations Research*, 175(1) :247–286, 2010.

[13] Christian Sannino, Jonathan Sprauel, and Christel Seguin. Modèles dysfonctionnels pour la gestion de la qualité de service des systèmes critiques. *3R-Recherche et Industrie 1*, 2016.

Réseaux de Neurones Récurrents Multi-tâches pour l'Analyse Automatique d'Arguments

Jean-Christophe Mensonides[1] Sébastien Harispe[1]
Jacky Montmain[1] Véronique Thireau[2]

[1] LGI2P, IMT Mines Ales, Univ Montpellier, Ales, France
[2] Université de Nîmes, CHROME, Rue du Dr Georges Salan, Nîmes, France

jean-christophe.mensonides@mines-ales.fr

Résumé

Dans cet article nous proposons une méthode d'extraction et d'analyse automatique d'arguments à partir de textes bruts, en nous affranchissant de l'utilisation de caractéristiques manuellement définies par des experts. Nous présentons un modèle multi-tâches faisant appel à des techniques d'apprentissage profond, composé de plusieurs couches de réseaux de neurones récurrents. Plus particulièrement, nous tirons parti de paramètres entraînés sur des tâches simples, comme l'étiquetage morpho-syntaxique ou le chunking, afin d'obtenir un modèle capable de traiter des tâches plus complexes nécessitant une compréhension fine du texte.

Mots Clef

Traitement Automatique du Langage Naturel, Extraction d'arguments, Réseaux de neurones récurrents, Apprentissage profond.

Abstract

In this article we propose a method performing automatic extraction and analysis of arguments from raw texts, without using handcrafted features. We introduce a multi-task deep learning model stacking several layers of recurrent neural networks. Specifically, we make use of weight parameters trained on simple tasks, such as Part-Of-Speech tagging and chunking, in order to obtain a model able to handle more complex tasks that require a detailed understanding of the text.

Keywords

Natural Language Processing, Argument mining, Recurrent neural networks, Deep learning.

1 Introduction

L'argumentation est un ensemble de techniques visant à faire adhérer un interlocuteur à un point de vue qui lui est présenté, en construisant un raisonnement à base d'arguments. Bien que l'étude de l'argumentation soit un champ étudié depuis longtemps dans des domaines tels que la philosophie ou la linguistique, l'extraction et l'analyse automatique d'arguments au sein de corpus textuels (aussi appelé *argument mining*) forment des axes de recherche relativement nouveaux. Un système d'argument mining a pour objectif la génération automatique d'un graphe d'arguments à partir de textes non structurés, et peut généralement être divisé en une séquence d'étapes comportant notamment la détection d'arguments et la modélisation des liens unissant ces derniers [1]. Nous nous limitons à une étude de la micro-structure argumentative, consistant à analyser la manière dont différents composants argumentatifs interragissent entre eux au sein d'un même texte.

De manière plus spécifique, Stab et Gurevych [2] ont proposé le corpus Argument Annotated Essays (version 2), contenant 402 dissertations extraites de essayforum.com. La structure argumentative de chaque dissertation a été manuellement annotée suivant un modèle de graphe orienté acyclique connexe, dans lequel les noeuds représentent des composants argumentatifs et les arcs des liens entre ces derniers. Le schéma d'annotation utilisé permet de distinguer trois types de composants argumentatifs : (i) les conclusions majeures, reflétant le point de vue global de l'auteur sur le sujet disserté, (ii) les conclusions intermédiaires, représentant des affirmations qui ne pourraient être acceptées sans justifications complémentaires, et (iii) les prémisses, servant de justifications aux conclusions intermédiaires avancées. Les arcs du graphe sont porteurs d'une étiquette "support" ou "attaque" selon que le composant argumentatif source corrobore ou réfute la cible. Les arcs ne peuvent exister que a) d'une prémisse vers une autre prémisse, b) d'une prémisse vers une conclusion (majeure ou intermédiaire), et c) d'une conclusion intermédiaire vers une autre conclusion (majeure ou intermédiaire).

Afin d'obtenir automatiquement un graphe synthétisant la structure argumentative d'une dissertation, Stab et Gurevych [2] ont proposé une chaîne de traitement constituée de quatre étapes : (1) Délimitation des frontières des composants argumentatifs, (2) Détermination du type

 ©AfIA 2018

de chaque composant argumentatif, (3) Détermination de l'existence d'un arc entre chaque paire ordonnée de composants argumentatifs, et (4) Etiquetage des arcs existants comme relation de support ou d'attaque.

Dans cet article, nous nous concentrons sur l'étude des tâches (1) et (2), en cherchant à nous affranchir de l'utilisation de caractéristiques définies manuellement par des experts. La section 2 présente un panorama des travaux antérieurs réalisés sur des tâches similaires à (1) et (2). La section 3 décrit le modèle que nous avons mis en place pour traiter les deux tâches évoquées ci-dessus. La section 4 présente les modalités d'entraînement du modèle. La section 5 est consacrée aux expérimentations que nous avons menées et aux résultats obtenus. La section 6 propose des directions et perspectives pour nos futures recherches.

2 Travaux antérieurs

La détection de composants argumentatifs consiste à déterminer les frontières séparant les unités textuelles porteuses d'arguments du reste du texte. Cette tâche est généralement considérée comme un problème de segmentation de texte supervisée au niveau du mot. Les modèles exploitant l'aspect séquentiel des mots, inhérent à la construction d'une argumentation convaincante, sont particulièrement adaptés et utilisés : Madnani et al (2012) utilisent un Conditional Random Field (CRF) afin d'identifier des segments non argumentatifs au sein de dissertations [3], Levy et al (2014) identifient les frontières d'unités textuelles représentant des conclusions supportant ou attaquant le sujet débattu dans des fils de discussions issus de Wikipedia [4], Ajjour et al (2017) utilisent des réseaux de neurones récurrents de type Long Short-Term Memory (LSTM) afin d'extraire des arguments issus de dissertations, d'éditoriaux et de commentaires générés par des internautes [5], Goudas et al (2014) identifient des phrases contenant des arguments avant de déterminer précisément leurs frontières au sein de médias sociaux à l'aide d'un CRF [6], Sardianos et al (2015) déterminent les limites de composants argumentatifs au sein d'articles de presse à l'aide d'un CRF [7], Stab et Gurevych (2017) utilisent un CRF afin d'isoler les composants argumentatifs au sein de dissertations [2], Eger et al (2017) ont recouru à des techniques d'apprentissage profond [8].

La tâche consistant à déterminer le type d'un composant argumentatif (prémisse, conclusion, etc.) a souvent été traité comme un problème de classification de texte supervisée. Eckle-Kohler et al (2015) distinguent des prémisses et des conclusions au sein d'articles de presse à l'aide de Naive Bayes, Random Forest et Support Vector Machine (SVM) [9], Park et Cardie (2014) utilisent un SVM pour déterminer à quel point des affirmations sont justifiées au sein de commentaires d'internautes relatifs à de nouveaux projets de législation [10], Stab et Gurevych (2017) classifient des composants argumentatifs en prémisses, conclusions intermédiaires et conclusions majeures dans des dissertations en utilisant un SVM [2], Persing et Ng (2016) utilisent un classifieur d'entropie maximale afin de déterminer le type de composants argumentatifs [11], Potash et al (2016) utilisent des réseaux de neurones récurrents dits "séquence à séquence" dans l'objectif d'inférer le type de composants argumentatifs [12].

L'étude de modèles multi-tâches, capables de traiter plusieurs problèmes différents en partageant un sous-ensemble commun de paramètres, a fait l'objet d'un engouement récent au sein de la communauté du traitement automatique du langage. Ce type de modèles est bio-inspiré : un être humain est capable de réaliser une multitude de tâches différentes et peut exploiter, quand cela est nécessaire, son savoir-faire acquis concernant la résolution d'un type de problème pour apprendre plus vite à résoudre d'autres types de problèmes. Ruder (2017) énonce les raisons pour lesquelles ce type de modèle est efficace d'un point de vue apprentissage automatique [13] : l'utilisation de plusieurs corpus différents induit une augmentation implicite du nombre d'exemples disponibles pendant la phase d'entraînement. De plus, le modèle doit rechercher des caractéristiques utiles pour l'ensemble des tâches à traiter, ce qui limite la modélisation du bruit dans les données et permet une meilleur généralisation.

Søgaard et Goldberg (2016) montrent qu'induire de la connaissance a priori dans un modèle multi-tâches en hiérarchisant l'ordre des tâches à apprendre permet d'obtenir de meilleurs performances [14]. Yang et al (2016) ont montré qu'entraîner un modèle multi-tâches et multi-langues permettait d'améliorer les performances sur des problèmes où les données ne sont que partiellement annotées [15], Hashimoto et al (2017) obtiennent des résultats compétitifs sur la majorité des tâches d'un même modèle [16]. Le bénéfice d'un modèle multi-tâches n'est cependant pas garanti, et dépend notamment de la distribution des données relatives aux différents problèmes traités (Mou et al (2016) [17], Alonso et Plank (2017) [18], Bingel et Søgaard (2017) [19]).

3 Modèle proposé

Nous proposons un modèle ayant pour objectif 1) de déterminer les frontières de composants argumentatifs présents dans un ensemble de dissertations et 2) de déterminer le type de chaque composant argumentatif dans lesdites dissertations. Nous nous inspirons du travail de Hashimoto et al [16] et optons pour un modèle multi-tâches s'affranchissant de la définition de caractéristiques manuellement définies. Plus particulièrement, nous utilisons des techniques issues de l'apprentissage profond et entraînons un modèle capable d'effectuer de l'étiquetage morpho-syntaxique (EMS), du chunking, de la détection de limites de composants argumentatifs et de la classification de com-

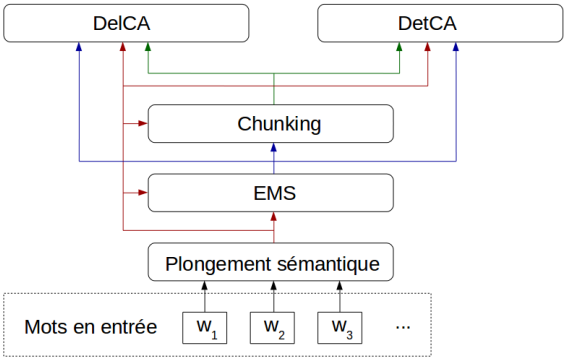


FIGURE 1 – Aperçu de l'architecture utilisée couche par couche. EMS, DelCA et DetCA sont respectivement des acronymes pour Etiquetage Morpho-Syntaxique, Délimitation des Composants Argumentatifs et Détermination du type des Composants Argumentatifs.

posants argumentatifs. Une illustration de l'architecture du modèle est proposée en Figure 1. Les différentes couches utilisées sont présentées ci-dessous.

3.1 Plongement sémantique

Nous utilisons une première couche de plongement sémantique assignant une représentation vectorielle e_t à chaque mot w_t donné en entrée du système. Nous utilisons Glove [20] afin d'obtenir un ensemble de représentations vectorielles entraînées de manière non-supervisée [1]. Les représentations vectorielles de mots sont continuellement optimisées au cours de l'entraînement du modèle sur les différentes tâches explicitées ci-dessous. Les mots pour lesquels nous ne disposons pas de représentation vectorielle pré-entraînée sont transformés en un mot spécial <*UNK*>.

3.2 Etiquetage morpho-syntaxique

La seconde couche du modèle correspond à une tâche d'étiquetage morpho-syntaxique (EMS), consistant à assigner pour chaque mot w_t en entrée du système une étiquette morpho-syntaxique (e.g, nom commun, verbe, déterminant, etc.). Nous utilisons un Gated Recurrent Unit (GRU) [21] bi-directionnel afin d'encoder les séquences de mots en entrée du système.

GRU est un réseau de neurones récurrent utilisant un mécanisme de déclenchement sans utilisation de cellule mémoire séparée. A l'instant t, GRU calcule l'état caché h_t de la manière suivante :

$$h_t = (1 - z_t)n_t + z_t h_{(t-1)}$$

avec

$$n_t = tanh(W_n x_t + b_n + r_t(W_{hn} h_{(t-1)} + b_{hn}))$$

1. Le modèle pré-entraîné est issu de https ://nlp.stanford.edu/projects/glove/

$$r_t = \sigma(W_r x_t + b_r + W_{hr} h_{(t-1)} + b_{hr})$$
$$z_t = \sigma(W_z x_t + b_z + W_{hz} h_{(t-1)} + b_{hz})$$

où x_t représente l'entrée à l'instant t, r_t, z_t et n_t sont respectivement les portes de réinitialisation, d'entrée et de nouveauté, σ représente la fonction sigmoïde, et W et b sont des matrices et vecteurs de paramètres.

En vue d'exploiter le contexte "passé" et "futur" d'un élément d'une séquence de N éléments $[x_1, x_1, ..., x_N]$, nous pouvons construire un encodage bi-directionnel par concaténation des états cachés obtenus par un encodage séquentiel "à l'endroit" (e.g, à l'instant $t = 1$, l'entrée est x_1, à l'instant $t = 2$, l'entrée est x_2, etc.) et un encodage "à l'envers" (e.g, à l'instant $t = 1$, l'entrée est x_N, à l'instant $t = 2$, l'entrée est x_{N-1}, etc.) :

$$\overrightarrow{h_t} = \overrightarrow{GRU}(x_t), t \in [1, N]$$
$$\overleftarrow{h_t} = \overleftarrow{GRU}(x_t), t \in [N, 1]$$
$$h_t = [\overrightarrow{h_t}; \overleftarrow{h_t}]$$

Nous utilisons les représentations vectorielles des mots constituant l'exemple en cours comme entrée de la couche EMS :

$$\overrightarrow{h_t^{(1)}} = \overrightarrow{GRU}(e_t)$$
$$\overleftarrow{h_t^{(1)}} = \overleftarrow{GRU}(e_t)$$
$$h_t^{(1)} = [\overrightarrow{h_t^{(1)}}; \overleftarrow{h_t^{(1)}}]$$

Ensuite pour chaque instant t, nous calculons la probabilité d'assigner l'étiquette k au mot w_t de la manière suivante :

$$p(y_t^{(1)} = k | h_t^{(1)}) = \frac{exp(W_{sm_{(1)}} fc_t^{(1)} + b_{sm_{(1)}})}{\sum_{c_1} exp(W_{sm_{(1)}} fc_t^{(1)} + b_{sm_{(1)}})} \quad (1)$$

$$fc_t^{(1)} = relu(W_{fc_{(1)}} h_t^{(1)} + b_{fc_{(1)}}) \quad (2)$$

Avec W et b matrices et vecteurs de paramètres, $relu$ la fonction Unité de Rectification Linéaire [22], et c_1 l'ensemble des classes possibles pour l'étiquette EMS.

3.3 Chunking

Le chunking consiste à assigner une étiquette chunk (chunk nom, chunk verbe, etc.) à chaque mot. Nous calculons les états cachés relatifs au chunking en exploitant ce que le modèle a appris pour la tâche EMS :

$$\overrightarrow{h_t^{(2)}} = \overrightarrow{GRU}([e_t; h_t^{(1)}; y_t^{(EMS)}])$$
$$\overleftarrow{h_t^{(2)}} = \overleftarrow{GRU}([e_t; h_t^{(1)}; y_t^{(EMS)}])$$
$$h_t^{(2)} = [\overrightarrow{h_t^{(2)}}; \overleftarrow{h_t^{(2)}}]$$

Avec $h_t^{(1)}$ l'état caché obtenu à l'instant t pour la tâche EMS et $y_t^{(EMS)}$ la représentation vectorielle pondérée de

©AfIA 2018

l'étiquette EMS. En suivant Hashimoto et al. [16], $y_t^{(EMS)}$ est définit comme suit :

$$y_t^{(EMS)} = \sum_{j=1}^{card(c_1)} p(y_t^{(1)} = j|h_t^{(1)})l(j) \quad (3)$$

où $l(j)$ est une représentation vectorielle de la j-ème étiquette EMS. Les représentations vectorielles des étiquettes sont pré-entraînées avec GloVe.

La probabilité d'assigner une étiquette chunk à un mot est ensuite calculée de manière similaire à celle pour les étiquettes EMS (équations (1) et (2)), mais avec un ensemble de paramètres propres à la couche chunking.

3.4 Délimitation des composants argumentatifs (DelCA)

L'objectif de cette tâche est de déterminer, au mot près, les frontières de chaque composant argumentatif au sein d'une dissertation. Nous suivons Stab et Gurevych [2] et traitons cette tâche comme un problème de segmentation de texte supervisée dont les étiquettes suivent un IOB-tagset [23] : le premier mot de chaque composant argumentatif porte l'étiquette "Arg-B", les mots restant dudit composant argumentatif portent l'étiquette "Arg-I", et les mots n'appartenant pas à un composant argumentatif portent l'étiquette "O".

Chaque dissertation est traitée comme une unique séquence de mots que nous encodons de la manière suivante :

$$\overrightarrow{h_t^{(3)}} = \overrightarrow{GRU}([e_t; h_t^{(1)}; y_t^{(EMS)}; h_t^{(2)}; y_t^{(chunk)}])$$

$$\overleftarrow{h_t^{(3)}} = \overleftarrow{GRU}([e_t; h_t^{(1)}; y_t^{(EMS)}; h_t^{(2)}; y_t^{(chunk)}])$$

$$h_t^{(3)} = [\overrightarrow{h_t^{(3)}}; \overleftarrow{h_t^{(3)}}]$$

où $y_t^{(chunk)}$ est la représentation vectorielle pondérée de l'étiquette chunk, calculée de manière similaire à celle de l'étiquette EMS (équation (3)).

La probabilité d'assigner une étiquette à un mot est ensuite calculée de manière similaire à celle pour les étiquettes EMS, mais avec un ensemble de paramètres propres à la couche DelCA.

3.5 Déterminer le type des composants argumentatifs (DetCA)

L'objectif de cette tâche est de déterminer le type de chaque composant argumentatif parmi prémisse, conclusion intermédiaire et conclusion majeure. Nous traitons cette tâche comme un problème d'étiquetage de segment. Nous considérons qu'un segment peut être la séquence des mots appartenant à un même composant argumentatif ou la séquence des mots appartenant à une même portion de texte continue dont les mots n'appartiennent pas à un composant argumentatif. La notion de segment est illustré en Figure 2.

[S1] The greater our goal is, the more competition we need. [S2] Take Olympic games which is a form of competition for instance, it is hard to imagine how an athlete could win the game without the training of his or her coach, and the help of other professional staffs such as the people who take care of his diet, and those who are in charge of the medical care [S3] . The winner is the athlete but the success belongs to the whole team. Therefore [S4] without the cooperation, there would be no victory of competition [S5] .

Consequently, no matter from the view of individual development or the relationship between competition and cooperation we can receive the same conclusion that ***[S6] a more cooperative attitudes towards life is more profitable in one's success****.*

FIGURE 2 – Un extrait d'une dissertation extrait du corpus. Les passages soulignés par un trait continu constituent des prémisses, ceux soulignés par un trait discontinu constituent des conclusions intermédiaires, et les passages en gras sont des conclusions majeures. Les numéros des segments [S#] sont rajoutés à titre indicatif. Le premier segment correspond à la portion du début du texte jusqu'à la première prémisse. Le second segment correspond à la première prémisse. Le troisième segment correspond à la portion non surlignée entre la première prémisse et la première conclusion intermédiaire, etc.

Nous encodons chaque segment $s_i, i \in [1, L]$ de la manière suivante :

$$\overrightarrow{h_{it}} = \overrightarrow{GRU}([e_{it}; h_{it}^{(1)}; y_{it}^{(EMS)}; h_{it}^{(2)}; y_{it}^{(chunk)}])$$

$$\overleftarrow{h_{it}} = \overleftarrow{GRU}([e_{it}; h_{it}^{(1)}; y_{it}^{(EMS)}; h_{it}^{(2)}; y_{it}^{(chunk)}])$$

$$h_{it} = [\overrightarrow{h_{it}}; \overleftarrow{h_{it}}]$$

où it représente l'instant t du segment s_i.

Afin que le modèle se concentre davantage sur les marqueurs potentiellement importants (comme "I firmly believe that" ou "we can receive the same conclusion that") nous utilisons un mécanisme d'attention [24], nous permettant de surcroît de synthétiser l'information portée par les états cachés d'un segment en un vecteur de taille fixe :

$$u_{it} = tanh(W_{att}h_{it} + b_{att})$$

$$\alpha_{it} = \frac{exp(u_{it}^{\intercal}u_{att})}{\sum_t exp(u_{it}^{\intercal}u_{att})}$$

$$sh_i = \sum_t \alpha_{it}h_{it}$$

Avec W_{att}, b_{att} et u_{att} respectivement matrices, biais et vecteurs de paramètres.

Nous encodons ensuite la dissertation à partir des états cachés synthétiques sh_i des segments :

$$\overrightarrow{h_j^{(4)}} = \overrightarrow{GRU}(sh_i), i \in [1, L]$$

$$\overleftarrow{h_j^{(4)}} = \overleftarrow{GRU}(sh_i), i \in [L, 1]$$

$$h_j^{(4)} = [\overrightarrow{h_j^{(4)}}; \overleftarrow{h_j^{(4)}}]$$

La probabilité d'assigner une étiquette à un segment est ensuite calculée de manière similaire à celle pour les étiquettes EMS, mais avec un ensemble de paramètres propres à la couche DetCA.

4 Entraînement du modèle

Nous entraînons le modèle en alternant les couches à chaque "epoch" dans l'ordre suivant : EMS, chunking, DelCA, DetCA. Afin d'évaluer la pertinence d'implémenter un modèle multi-tâches, nous avons entraîné une version du modèle en omettant l'optimisation des couches EMS et chunking (nommée "w/o EMS & chunking") et une version du modèle en optimisant l'ensemble des couches (nommée "w/ EMS & chunking"). Les détails de l'entraînement de chaque couche sont explicités ci-dessous.

4.1 Couche EMS

Nous suivons Hashimoto et al. [16] et notons $\theta_{EMS} = (W_{EMS}, b_{EMS}, \theta_e)$ l'ensemble des paramètres intervenant dans la couche EMS. W_{EMS} représente l'ensemble des matrices de paramètres de la couche EMS, b_{EMS} l'ensemble des biais de la couche EMS et θ_e l'ensemble des paramètres de la couche de plongement sémantique des mots. La fonction de coût est définie par :

$$J^{(1)} = -\sum_s \sum_t log\, p(y_t^{(1)} = k|h_t^{(1)}) + \lambda \left\|W_{EMS}\right\|^2 + \delta \left\|\theta_e - \theta'_e\right\|^2$$

Avec $p(y_t^{(1)} = k|h_t^{(1)})$ la probabilité d'assigner la bonne étiquette k au mot w_t de la séquence de mots s, $\lambda \left\|W_{EMS}\right\|^2$ est la régularisation L2 et $\delta \left\|\theta_e - \theta'_e\right\|^2$ un régularisateur successif. λ et δ sont des hyper-paramètres.

Le régularisateur successif a pour vocation de stabiliser l'entraînement en empêchant θ_e d'être trop modifié spécifiquement par la couche EMS. θ_e étant partagé par l'ensemble des couches du modèle, des modifications trop importantes apportées par l'entraînement de chaque couche empêcherait le modèle d'apprendre convenablement. θ'_e est l'ensemble des paramètres intervenant dans la couche de vectorisation des mots à l'époch précédente.

4.2 Couche chunking

Nous notons $\theta_{chunk} = (W_{chunk}, b_{chunk}, E_{EMS}, \theta_e)$ l'ensemble des paramètres intervenant dans la couche chunking. W_{chunk} et b_{chunk} sont respectivement les matrices de paramètres et biais de la couche chunking, incluant ceux de θ_{EMS}. E_{EMS} est l'ensemble des paramètres caractérisant la représentation vectorielle des étiquettes EMS. La fonction de coût est définie de la manière suivante :

$$J^{(2)} = -\sum_s \sum_t log\, p(y_t^{(2)} = k|h_t^{(2)}) + \lambda \left\|W_{chunking}\right\|^2 + \delta \left\|\theta_{EMS} - \theta'_{EMS}\right\|^2$$

Avec $p(y_t^{(2)} = k|h_t^{(2)})$ la probabilité d'assigner la bonne étiquette k au mot w_t de la séquence de mots s. θ'_{EMS} est l'ensemble des paramètres de la couche EMS obtenus avant d'entamer l'"epoch" courante d'entraînement de la couche chunking.

4.3 Couche DelCA

Notons $\theta_{DelCA} = (W_{DelCA}, b_{DelCA}, E_{EMS}, Echunk, \theta_e)$ l'ensemble des paramètres intervenant dans la couche DelCA, avec W_{DelCA} et b_{DelCA} respectivement matrices de paramètres et biais de la couche DelCA, incluant ceux de la couche chunking et EMS. E_{chunk} est l'ensemble des paramètres caractérisant la représentation vectorielle des étiquettes de la couche chunking. La fonction de coût est définie de la manière suivante :

$$J^{(3)} = -\sum_d \sum_t log\, p(y_t^{(3)} = k|h_t^{(3)}) + \lambda \left\|W_{DelCA}\right\|^2 + \delta \left\|\theta_{chunk} - \theta'_{chunk}\right\|^2$$

Avec $p(y_t^{(3)} = k|h_t^{(3)})$ la probabilité d'assigner la bonne étiquette k au mot w_t de la dissertation d. θ'_{chunk} est l'ensemble des paramètres de la couche chunking obtenus avant d'entamer l'"epoch" courante d'entraînement de la couche DelCA.

4.4 Couche DetCA

Notons $\theta_{DetCA} = (W_{DetCA}, b_{DetCA}, E_{EMS}, Echunk, \theta_e)$ l'ensemble des paramètres intervenant dans la couche DetCA, avec W_{DetCA} et b_{DetCA} respectivement matrices de paramètres et biais de la couche DetCA, incluant ceux de la couche chunking et EMS. La fonction de coût est définie de la manière suivante :

$$J^{(4)} = -\sum_d \sum_i log\, p(y_i^{(4)} = k|sh_i^{(4)}) + \lambda \left\|W_{DetCA}\right\|^2 + \delta \left\|\theta_{chunk} - \theta'_{chunk}\right\|^2$$

Avec $p(y_i^{(4)} = k|sh_i^{(4)})$ la probabilité d'assigner la bonne étiquette k au segment s_i de la dissertation d.

5 Expérimentations et résultats

5.1 Hyper-paramètres et données utilisées

Optimisation. Nous entraînons le modèle en alternant les couches, suivant l'ordre suivant : EMS, chunking, DelCA, DetCA. Chaque couche est entraînée pendant une "epoch" avant de passer à la couche suivante. Nous utilisons Adam [25] comme algorithme d'apprentissage, avec

©AfIA 2018

$\beta_1 = 0.9$, $\beta_2 = 0.999$ et $\epsilon = 10^{-8}$. Le coefficient d'apprentissage est commun à toutes les couches et fixé à 10^{-3} au début de l'entraînement, puis multiplité par 0.75 toutes les 10 "epoch". Afin de limiter le problème d'explosion du gradient, nous redimensionnons sa norme avec une stratégie de gradient clipping [26]. Nous suivons [16] et appliquons un gradient clipping de $min(3.0, profondeur)$, où $profondeur$ représente le nombre de GRU impliquées dans la couche entraînée.

Initialisation des paramètres. Afin de faciliter la propagation du gradient lors de l'entraînement, nous utilisons des matrices orthogonales générées aléatoirement comme états initiaux pour les matrices de paramètres des GRU, comme préconisé par Saxe et al. [27]. Les autres matrices de paramètres sont initialisées avec des valeurs issues d'une loi normal $\mathcal{N}(0, \sqrt{2/n_{in}})$, où n_{in} représente le nombre de neurones entrant dans la couche concernée, comme proposé par He et al [28]. Les vecteurs de biais sont initialisés en tant que vecteurs nuls.

Dimensions vectorielles utilisées. La représentation vectorielle utilisée pour les mots en entrée du système et les représentations vectorielles des étiquettes EMS et chunking sont de dimension 50. Les états cachés des GRU sont de dimension 100 pour toutes les couches du modèle.

Régularisation. En suivant [16], nous fixons les coefficients λ à 10^{-6} pour les matrices de paramètres des GRU et 10^{-5} pour les autres matrices de paramètres. Le coefficient de régularisation successif δ est fixé à 10^{-2} pour toutes les couches. Nous utilisons aussi Dropout [29] sur toutes les couches, avec taux de neurones affectés de 0.2.

Données d'entraînement pour les couches EMS et chunking. Nous utilisons le corpus issu de la tâche partagée CoNLL-2000 [30] avec les étiquettes associées pour entraîner les couches EMS et chunking.

Données d'entraînement pour les couches DelCA et DetCA. Nous utilisons le corpus Argument Annotated Essays (version 2) partagé par Stab et Gurevych [2] en suivant le découpage entraînement/test fourni pour l'entraînement des couches DelCA et DetCA.

Arrêt de l'entraînement. Dans un cas d'entraînement uni-tâche, une pratique généralement adoptée est d'arrêter l'entraînement du modèle peu avant le surapprentissage. Dans le cas de notre modèle, il n'est pas évident de déterminer le meilleur moment pour arrêter l'entraînement, puisque le modèle peut surapprendre sur une tâche particulière, mais pas sur les autres. Ainsi, nous arrêtons l'entraînement du modèle lorsqu'il surapprend sur les couches DelCA et DetCA, et reportons les meilleurs résultats obtenus pour chaque tâche avant le surapprentissage de celle-ci.

DetCa simple. Nous nommons DetCa simple la tâche DetCa avec la modification suivante : tous les segments des dissertations correspondant à des composants argumentatifs sont traités comme ne comportant qu'un unique mot spécial <*VIDE*>. L'hypothèse est que cette transformation forcera le modèle à se concentrer sur le contexte entourant les composants argumentatifs, et l'empêchera donc de se surentraîner en considérant les mots à l'intérieur des composants.

Tâche	w/o EMS & chunking	w/ EMS & chunking
DelCA	0.5934	0.8688
DetCA	0.7464	0.7105
DetCA simple	0.7529	0.7911

TABLE 1 – Macros f1-scores obtenus sur les différentes tâches.

Tâche	F1-score obtenus en [2]	F1-score humain
DelCA	0.867	0.886
DetCA	0.826	0.868

TABLE 2 – F1-scores obtenus sur les tâches DelCA et DetCA par Stab et Gurevych [2] et des agents humains.

5.2 Résultats obtenus

Les résultats obtenus sur les données de test pour les tâches DelCA, DetCA et DetCA simple sont présentés en Table 1. La colonne "w/o EMS & chunking" fait référence à la version du modèle pour laquelle l'optimisation des couches EMS et chunking a été omise. La colonne "w/ EMS & chunking" fait référence à la version du modèle pour laquelle l'optimisation des couches EMS et chunking a été réalisée. Nous prenons comme référence les performances atteintes par des agents humains [2] ainsi que les résultats présentés par Stab et Gurevych [2], illustrés en Table 2.

Evaluation générale des performances. Nous obtenons un macro f1-score de 0.8688 sur DelCA avec la version "w/ EMS & chunking". Ces résultats sont obtenus sans définition de caractéristiques manuelles et sont comparables à ceux enregistrés en [2] ; ils atteignent 98,06% de la performance humaine. Concernant la classification des composants argumentatifs, nous obtenons un macro f1-score de 0.7911 avec DetCA simple pour la version "w/ EMS & chunking", ce qui représente 95,8% des performances obtenues en [2] et 91,1% de la performance humaine.

Pertinence de DetCA simple. Selon nous, les mots formant un composant argumentatif ne sont pas réellement caractéristiques de sa classe, et en se concentrant dessus, le modèle peut être amené à modéliser du bruit l'empêchant de généraliser correctement. En revanche, le contexte dans lequel apparaissent les composants semble très important. Par exemple, des mots tels que "we can receive the same conclusion that" semblent indiquer que l'auteur va annoncer une conclusion intermédiaire ou majeure. Cela peut expliquer la différence de performances entre DetCA et DetCA simple, notamment pour la version "w/ EMS

2. La performance humaine correspond à la moyenne des résultats obtenus par des annotateurs humains, tels que présentés en [2]

& chunking", avec respectivement un f1-score de 0.7105 contre 0.7911, soit une amélioration de 11,3%.

Pertinence du modèle multi-tâches. Les macro f1-scores sur les tâches DelCA et DetCA simple sont respectivement de 0.5934 et 0.7529 pour la version "w/ EMS & chunking" et de 0.8688 et 0.7911, soit des améliorations de 46,4% et 5,1%. Ces résultats permettent donc de valider l'intérêt d'entraîner un modèle multi-tâches et incitent à l'ajout de tâches auxiliaires supplémentaires.

6 Travaux à venir et perspectives

Les résultats obtenus sont encourageants et pourraient sûrement être améliorés, notamment avec une recherche plus d'approfondie d'hyper-paramètres optimaux. La différence de performances entre les versions du modèle "w/ EMS & chunking" et "w/o EMS & chunking" portent à croire qu'implémenter davantage de tâches auxiliaires pourrait être bénéfique. Une piste serait d'introduire une couche modélisant un arbre de dépendances syntaxiques en complément de la couche chunking, comme effectué en [16].

En vue d'implémenter un système complet d'argument mining tel que présenté par Stab et Gurevych [2], nous prévoyons d'implémenter des couches permettant la génération automatique de graphes d'arguments. A cette fin il est nécessaire de déterminer s'il existe un arc entre chaque paire ordonnée de composants argumentatifs, ainsi que d'inférer l'étiquette portée par ledit arc.

7 Conclusion

Cet article a présenté une méthode d'extraction et d'analyse automatique d'arguments à partir de textes bruts. L'utilisation de techniques d'apprentissage profond nous permet de nous affranchir de la définition de caractéristiques manuellement définies. Par ailleurs, l'amélioration des performances de notre système par l'exploitation de paramètres optimisés sur des tâches auxiliaires met en avant l'intérêt de l'utilisation d'un modèle multi-tâches. Nous avons comme perspective la complétion de la chaîne de traitement existante en vue d'obtenir un système capable de synthétiser une dissertation par modélisation automatique d'un graphe d'arguments.

Références

[1] M. Lippi et P. Torroni, Argumentation mining : State of the art and emerging trends, *ACM Transactions on Internet Technology (TOIT)*, 16(2), p.10, 2016.

[2] C. Stab et I. Gurevych, Parsing argumentation structures in persuasive essays, *Computational Linguistics*, 43(3), pp.619-659, 2017.

[3] N. Madnani, M. Heilman, J. Tetreault et M. Chodorow, Identifying high-level organizational elements in argumentative discourse, *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies*, pp. 20-28, Association for Computational Linguistics, 2012.

[4] R. Levy, Y. Bilu, D. Hershcovich, E. Aharoni et N. Slonim, Context dependent claim detection, *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics : Technical Papers*, pp. 1489-1500, 2014.

[5] Y. Ajjour, W.F. Chen, J. Kiesel, H. Wachsmuth et B. Stein, Unit Segmentation of Argumentative Texts, *Proceedings of the 4th Workshop on Argument Mining*, pp. 118-128, 2017.

[6] T. Goudas, C. Louizos, G. Petasis et V. Karkaletsis, Argument extraction from news, blogs, and social media, *Hellenic Conference on Artificial Intelligence*, pp. 287-299, Springer, Cham, 2014.

[7] C. Sardianos, I.M. Katakis, G. Petasis et V. Karkaletsis, Argument extraction from news, *Proceedings of the 2nd Workshop on Argumentation Mining*, pp. 56-66, 2015.

[8] S. Eger, J. Daxenberger et I. Gurevych, Neural End-to-End Learning for Computational Argumentation Mining, *arXiv preprint arXiv :1704.06104*, 2017.

[9] J. Eckle-Kohler, R. Kluge et I. Gurevych, On the role of discourse markers for discriminating claims and premises in argumentative discourse, *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 2236-2242, 2015.

[10] J. Park et C. Cardie, Identifying appropriate support for propositions in online user comments. *Proceedings of the 1st Workshop on Argumentation Mining*, pp. 29-38, 2014.

[11] I. Persing et V. Ng, End-to-End Argumentation Mining in Student Essays, *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies*, Association for Computational Linguistics, pages 1384–1394, 2016.

[12] P. Potash, A. Romanov et A. Rumshisky, Here's My Point : Joint Pointer Architecture for Argument Mining, *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp. 1364-1373, 2017.

[13] S. Ruder, An overview of multi-task learning in deep neural networks, *CoRR*, abs/1706.05098, 2017.

[14] A. Søgaard et Y. Goldberg, Deep multi-task learning with low level tasks supervised at lower layers, *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2 : Short Papers)*, Vol. 2, pp.231-235, 2016.

[15] Z. Yang, R. Salakhutdinov et W. Cohen, Multi-task cross-lingual sequence tagging from scratch, *arXiv preprint arXiv :1603.06270*, 2016.

©AfIA 2018

[16] K. Hashimoto, C. Xiong, Y. Tsuruoka et R. Socher, A joint many-task model : Growing a neural network for multiple nlp tasks, *Empirical Methods in Natural Language Processing (EMNLP)*, 2017.

[17] L. Mou, Z. Meng, R. Yan, G. Li, Y. Xu, L. Zhang et Z. Jin, How transferable are neural networks in nlp applications ?, *Empirical Methods in Natural Language Processing (EMNLP)*, pp. 479–489, 2016.

[18] H.M Alonso et B. Plank, When is multitask learning effective ? Semantic sequence prediction under varying data conditions, *15th Conference of the European Chapter of the Association for Computational Linguistics*, 2017.

[19] J. Bingel et A. Søgaard, Identifying beneficial task relations for multi-task learning in deep neural networks, *arXiv preprint arXiv :1702.08303*, 2017.

[20] J. Pennington, R. Socher et C. Manning, Glove : Global vectors for word representation. *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pp. 1532-1543, 2014.

[21] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk et Y. Bengio, Learning phrase representations using RNN encoder-decoder for statistical machine translation, *arXiv preprint arXiv :1406.1078*, 2014.

[22] V. Nair et G.E. Hinton, Rectified linear units improve restricted boltzmann machines, *Proceedings of the 27th international conference on machine learning (ICML-10)*, pp. 807-814, 2010.

[23] L.A. Ramshaw et M.P. Marcus, Text chunking using transformation-based learning, *Natural language processing using very large corpora*, pp. 157-176, Springer, Dordrecht, 1999.

[24] D. Bahdanau, K. Cho et Y. Bengio, Neural machine translation by jointly learning to align and translate, *ICLR*, 2015.

[25] D.P. Kingma et J. Ba, Adam : A method for stochastic optimization, *ICLR*, 2015.

[26] R. Pascanu, T. Mikolov et Y. Bengio, On the difficulty of training recurrent neural networks, *Proceedings of The 30th International Conference on Machine Learning*, pp. 1310–1318, 2013.

[27] A.M. Saxe, J.L. McClelland, S. Ganguli, Exact solutions to the nonlinear dynamics of learning in deep linear neural networks, *Proceedings of the International Conference on Learning Representations (ICLR)*, 2014.

[28] K. He, X. Zhang, S. Ren et J. Sun, Delving Deep into Rectifiers : Surpassing Human-Level Performance on ImageNet Classification, *ICCV*, 2015.

[29] N. Srivastava, G.E. Hinton, A. Krizhevsky, I. Sutskever et R. Salakhutdinov, Dropout : a simple way to prevent neural networks from overfitting, *Journal of machine learning research*, 15(1) :1929–1958, 2014.

[30] E.F.T.K. Sang, S. Buchholz, Introduction to the CoNLL-2000 shared task : chunking, *Proceedings of the 2nd Workshop on Learning Language in Logic and the 4th Conference on Computational Natural Language Learning*, Lisbon, Portugal, vol. 7, 2000, pp. 127–132, 2000.

#AIDS Analyse Information Dangers Sexualité : caractériser les discours à propos du VIH dans les forums de santé

Yves Mercadier[1] Jérôme Azé[1] Sandra Bringay[1,2] Viviane Clavier[3]
Erick Cuenca[1] Céline Paganelli[4] Pascal Poncelet[1] Arnaud Sallaberry[1,2]

[1] LIRMM, Université de Montpellier, CNRS, Montpellier, France, prenom.nom@lirmm.fr
[2] AMIS, Université Paul Valéry, Montpellier, France
[3] GRESEC EA 608 -– Université Grenoble Alpes, France
[4] LERASS EA 827, Université Paul Valéry, F34000, Montpellier, France

Résumé

Contexte : *Les forums de discussion consacrés au VIH remplissent trois fonctions. Tout d'abord, ils sont utilisés par les internautes comme sources d'information pour se renseigner sur les traitements, les risques d'infection liés au sida ou le « vivre avec » la maladie ou le virus. Ensuite, ils servent de média pour les institutions de santé ou les associations afin de communiquer des informations de prévention sur le VIH. Enfin, ils apparaissent comme sources de connaissances pour les professionnels de santé (médecins, soignants) pour comprendre les comportements liés au VIH, et pour les professionnels de la prévention, pour modérer les forums et adapter les campagnes de prévention aux différents publics.*
Objectif : *Notre objectif dans cet article est de proposer un processus d'ingénierie des connaissances complet, permettant : 1) de collecter des messages dans des médias sociaux ; 2) de classifier ces messages de manière semi-automatique selon le genre, le niveau d'expertise, le niveau informationnel, le type de risque ainsi que les émotions exprimées ; 3) de visualiser ces nouvelles connaissances dans une représentation originale prenant en compte la temporalité et la hiérarchie sous-jacente à la classification. Cette visualisation pourrait permettre aux gestionnaires de sites de forums et aux professionnels de santé de naviguer dans le flot de messages pour suivre l'évolution de l'importance de ces thématiques.*
Méthodes : *Notre approche combine une démarche qualitative et quantitative. Nous avons travaillé sur plus de* 226 252 *messages issus du forum Sida Info Service. Deux chercheurs en sciences de l'information et de la communication ont élaboré une grille d'analyse, puis ont annoté* 4 481 *messages selon cette grille. Ces données ont été utilisées pour apprendre des classifieurs qui ont permis d'étiqueter l'ensemble des messages du site à notre disposition. Afin de définir les meilleurs classifieurs, nous avons comparé l'efficacité des méthodes de classification traditionnelles statistiques et plusieurs architectures d'apprentissage profond. Une fois les messages étiquetés, nous avons utilisé une visualisation de type streamgraph, combinée avec un outil de navigation hiérarchique, pour visualiser l'évolution de ces annotations dans le temps.*
Résultats : *Les résultats sont prometteurs et montrent l'efficacité des méthodes d'apprentissage profond pour caractériser les messages des forums de manière automatique. La méthode de visualisation mise en place permet d'explorer les résultats de ces méthodes et ainsi faciliter l'accès aux connaissances.*

Mots Clef

Classification automatique de textes, Apprentissage profond, Visualisation, VIH

Abstract

It's the English version of the abstract. Exactly as in French it must be short. It must speak of the same topics...

Keywords

Example, model, template.

1 Introduction

À l'heure où 3,9 milliards d'individus dans le monde sont connectés à internet (soit 51% de la population mondiale), et où 2,91 milliards d'entre eux sont inscrits sur des réseaux sociaux (soit 29% de la population mondiale) [1], les pratiques pour s'informer et communiquer ont considérablement évolué et sont marquées par les innovations technologiques récentes et la place occupée par le numérique. En 2017, 74% des français accèdent tous les jours à Internet (95% des 18-24 ans) et passent en moyenne 1 heure 15 par jour sur les réseaux sociaux.
Les travaux présentés dans cet article se déroulent dans le cadre du projet #AIDS (Analyse Information Dangers

1. https://www.blogdumoderateur.com/chiffres-reseaux-sociaux/

 ©AfIA 2018

Sexualité) qui réunit une équipe de chercheurs pluridisciplinaire en sciences de l'information et de la communication et en informatique, ainsi que des acteurs de la santé et de la prévention. Ce projet vise à analyser semi-automatiquement les contenus échangés sur les médias sociaux (forums de discussion, réseaux sociaux, plateformes de microblogging) consacrés au VIH pour en extraire de la connaissance. Avec plus de 35 millions de morts à ce jour, le VIH continue de représenter un problème mondial majeur de santé publique. En 2016, 1 million de personnes sont décédées d'une ou des causes liées au VIH dans le monde [2]. Toujours en 2016, en France, 6 003 personnes ont découvert leur séropositivité.

Les médias sociaux consacrés aux VIH s'adressent principalement aux personnes infectées et aux malades dans une démarche de prévention ou d'accompagnement mais également aux internautes soucieux de connaître les risques de contamination. Les publications peuvent prendre plusieurs formes comme des dispositifs de prévention, d'accompagnement, d'amélioration de la santé mentale et physique, des espaces d'échanges où sont évoqués les traitements, le "vivre avec" la maladie ou le virus, et plus largement des questions de la vie de tous les jours ou encore donner lieu à des usages sexuels d'Internet.

Un grand nombre de ressources en ligne sont disponibles, à toute heure. De nombreux blogs comme *Je VIH avec* [3] alimentés par des médecins, des porteurs du virus, des associations de malades sont présents sur le web. De nombreux groupes de discussion sont consacrés à ces thématiques sur des réseaux sociaux destinés aux patients comme Carenity [4] ou encore sur le site généraliste Doctissimo [5]. On dénombre, au 1er mars 2018, une centaine de groupes facebook français en lien avec le VIH. Plus de 500 comptes twitter mentionnent les hashtags #VIH ou #SIDA dans l'intitulé de leur compte, au 1er mars 2018. Ces comptes émanent d'associations, de particuliers, d'institutions ou de médecins. Ils portent pour la plupart sur les traitements ou sur la vie quotidienne avec le VIH. Dans ce travail, nous nous focaliserons sur les messages qui nous ont été fournis par les gestionnaires du forum associé au site Sida Info Service (SIS) [6].

Si les médias sociaux ont largement été étudiés, peu de travaux ont jusqu'à présent été menés sur les médias sociaux consacrés au VIH et en particuliers les forums de santé. L'originalité de ce travail est de considérer, d'une part ces médias comme des dispositifs info-communicationnels sur lesquels se développent de nouvelles formes d'échanges et qui véhiculent des informations n'ayant pas leur place dans d'autres espaces (cabinet médical, discussions familiales, médias traditionnels...). D'autre part, les discours produits sur ces médias reflètent des pratiques ordinaires relatives aux questions de sexualité, de consommations de produits illicites et d'addictions et permettent de détecter des pratiques émergentes, de nouvelles conduites à risque. **Notre objectif dans cet article est de proposer une chaîne de traitements incluant une phase de classification et un outil de visualisation pour mieux analyser et comprendre les contenus de ces discours. Ces informations sont essentielles pour la modération du site.**

Notre approche combine une démarche qualitative et quantitative. Nous avons travaillé sur plus de 226 252 messages issus du forum SIS. Deux professionnels de l'information et de la communication ont élaboré une grille d'annotation, puis ont annoté 4 481 messages selon cette grille. Ces données ont été utilisées pour apprendre des classifieurs supervisés qui ont permis d'étiqueter dans un deuxième temps, les 226 252 messages de manière automatique. Une fois les messages étiquetés, nous avons utilisé une approche *MultiStream* [8] adaptée aux séries temporelles multiples. L'originalité est ici de pouvoir visualiser simultanément la hiérarchie associée aux annotations et leur évolution dans le temps.

Concernant la phase de classification automatique de textes il s'agit d'un sujet classique de Traitement Automatique de la Langue (TAL), qui consiste à assigner des catégories prédéfinies à des documents en texte libre. Les approches de classification classiques se concentrent sur le choix du meilleur classifieur (e.g. SVM ou régression logistique) et sur la définition des meilleures caractéristiques prises en entrée de ces classifieurs. La plupart des techniques sont basées sur les mots, les lexiques et sont spécifiques à une tâche particulière. Ces modèles ont été appliqués avec succès sur de très hautes caractéristiques dimensionnelles parfois éparses. Dernièrement, pour de nombreuses tâches de classification de textes, les méthodes d'apprentissage profond se sont révélées efficaces et ont permis de faire des progrès importants dans des domaines tels que la reconnaissance de formes (pattern recognition) [4] ou la bioinformatique [23].

Cette tendance s'est confirmée avec le succès des word embeddings [21, 22] et des méthodes d'apprentissage profond [35]. L'apprentissage profond permet l'apprentissage automatique de représentations à plusieurs niveaux.

[7] ont démontré qu'une architecture d'apprentissage profond, même simple, surpasse la plupart des approches classiques pour des tâches variées de TAL telles que la reconnaissance d'entités nommées (NER) [1], le Parsing [39, 45], l'étiquetage de rôles sémantiques (SRL) (Semantic Role Labeling) [13, 36], le marquage POS [3, 18], la classification de sentiments [32, 25, 15, 16], la traduction automatique [37]. Depuis, de nombreux algorithmes complexes d'apprentissage profond ont été proposés pour résoudre ces tâches difficiles. **Dans cet article, nous allons entre autres comparer l'efficacité des méthodes de classification traditionnelles et plusieurs architectures d'apprentissage profond.**

2. http://www.who.int/mediacentre/factsheets/fs360/fr/
3. http://je-vih-avec.over-blog.com/
4. https://www.carenity.com/
5. http://www.doctissimo.fr/
6. https://www.sida-info-service.org/

De nombreux travaux en visualisation portent sur la représentation de données textuelles [7] [17] et/ou temporelles [8] [2]. Une approche prometteuse, permettant de représenter des données temporelles issues de textes, a initialement été proposée par [12] sous le nom de *ThemeRiver*. Généralisée sous le nom de *StreamGraphs*, cette approche consiste à superposer des flots représentant l'évolution d'une valeur numérique (le plus souvent issue de textes) dans le temps [5]. De nombreuses évolutions ont été proposées, e.g. [9, 38, 42]. La principale limite de ces approches est la difficulté de représenter un grand nombre de flots et un long intervalle de temps. La méthode utilisée dans cet article appelée *MultiStream* [8], pallie ces problèmes en décrivant une approche de type focus+contexte combinant un ensemble de vues interactives. **Dans cet article, nous montrerons comment cette visualisation aide le chercheur en science de l'information et de la communication à analyser l'ensemble des messages du site.**

L'article est organisé en quatre sections. Dans la section 2, nous présentons notre processus. Dans la section 3, nous présentons l'évaluation de la chaîne de traitements proposée. Dans la section 4, nous discutons ces résultats et ouvrons des perspectives.

2 Méthodes

La figure 1 illustre la méthode proposée, structurée en 4 étapes : (1) Collecte et nettoyage de données, (2) Annotation manuelle par deux experts, (3) Prédiction automatique de catégories, (4) Visualisation temporelle et hiérarchique de ces catégories.

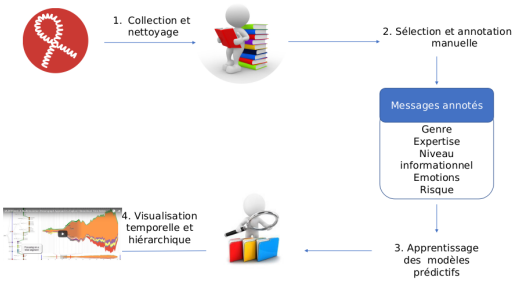


FIGURE *1 – Processus en 4 étapes*

2.1 Étape 1 : Collection et nettoyage des données

Nous avons effectué une analyse de contenu de 226 252 messages en français postés sur le forum SIS qui nous ont été fournis par les gestionnaires du forum [9]. Ces messages ont été anonymisés en supprimant les pseudonymes, noms, prénoms et localités. Puis les prétraitements suivants ont été appliqués : suppression des ponctuations et des caractères spéciaux, changement des majuscules en minuscules, suppression des mots trop fréquents pour apporter de l'information à partir d'une liste de stop words et enfin stemmatisation consistant à ne conserver que le radical des mots, pour regrouper sous le même radical tous les mots d'une famille.

2.2 Étape 2 : Annotation manuelle

Deux chercheurs en sciences de l'information et de la communication impliqués dans le projet, ont élaboré une grille d'annotation permettant de caractériser les types d'informations véhiculées par les messages postés sur les médias sociaux à partir d'une typologie testée précédemment sur des forums de discussion de santé et en tenant compte de la temporalité de l'écriture, des thématiques abordées, des micro-échanges [28, 6]. 4 481 messages ont été annotés manuellement par les deux experts [10] (sans croisement des annotations). Les chercheurs ont systématiquement annoté l'un des forums général de SIS (le plus volumineux), qui comporte plusieurs dizaines de fils de discussion, sans faire de sélection *a priori* des thématiques. La grille s'organise selon 5 dimensions :

1. le **genre** de l'auteur décliné en femme et homme ;
2. le **niveau d'expertise** de l'auteur. On distingue les messages des modérateurs, correspondant à des savoirs, de ceux des patients, qui parfois citent leur source ou rapportent des discours ;
3. le **niveau informationnel** du message. Les messages peuvent correspondre à des partages (d'expérience ou de connaissances), à des questions formulées explicitement ou non, à des témoignages faisant référence au vécu, à l'expérience, à des informations spécialisées (qui rapportent notamment des discours professionnels) ou scientifique (contenu supposé objectif dont la validité ne s'appuie pas sur le vécu), à des conseils (informations pratiques, procédures, protocoles) ;
4. les **émotions** exprimées par les participants en considérant un spectre émotionnel : la panique, l'angoisse, la peur, l'inquiétude, la réassurance (le fait de rassurer quelqu'un) ou le fait d'être rassuré.
5. le **risque**. Un message peut faire référence à un risque encouru par les internautes lors d'un événement passé, présent ou à venir, déclencheur d'une sensation de prise de risque. Les types de situation qui déclenchent une sensation de risque n'ont pas été annotées compte tenu du fait que ces situations sont bien connues des associations et figurent dans la rubrique « Questions fréquentes ». En revanche, nous avons laissé la possibilité à l'annotateur d'introduire un commentaire libre sur l'éventuelle prise

7. http://textvis.lnu.se/
8. https://vcg.informatik.uni-rostock.de/~ct/timeviz/timeviz.html
9. Accord CNIL, Certificat d'enregistrement 2-17031
10. Avec l'aide de Amal Jaouzi, stagiaire, doctorante en sciences de l'information et la communication.

©AfIA 2018

d'alcool et de drogue durant l'événement déclencheur.

Un message peut donc être décrit selon ces 5 dimensions. Dans la Table 1, nous résumons la liste des classes définies après la phase d'annotation manuelle, ainsi que le nombre de messages annotés avec ces étiquettes. Chaque message peut être associé à zéro, une ou plusieurs classes. Ces classes seront utilisées comme sorties pour la prédiction de catégories.

TABLE *1 – Sortie de la prédiction*

Index	Dimensions d'analyse	Classe	Effectif
C1	genre	Femme	36
C2	genre	Homme	67
C3	expertise	modérateur	46
C4	expertise	savoir patient	222
C5	expertise	sources citées	19
C6	expertise	discours rapporté	8
C7	information	demande	402
C8	information	informations scientifiques ou pratiques	661
C9	information	apport-réponse	874
C10	information	témoignage	284
C11	information	information vide	1144
C12	information	conseils	159
C13	émotion	peur	66
C14	émotion	inquiétude	109
C15	émotion	panique	34
C16	émotion	général	28
C17	émotion	angoisse	142
C18	émotion	réassurance	57
C19	émotion	rassuré	22
C20	risque	anticipé	8
C21	risque	passé	92
C22	risque	actuel	1

2.3 Étape 3 : Modèles pour la prédiction de catégories

Dans cette étude, nous explorons l'utilisation de méthodes d'apprentissage profond pour prédire les 22 catégories des messages collectés. Nous avons comparé les résultats des modèles d'apprentissage profond aux méthodes de classification traditionnelles. Dans cette section, nous détaillons comment chaque modèle est construit.

Modèles. Les modèles utilisés dans notre étude sont résumés dans le tableau 2. Les modèles d'apprentissages classiques comparés sont de type SVM (LSVC, SSVC), Naïve Bayes (MNB, BNB), Régression linéaire (SGDC), Arbre de décision (DT, AB, RF), K-plus-proches-voisins (KNC, PA). Dans ce qui suit, nous discutons plus précisément les modèles d'apprentissage profond. Tous nos modèles sont construits à partir de la même couche d'entrée et de la même couche de sortie. La couche d'entrée est une couche de type embedding permettant de caractériser chaque mot du document analysé par un vecteur de dimension 300. La couche de sortie contient elle, autant de neurones qu'il y a de classes dans notre ensemble de données. Ensuite, nous décrivons les couches intermédiaires de nos différents modèles de réseaux neuronaux. Nous invitons le lecteur à consulter des revues plus détaillées telles que celle de [33].

— *Basic neural network (BNN)* connecte directement la couche d'entrée avec la couche de sortie.

TABLE *2 – Modèles de Classification : classiques vs apprentissage profond*

Classiques	Apprentissage profond
LinearSVC (LSVC)	Basic neural network (BNN)
SigmoïdSVC (SSVC)	Multi-layer perceptrons (MLP)
MultinomialNB (MNB)	Deep Multi-layer perceptrons (DMLP)
BernoulliNB (BNB)	Long short-term memory (LSTM) RNN
SGDClassifier (SGDC)	Bidirectional Long short-term memory (BLSTM) RNN
Decision Tree (DT)	Gated Recurrent Unit (GRU) RNN
AdaBoost (AB)	Long-term Recurrent Convolutional Networks (LRCN)
Random Forest (RF)	Convolutional neural network (CNN)
KNeighborsClassifier (KNC)	Serial Convolutional neural network (SCNN)
	Multi Channel Convolutional neural network (MCNN)

— *Multilayer Perceptrons (MLP)* est le modèle le plus simple d'apprentissage profond. Il est réalisé par une couche cachée de neurones entre la couche d'entrée et de sortie.

— *Deep Multilayer Perceptrons (DMLP)* est basé sur 20 couches cachées.

— *Long Short-Term Memory (LSTM)* est un sous-type de RNN. Les réseaux neuronaux récurrents permettent l'étude de séquences de données. Ces réseaux, composés de plusieurs couches, estiment leurs sorties en fonction des états de leurs couches précédentes en utilisant une mémoire intermédiaire. LSTM est basé sur des blocs de mémoire qui sont utilisés comme unités dans la couche récurrente pour capturer des dépendances à plus longue portée.

— *Bidirectional Long Short-Term Memory (BLSTM)* permet d'utiliser l'information après et avant les données étudiées par le réseau au temps t.

— *Gated Recurrent Unit (GRU)* est un réseau de neurones permettant une étude plus rapide que le LSTM tout en conservant les qualités de celui-ci.

— *Long-term Recurrent Convolutional Networks (LRCN)* associe en série un réseau convolutif et un réseau LSTM.

— *Convolutional Neural Network (CNN)* sont structurés par deux opérations : convolution puis max-pooling. La convolution est basée sur plusieurs filtres combinés pour extraire les nombreuses propriétés associées aux données. La seconde opération compresse les résultats de l'opération précédente pour extraire des informations denses.

— *Serial Convolutional neural network (SCNN)* place séquentiellement quatre réseaux convolutionnels les uns après les autres.

— *Multi Channel Convolutional neural network (MCNN)* est basé sur six réseaux de convolution branchés sur la couche d'entrée. Chaque réseau de convolution possède un noyau différent des autres. Ensuite, nous associons la sortie des six réseaux de convolution à une seule couche pour construire une sortie mono-couche entièrement connectée.

Entrées des modèles : Les entrées des modèles sont les messages. Nous utilisons deux pré-traitements distincts

pour les deux types de méthodes de classification étudiées. Pour les méthodes de classification traditionnelles, nous appliquons une vectorisation basée sur la mesure Tf-Idf [30]. Pour les méthodes d'apprentissage profond, les données sont préparées différemment. Nous conservons les 5000 mots les plus fréquents (hors stop words) et nous représentons chaque document par une séquence de mots ce qui permet de conserver l'ordre des mots.

Sorties des Modèles : Les effectifs dans les classes sont très déséquilibrés. Nous avons donc mis en place deux stratégies que nous appelons *précise* et *imprécise*. Dans la stratégie *précise*, l'objectif est de décrire finement un message, par exemple selon la dimension émotion, en l'associant à l'étiquette *peur* ou *angoisse*. Dans la stratégie *imprécise*, le niveau de description attendu est plus général et le message pourrait être associé à une *émotion négative*. Nous avons donc regroupé manuellement, après discussion avec les experts, certaines classes. Par ailleurs, un grand nombre de messages ne sont pas étiquetés comme le montre le tableau 1. Pour le niveau d'expertise, les émotions et le risque, nous avons considéré une classe *Autres* correspondant à des messages non étiquetés. Nous effectuons sur ces échantillons un ré-échantillonnage de la taille de la classe majoritaire (non étiquetée). Une description des classes finales est donnée dans la table 3.

TABLE *3 – Classes utilisées dans les modèles selon les deux stratégies précise et imprécise*

Stratégie	Précise (Nombre de posts)	Imprécise (Nombre de posts)
Genre	*genre_1* C1 (57) C2 (28)	
Niveau d'expertise	*expe_1* C3 (222) C4 (46) C5/C6 (74)	*expe_2* C3/C4/C5/C6 (342) Autres (342)
Niveau informationnel	*info_1* C7 (402) C8 (661) C9 (874) C10 (284) C11 (1144) C12 (159)	*info_2* C7 (402) C8/C9/C10/C12 (1978) C11 (1144)
Émotion	*emot_1* C13/C14 (65) C15/C17 (544) C18/C19 (30)	*emot_2* C13/C14/C15/C16 (609) C18/C19 (30) Autres (609)
Risque		*risq* C20/C21/C22 (101) Autres (101)

Partitionnement des données et entraînement Afin d'évaluer les différents modèles, nous appliquons une validation croisée en dix plis. L'ensemble de données est divisé en dix sous-ensembles. Nous utilisons neuf sous-ensembles pour la phase d'apprentissage et un sous-ensemble pour la phase de validation. Nous avons répété ce processus dix fois. Pour chaque pli, nous utilisons un sous-ensemble de différentes phases de validation et nous avons calculé une métrique pour évaluer la performance.

Pour les algorithmes de classification classiques, nous utilisons l'outil sklearn [1] avec les paramètres par défaut. Pour toutes les architectures d'apprentissage profond, nous avons utilisé les paramètres par défaut pour la taille du Mini-batch (c.-à-d. Le nombre d'instances d'entraînement à considérer en même temps), la dimension d'intégration (embedding dimension, chaque mot est décrit par un vecteur de dimension n), le nombre d'epoch (le nombre d'itérations sur l'ensemble d'entraînement), la fonction d'activation et le taux d'abandon (Dropout ratio, ratio d'unités cachées à désactiver dans chaque formation de Mini-batch). Ensuite, nous utilisons une configuration avec un Mini-batch de taille 32, une dimension d'intégration de 300, une couche cachée de taille 256, un nombre d'epoch de 25, la fonction d'activation selu et un Dropout de 0,3.

Mesure de la performance Pour tous les modèles, nous nous sommes basés sur une métrique couramment utilisée dans le domaine de la classification. Chaque sortie est un vecteur de dimension n correspondant aux n classes à prédire. Nous mesurons la qualité des classifieurs avec la métrique *exactitude* (accuracy). Celle-ci exprime le nombre de classes prédites justes en regard du nombre total de prédictions de classes réalisées. Nous avons choisi d'utiliser une métrique stricte en considérant une sortie comme juste uniquement si l'intégralité des classes prédites sont exactes.

Exemple 1 *Considérons 3 documents décrits par un vecteur d'annotations pour 4 classes. Chaque élément i de ce vecteur correspond à la présence (1) ou à l'absence (0) de la classe i pour le document.*
$Document1 \longrightarrow 0010$
$Document2 \longrightarrow 1000$
$Document3 \longrightarrow 0100$
Considérons maintenant un modèle retournant 3 vecteurs de prédictions de ces 4 classes. Chaque élément i de ce vecteur correspond à la prédiction (1) ou à l'absence de prédiction (0) de la classe i pour le document.
$prédiction1 \longrightarrow 0010$
$prédiction2 \longrightarrow 0010$
$prédiction3 \longrightarrow 1100$
Seule la première prédiction sera considérée comme juste pour les trois prédictions émises. L'exactitude du modèle sera égale à : $acc = \frac{1}{3} = 0.333$

2.4 Étape 4 : Visualisation temporelle et hiérarchique

Le flux de messages peut être représenté sous la forme d'une série temporelle multiple, c'est-à-dire un ensemble de variables quantitatives associées à un même intervalle temporel. Au cours des dernières années, les visualisations de type *Streamgraphs* [5] ont été largement utilisées pour représenter l'évolution de plusieurs séries temporelles. Cependant, les *Streamgraphs* posent deux problèmes majeurs : (1) Comment gérer un long intervalle de

1. http://scikit-learn.org/stable/

©AfIA 2018

temps ? (2) Comment gérer un grand nombre de séries temporelles ?

Pour résoudre le problème (1), des méthodes de type vue d'ensemble+détail ou encore focus+contexte peuvent être mises en place (voir [24] pour une introduction). Les méthodes vue d'ensemble+détail consistent à combiner deux vues, une montrant l'ensemble des données de façon agrégée, l'autre montrant en détail un sous ensemble sélectionné. Les méthodes de type focus+contexte consistent à intégrer les parties agrégées (contexte) et les parties détaillées (focus) dans une seule et même vue. La plus connue de ces méthodes est sans doute le *fisheye* qui consiste à appliquer une distortion sur la visualisation de façon à agrandir la zone du focus en réduisant la zone du contexte.

Pour résoudre le problème (2), les séries temporelles peuvent être aggrégées selon leur proximité pour former une structure hiérarchique visualisable sous forme d'arbre. Cet arbre peut aider à sélectionner les niveaux de détail affichés.

La méthode que nous utilisons ici [**?**] repose sur 3 vues permettant de traiter les problèmes susmentionés. Une première vue (voir Fig. 2 cadre 1) permet de voir les séries agrégées sur l'ensemble de l'intervalle de temps, et de sélectionner un sous-intervalle (vue d'ensemble). Une seconde vue (voir Fig. 2 cadre 2) affiche les séries sur l'intervalle sélectionné (détail). Cette vue inclut un *fisheye*, i.e. une zone dans laquelle les séries sont détaillées (focus), entourée par des zones dans lesquelles les séries sont agrégées (contexte). La troisième vue (voir Fig. 2 cadre 3) permet de visualiser l'arbre d'agrégation, dans notre cas la hiérarchie des classes présentées dans la table 1 et de sélectionner les niveaux devant être utilisés pour les zones agrégées et pour les zones détaillées.

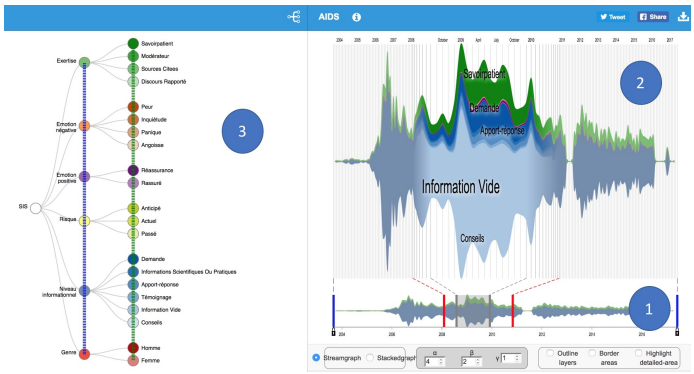


FIGURE *2 – Visualisation des séries temporelles avec un focus sur la période 2008-2011.*

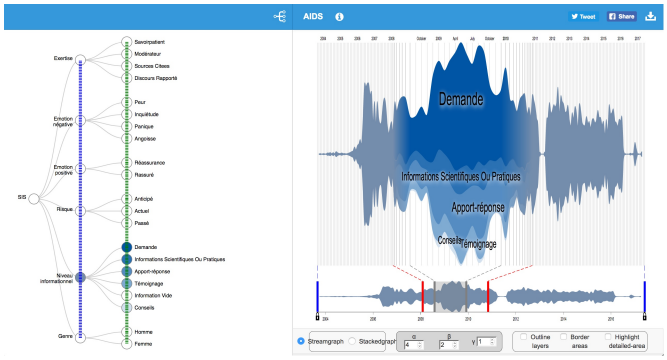


FIGURE *3 – Visualisation de la dimension "Niveau informationnel" dans laquelle les messages relatifs à la classe "Information vide" ont été filtrés grâce à l'arbre d'agrégation.*

TABLE *4 – Comparaison des performances des classifieurs.*

Stratégie	Précise				Imprécise			
jeux	emot_1	genre_1	expe_1	info_1	emot_2	expe_2	info_2	risq
LSVC	0.318	0.621	0.735	0.255	0.533	0.659	0.488	**0.638**
SSVC	**0.325**	0.624	0.735	0.242	0.532	0.422	0,000	0,440
MNB	0,020	0.631	0.701	0.175	0.472	0.662	0.406	0.603
BNB	0.058	**0.660**	0.725	0.280	**0.563**	**0.672**	0.371	0.603
SGDC	0.309	0.640	0.711	0.238	0.518	0.630	**0.490**	0.608
DT	0.244	0.533	0.664	0.206	0.468	0.625	0.362	0.564
AB	0.300	0.563	0.718	0.248	0.493	0.600	0.450	0.633
RF	0.0	0.650	**0.752**	0.0	0.116	0.379	0.0	0.440
KNC	0.227	0.533	0.677	**0.323**	0.439	0.589	0.281	0.461
BNN	0.718	0.945	0.966	0.372	0.841	0.894	0.883	0.883
MLP	0.662	0.945	0.966	0.356	0.745	0.901	0.909	**0.921**
DMLP	0.627	0.658	0.788	0.308	0.801	0.342	0.293	0.376
LSTM	**0.827**	**0.990**	**0.969**	**0.564**	**0.882**	**0.922**	**0.940**	0.906
BLSTM	0.776	0.927	0.953	0.521	0.855	0.906	0.898	0.902
GRU	0.502	0.684	0.884	0.355	0.441	0.805	0.745	0.805
LRCN	0.795	0.954	0.926	0.516	0.832	0.881	0.893	0.907
CNN	0.713	0.954	0.963	0.408	0.841	0.900	0.883	0.897
SCNN	0.741	0.954	0.959	0.436	0.829	0.905	0.893	0.892
MCNN	0.734	0.963	0.956	0.382	0.847	0.905	0.881	0.892

3 Résultats

3.1 Comparaison des modèles de classification

Le tableau 4 synthétise les résultats des différents classifieurs sur les différentes dimensions et selon les 2 straté-

gies. Comme l'on pouvait s'y attendre, les résultats pour les stratégies de classification précises sont inférieurs aux résultats pour les stratégies de classification imprécises. Par exemple, pour la dimension relative au niveau informationnel, l'exactitude pour la classification info_1 (stratégie précise) est au mieux de 0,564 alors que pour info_2 (stratégie imprécise) est de 0,906. Pour chaque type de classification, les meilleurs classifieurs traditionnels ont obtenu des résultats inférieurs aux classifieurs en apprentissage profond. Par exemple, pour la classification du genre, le meilleur classificateur traditionnel obtient une exactitude de 0,660 alors que le meilleur classifieur en apprentissage profond obtient une exactitude de 0,990. Pour 7 types de classification sur 8, la meilleure architecture est LSTM. Ce résultat est cohérant avec la littérature. Par exemple, [14] obtiennent également des résultats performants sur des textes avec LSTM. La classification du genre, de l'expertise, du risque et du niveau d'expérience obtiennent en apprentissage profond des exactitudes supérieures à 0,9 dans le cas imprécis et même parfois précis. Les résultats sont plus mitigés pour les émotions et le niveau informationnel qui correspondent aux classes très déséquilibrées, en particulier avec les classifieurs classiques. Par exemple, le classifieur RF ne fonctionne pas sur les classes déséquilibrées telles que emot_1 ou emot_2. Il serait intéressant de réannoter un nombre suffisant d'exemples pour rééquilibrer ces classes.

3.2 Visualisation

La visualisation obtenue à partir des 226 252 messages annotés automatiquement à partir du meilleur modèle (LSTM) est disponible sur le site du projet Multistream [11]. La figure 2 permet de repérer un pic d'activité du forum en 2009 et une majorité d'information de type "Information vide" révélant la part importante de messages de convivialité ou de soutien, les conseils étant également très recherchés dans des forums modérés par des professionnels de la prévention. La figure 3 est obtenue après sélection des classes à afficher dans l'arbre d'agrégation, en filtrant les messages de type "Information vide". On remarque alors que l'information de type "Demande" domine le pic de messages de 2009, ce qu'il n'était pas possible de visualiser avant le filtrage. Ces deux figures illustrent le processus de navigation dynamique dans le flux de messages, rendu possible par l'outil *MultiStream*.

Nous allons détailler dans la suite les 5 dimensions identifiées lors de la phase d'annotation manuelle (le genre, le niveau d'expertise, le niveau informationnel, le risque et les émotions) et illustrer avec des exemples ce qu'il est possible de repérer dans ces flots d'informations, et l'exploitation éventuelle qui peut en être faite.

Concernant la dimension *Genre*, on remarque qu'il y a plus de messages d'hommes que de femmes. Cette situation est spécifique au VIH, puisque les femmes sont souvent plus présentes sur les forums de santé [27]. Dans le cas du sida, les hommes ayant des relations sexuelles avec d'autres hommes (HARSAH), en termes de prévalence, sont les plus touchés par le VIH, ils sont aussi la cible des campagnes de prévention et sont bien informés des risques d'exposition au sida [10]. Ces résultats nous laissent penser que les forums constituent des médias alternatifs et complémentaires [31] indispensables pour des hommes plutôt moins acculturés aux messages de prévention, qui se renseignent pour des prises de risques plus exceptionnelles (par exemple dans le cadre de relations extra-conjugales hétéro ou homosexuelles).

Concernant la dimension *Niveau d'expertise*, les messages annotés montrent que se développent sur les forums de nouvelles formes d'expertise, hors de la sphère médicale et relayées par le biais des participants ou des modérateurs. Ces informations sont de première importance pour les professionnels (de santé et de la prévention) qui peuvent ainsi appréhender les connaissances qui circulent sur le VIH. Ainsi, la majorité des messages correspondent à des "savoirs patients" et dans une moindre proportion à des messages de modérateurs ou à des messages contenant des sources citées et des discours rapportés. Dans le cas du VIH, les patients apparaissent comme experts de leur propre maladie, de leur cas singulier bien qu'ils ne soient pas spécialistes de la maladie en général ni de la médecine. Ces derniers ont été conduits à côtoyer des professionnels de santé dont ils rapportent les propos, ce qui légitime les conseils qu'ils donnent aux autres participants : *"Merci de m'avoir répondu, c'est vrai vous m'avez bien rassurée ms je ne peux m'empecher de penser qu'il y a un risque , j'en ai parlé à 2 médecins qui travaillent avec moi , ils me disent que le risque est minime ms qu'il y en a un , ils ont l'habitude de soigner des séropositifs puisqu'ils accueillent des patients ayant eu un accident exposant au sang comme on dit chez nous"*. Les forums sont aussi le moyen de transmettre des sources spécialisées qui participent à la circulation des savoirs sur le VIH : *"En tout cas pour tout ce qui est des modes de transmission du vih tu trouveras tout ce qu'il te faut savoir ici : http ://www.sida-info service.org/forums ... hp4 ?t=3257[. . .]."*.

Concernant le *Niveau informationnel*, la majorité des messages sont vides, au sens où ils renvoient à des messages de convivialité (ex. *"Bonjour tous", "@ bientôt", "Merci++"*). Une fois ces messages filtrés, deux types d'informations apparaissent : des demandes ou des apports d'information associés à ces demandes. Ces messages concernent essentiellement des informations pratiques ou scientifiques : *"Bonjour, J'ai lu que les anticorps apparaissent generalement entre 20 et 45 jours apres un risque. Ca veut donc dire que ||nom|| n'est tres fiable qu'apres 6 semaines ? Je sens que je vais avoir la reponse sybilline "1 mois=bonne indication, 3 mois= sur", mais j'aimerais savoir ce que ca veut dire bonne indication (je suis negatif a 4 semaines avec*

11. http://advanse.lirmm.fr/multistream/ - jeu de données AIDS

©AfIA 2018

des ganglions et des courbatures partout... Et pouvezvous aussi me dire si le test ||nom|| detecte le VIH-1 et le VIH-2 ? Merci pour vos reponses!!", et dans une moindre mesure des témoignages relatant le vécu des porteurs du virus. Les participants au forum directement concernés par le VIH cherchent en premier lieu des informations relatives à cette maladie, aux symptômes, aux traitements. Ils cherchent également à mieux comprendre le discours des experts en amont d'un rendez-vous médical ou en aval pour trouver la définition de termes spécialisés ou confirmer le discours des professionnels de santé. Pour les proches non directement concernés par le virus, la recherche d'information permet de découvrir des conseils pratiques pour accompagner une personne malade.

Concernant le *risque*, il faut distinguer une dimension objective du risque, incarnée dans une connaissance plus ou moins partagée des facteurs de prise de risque, qui se confronte à des perceptions variables du risque en lien notamment avec l'expérience personnelle de l'individu, son environnement social, affectif et culturel. La dimension subjective du risque motive l'essentiel des demandes : *"Bonjour, ma copine est séropositive depuis peu, mon 1er test est négatif. J'attends 3 mois pour confirmer le résultat. Comme j'entends tout et n'importe quoi a propos des fellations, je voulais savoir si je risquais un risque ou pas lors d'une fellation dans mon cas. Merci pour votre réponse. C'est nouveau pour moi et je gère assez mal la situation."* Ainsi, les trois quarts des messages portent sur un risque passé, peu de personnes posant des questions avant une situation présumée risquée. Dans les forums où priment l'émotion, c'est la panique, le stress qui dominent. Cette situation engendre de la part des répondants des messages de réassurance, la catégorie d'émotion qui est la plus représentée en termes de pourcentage d'apports d'information. En effet, de nombreux messages d'encouragements, par exemple lors de l'attente d'un dépistage viennent soutenir les membres ayant pris un risque. Peu de messages ont cependant été étiquetés selon le risque lié à l'alcool et à la prise de substance psychotrope. Ce constat est certainement lié au fait qu'un forum institutionnel de type SIS n'est pas le lieu pour évoquer les pratiques à risques autres que celles liées à la sexualité.

3.3 Limites de l'étude

Notre approche préliminaire présente de nombreuses limites que nous détaillons dans cette section.

Concernant la phase d'annotation, la grille utilisée pourrait être retravaillée. La dimension relative au *Niveau Informationnel* pourrait être scindée en deux pour distinguer les messages relevant de l'interaction (demande ou apport de réponse à une question) des messages uniquement destinés à l'apport d'information (Information Scientifique et pratique, témoignage et conseil) ou encore les messages vides au niveau informationnel. Cela est tout à fait possible car la visualisation utilisée permet d'afficher plusieurs niveaux de hiérarchie.

Concernant la phase de classification, la principale limite est que nous avons utilisé uniquement l'exactitude comme métrique pour évaluer la performance des classifieurs. Le rappel et la précision sont deux mesures importantes pour évaluer la qualité d'un classifieur et devrait être prises en considération. Une autre limitation porte sur l'interprétation de l'apprentissage profond. Ces modèles sont des "boîtes noires". Ils ne fournissent pas d'explication même si la prédiction est efficace [34]. Cependant, malgré un nombre de messages manuellement annotés limité, cette étude a montré que la performance des modèles traditionnels était beaucoup plus faible que la performance de l'apprentissage profond. De nouvelles techniques d'apprentissage profond sont étudiées pour faciliter l'interprétation de tels modèles [20, 19] et pourraient s'avérer judicieuses dans ce contexte.

Nous avons également identifié des limites liées à la temporalité utilisée pour la visualisation de l'ensemble des données du site SIS. Actuellement, la temporalité est le mois. Or, la période temporelle d'interaction est généralement le jour ou la semaine dans le contexte des échanges sur les forums. Une possibilité de focus sur une période plus courte permettrait de repérer des épisodes liés à des questions sur une prise de risque et le délai pour se faire dépister. En effet, 48h après une prise de risque, il existe un traitement préventif qui peut être suggéré par les internautes. Les résultats du premier test permettant de savoir s'il y a eu contamination sont obtenus entre 6 semaines et 90 jours. Cette période s'accompagne de messages contenant des émotions négatives, de la réassurance et dans la majorité des cas se finit par des messages de la personne rassurée comme nous avons pu l'observer dans de nombreux fils de discussion. De même, la visualisation pourrait être améliorée par l'affichage d'informations supplémentaires comme l'heure des messages. Cette information est essentielle notamment lorsque l'on s'intéresse aux émotions exprimées dans les messages. Par exemple, les angoisses s'expriment généralement la nuit. La quantification de ces phénomènes est important pour le site SIS en terme de planification de leurs équipes de modérateurs.

Pour finir, une dernière limite de notre étude est liée à la généralisation de nos résultats. En effet, la tâche est très spécifique au sujet d'étude, le VIH, au type de textes que sont les messages des forums, ce qui rend difficile la généralisation de notre approche. Toutefois, notre méthodologie et les résultats peuvent être utilisés comme référence pour d'autres études sur l'identification automatique de catégories à partir de données sociales.

4 Conclusion et perspectives

Notre étude a souligné l'efficacité des architectures d'apprentissage profond pour prédire des thèmes sur des données de forum. La visualisation originale sous forme de StreamGraph permet d'explorer de manière efficace tout en quantifiant les différentes catégories de

messages. La faisabilité de notre approche peut conduire à de nouvelles applications en santé basées sur les médias sociaux destinés aux patients et aux professionnels de santé.

Nos résultats montrent clairement le besoin d'études capables d'analyser automatiquement des forums et d'en extraire des informations utiles. Nous proposons l'utilisation de l'apprentissage automatique et de la visualisation interactive pour relever ces défis. Cette étude reste préliminaire. Une étude plus approfondie sur les différents types de préparation des données, paramètres des algorithmes, modèles d'apprentissage permettrait d'affiner l'interprétation des résultats de la phase 3. En particulier, différents regroupement de classes et l'interprétation des liens entre classes serait pertinent. Quels sont les sentiments associés à une prise de risque ? Est-ce que les réponses apportées dans les fils sont rassurantes ou satisfaisantes ? Comment s'informent les participants dans les forums ? Pour finir, une étude poussée sur l'utilité et l'utilisabilité de la visualisation présentée en phase 4 est également nécessaire. Nous suggérons 4 perspectives.

Premièrement, nous prévoyons d'entreprendre une analyse à grande échelle en utilisant une collection de médias sociaux plus large (Autres forums, Facebook, Twitter, ...). Cette analyse inclura l'application de méthode d'apprentissage non supervisé de type Latent Dirichlet Allocation (LDA) [29, 41] pour extraire les thèmes émergeant des discussions et l'exploration du style linguistique des différents utilisateurs [44, 43, 40]. Une attention particulière sera portée sur l'identification d'une typologie des risques encourus en lien avec le VIH (e.g. consommation de drogue, comportements sexuels atypiques, etc).

Deuxièmement, nous pensons que lorsque les ensembles de données sont petits, l'apprentissage est difficile. Une amélioration significative serait la mise en œuvre de techniques d'apprentissage actif [26]. En effet, dans ce type de tâche, il est important d'optimiser les informations disponibles afin que les systèmes de classification puissent les utiliser le plus efficacement possible pendant la phase d'apprentissage tout en préservant l'acquisition de nouveaux échantillons étiquetés [11]. L'utilisation de la visualisation pour guider les annotateurs vers des messages à annoter pourrait s'avérer également intéressante.

Troisièmement, au sein d'un ensemble de données suffisamment important, nous pouvons tirer parti des modèles d'apprentissage automatique pour utiliser des fonctionnalités plus complexes pour caractériser les utilisateurs qui postent ces messages. Nous suggérons de mettre l'accent sur les groupes d'utilisateurs, y compris les professionnels de la santé, les célébrités, le grand public et les associations. Cela nous amènera à comprendre quel groupe d'utilisateurs est important, peut jouer le rôle d'influenceur, les incitant à partager leurs messages, à les aimer et à leur répondre.

Pour finir, nous prévoyons d'étudier la distribution temporelle des messages pour nous concentrer sur la dynamique des thématiques au fils du temps. Nous pouvons étudier les corrélations temporelles entre les réactions des internautes et les événements du monde réel comme les soirées de type Sidaction. Cette analyse exploratoire pourrait aider à identifier les facteurs contribuant à la sensibilisation. Au-delà, nous pouvons également analyser la répartition géographique des messages.

Ce type d'étude est importante pour convaincre les parties prenantes, les professionnels de la santé et le grand public de s'impliquer et d'utiliser le Web 3.0 comme intelligence collective pour repousser les maladies telles que le VIH.

5 Remerciements

Ce travail s'intègre dans le projet #AIDS et a été soutenu par une subvention ANRS [12] en 2016. Les auteurs souhaitent remercier les gestionnaires du site Sida-Info-Service pour le partage des données et ses participants pour leur engagement à combattre le VIH.

Références

[1] *Joint Named Entity Recognition and Disambiguation* (September 2015).

[2] AIGNER, W., MIKSCH, S., SCHUMANN, H., AND TOMINSKI, C. *Visualization of Time-Oriented Data*. Springer, 2011.

[3] ANDOR, D., ALBERTI, C., WEISS, D., SEVERYN, A., PRESTA, A., GANCHEV, K., PETROV, S., AND COLLINS, M. Globally normalized transition-based neural networks, 2016. cite arxiv :1603.06042.

[4] BACCOUCHE, M., MAMALET, F., WOLF, C., GARCIA, C., AND BASKURT, A. Sequential deep learning for human action recognition. In *Proceedings of the Second International Conference on Human Behavior Unterstanding* (Berlin, Heidelberg, 2011), HBU'11, Springer-Verlag, pp. 29–39.

[5] BYRON, L., AND WATTENBERG, M. Stacked Graphs - Geometry & Aesthetics. *IEEE Transactions on Visualization and Computer Graphics 14*, 6 (2008), 1245–1252.

[6] CLAVIER, V., AND PAGANELLI, C. Une approche méthodologique croisée du traitement des données de la recherche : le cas d'un corpus d'échanges issus de médias sociaux dans le domaine de la santé. In *Colloque COSSI 2017, Méthodes et stratégies de gestion de l'information par les organisations : des "Big Data" aux "Thick Data", 85ème congrès de l'ACFAS, Université McGill* (2017).

[7] COLLOBERT, R., WESTON, J., BOTTOU, L., KARLEN, M., KAVUKCUOGLU, K., AND KUKSA, P. Natural language processing (almost) from scratch. *J. Mach. Learn. Res. 12* (Nov. 2011), 2493–2537.

12. http://www.anrs.fr/fr

©AfIA 2018

[8] CUENCA, E., SALLABERRY, A., WANG, F. Y., AND PONCELET, P. MultiStream : A Multiresolution Streamgraph Approach to Explore Hierarchical Time Series. *IEEE Transactions on Visualization and Computer Graphics*, to appear (2018).

[9] CUI, W., LIU, S., TAN, L., SHI, C., SONG, Y., GAO, Z., QU, H., AND TONG, X. TextFlow : Towards Better Understanding of Evolving Topics in Text. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (2011), 2412–2421.

[10] DE OLIVEIRA, J.-P. *Communication publique et formes de gouvernabilité contemporaines de l'Etat : le cas de l'homosexualité dans les campagnes de prévention du sida en France (1987-2007)*. PhD thesis, Université Stendhal, Grenoble, 2012.

[11] DUCOFFE, M., AND PRECIOSO, F. QBDC : query by dropout committee for training deep supervised architecture. *CoRR abs/1511.06412* (2015).

[12] HAVRE, S., HETZLER, E., AND NOWELL, L. ThemeRiver : Visualizing Theme Changes over Time. In *Proceedings of the IEEE Symposium on Information Visualization* (2000), IEEE, pp. 115–123.

[13] HE, L., LEE, K., LEWIS, M., AND ZETTLEMOYER, L. Deep semantic role labeling : What works and what's next. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics* (2017).

[14] JOHNSON, R., AND ZHANG, T. Supervised and semi-supervised text categorization using lstm for region embeddings. In *Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48* (2016), ICML'16, JMLR.org, pp. 526–534.

[15] KALCHBRENNER, N., GREFENSTETTE, E., AND BLUNSOM, P. A convolutional neural network for modelling sentences. In *ACL (1)* (2014), The Association for Computer Linguistics, pp. 655–665.

[16] KIM, Y. Convolutional neural networks for sentence classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL* (2014), A. Moschitti, B. Pang, and W. Daelemans, Eds., ACL, pp. 1746–1751.

[17] KUCHER, K., AND KERREN, A. Text visualization techniques : Taxonomy, visual survey, and community insights. In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis)* (2015), pp. 117–121.

[18] KUMAR, A., IRSOY, O., ONDRUSKA, P., IYYER, M., BRADBURY, J., GULRAJANI, I., ZHONG, V., PAULUS, R., AND SOCHER, R. Ask me anything : Dynamic memory networks for natural language processing. In *Proceedings of The 33rd International Conference on Machine Learning* (New York, New York, USA, 20–22 Jun 2016), M. F. Balcan and K. Q. Weinberger, Eds., vol. 48 of *Proceedings of Machine Learning Research*, PMLR, pp. 1378–1387.

[19] LIPTON, Z. C. The mythos of model interpretability. *CoRR abs/1606.03490* (2016).

[20] LIU, W., WANG, Z., LIU, X., ZENG, N., LIU, Y., AND ALSAADI, F. E. A survey of deep neural network architectures and their applications. *Neurocomputing 234* (2017), 11 – 26.

[21] MIKOLOV, T., KARAFIÁT, M., BURGET, L., CERNOCKÝ, J., AND KHUDANPUR, S. Recurrent neural network based language model. In *INTERSPEECH* (2010), T. Kobayashi, K. Hirose, and S. Nakamura, Eds., ISCA, pp. 1045–1048.

[22] MIKOLOV, T., SUTSKEVER, I., CHEN, K., CORRADO, G. S., AND DEAN, J. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems 26*, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2013, pp. 3111–3119.

[23] MIN, S., LEE, B., AND YOON, S. Deep learning in bioinformatics. *CoRR abs/1603.06430* (2016).

[24] MUNZNER, T. *Visualization Analysis and Design*. A.K. Peters visualization series. A K Peters, 2014.

[25] NAKOV, P., RITTER, A., ROSENTHAL, S., SEBASTIANI, F., AND STOYANOV, V. Semeval-2016 task 4 : Sentiment analysis in twitter. In *Proceedings of the 10th International Workshop on Semantic Evaluation* (2016), pp. 1–18.

[26] OLSSON, F. A literature survey of active machine learning in the context of natural language processing. Tech. Rep. T2009 :06, 2009.

[27] PAGANELLI, C., AND CLAVIER, V. Le forum de discussion : une ressource informationnelle hybride entre information grand public et information spécialisée. *Yasri-Labrique Eleonore. Les forums de discussion : agoras du XXIe siècle ? Théories, enjeux et pratiques discursives, L'harmattan (collection Langue et Parole)* (2011), 39–55.

[28] PAGANELLI, C., AND CLAVIER, V. S'informer via des médias sociaux de santé : quelle place pour les experts ? 141–143.

[29] PENNACCHIOTTI, M., AND GURUMURTHY, S. Investigating topic models for social media user recommendation. In *Proceedings of the 20th International Conference Companion on World Wide Web* (New York, NY, USA, 2011), WWW '11, ACM, pp. 101–102.

[30] RAMOS, J. Using tf-idf to determine word relevance in document queries, 1999.

[31] RENAHY, E., AND CHAUVIN, P. Internet uses for health information seeking : A literature review. *Revue Epidémiologique de Santé Publiqu 54*, 3 (2006), 263–275.

[32] ROSENTHAL, S., NAKOV, P., KIRITCHENKO, S., MOHAMMAD, S., RITTER, A., AND STOYANOV, V. Semeval-2015 task 10 : Sentiment analysis in twitter. In *Proceedings of the 9th International Workshop on Semantic Evaluation* (2015), pp. 451–463.

[33] SCHMIDHUBER, J. Deep learning in neural networks : An overview. *Neural Networks 61* (2015), 85–117. Published online 2014; based on TR arXiv :1404.7828 [cs.NE].

[34] SHWARTZ-ZIV, R., AND TISHBY, N. Opening the black box of deep neural networks via information. *CoRR abs/1703.00810* (2017).

[35] SOCHER, R., PERELYGIN, A., WU, J. Y., CHUANG, J., MANNING, C. D., NG, A. Y., AND POTTS, C. Recursive deep models for semantic compositionality over a sentiment treebank. In *EMNLP* (2013), vol. 1631, p. 1642.

[36] SRIVASTAVA, R. K., GREFF, K., AND SCHMIDHUBER, J. Training very deep networks. In *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2* (Cambridge, MA, USA, 2015), NIPS'15, MIT Press, pp. 2377–2385.

[37] SUKHBAATAR, S., SZLAM, A., WESTON, J., AND FERGUS, R. End-to-end memory networks. In *Advances in Neural Information Processing Systems 28*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, Eds. Curran Associates, Inc., 2015, pp. 2440–2448.

[38] SUN, G., WU, Y., LIU, S., PENG, T.-Q., ZHU, J. J. H., AND LIANG, R. EvoRiver : Visual Analysis of Topic Coopetition on Social Media. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 1753–1762.

[39] VINYALS, O., KAISER, L. U., KOO, T., PETROV, S., SUTSKEVER, I., AND HINTON, G. Grammar as a foreign language. In *Advances in Neural Information Processing Systems 28*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, Eds. Curran Associates, Inc., 2015, pp. 2773–2781.

[40] WANG, S., J.PAUL, M., AND DREDZE, M. Exploring health topics in chinese social media : an analysis of sina weibo. *AAAI Workshop on the World Wide Web and Public Health Intelligence 23* (2014), 20–23.

[41] WANG, Y., AGICHTEIN, E., AND BENZI, M. Tm-lda : Efficient online modeling of latent topic transitions in social media. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (New York, NY, USA, 2012), KDD '12, ACM, pp. 123–131.

[42] WU, Y., LIU, S., YAN, K., LIU, M., AND WU, F. OpinionFlow : Visual Analysis of Opinion Diffusion on Social Media. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 1763–1772.

[43] ZENG, Q. T., AND TSE, T. Viewpoint paper : Exploring and developing consumer health vocabularies. *JAMIA 13*, 1 (2006), 24–29.

[44] ZHAN, Y., LIU, R., LI, Q., LEISCHOW, S., AND ZENG, D. Identifying topics for e-cigarette user-generated contents : a case study from multiple social media platforms. *J Med Internet Res 19*, 1 (2017), e24.

[45] ZHU, M., ZHANG, Y., CHEN, W., ZHANG, M., AND ZHU, J. Fast and accurate shift-reduce constituent parsing, 2013.

©AfIA 2018

Optimisation en temps réel de la gestion des convergences ferroviaires

Valentina Pozzoli[1] François Ramond[1]

[1]SNCF Innovation et Recherche
40, avenue des Terroirs de France
75611 Paris Cedex

Valentina.Pozzoli@sncf.fr
Francois.Ramond@sncf.fr

Résumé

L'exploitation ferroviaire en zone dense est une tâche complexe. La densité des circulations provoque une propagation très rapide des retards, même de petite amplitude. Afin de prendre des décisions d'exploitation rapides et optimales il est nécessaire de se doter d'outils d'aide à la décision performants.
Dans cet article, nous présentons un système d'aide à la gestion optimale en temps réel des zones de convergences ferroviaires, afin d'améliorer la ponctualité des voyageurs de la ligne. Nous détaillons les enjeux liés au projet, ainsi que les choix de modélisation et l'architecture du prototype. Le système a été testé avec succès sur un site pilote en Île-de-France : nous montrons les résultats de ces expérimentations et donnons les perspectives pour la suite du projet.

Mots Clef

Aide à la décision, Optimisation, Exploitation ferroviaire, Algorithmes en temps réel, Convergences ferroviaires

Abstract

Railway traffic management in mass transit system is a complex task. Trains run so frequently and so close to each other that delays propagate very fast. Operators have to take efficient decisions depending on a multitude of factors, that impact a large number of trains, and thus of passengers. Artificial intelligence methods come in help in the decision process.
In this paper, a real-time decision-support tool for the optimal management of junction areas is presented. It proposes the best order of trains to run through a junction, such that the passenger punctuality is maximized. Details are given on the modelling and the architecture of the tool, as well as on the goals and ambitions of the project. The results of promising experiments on a case-study on a junction in a Parisian area are presented.

Keywords

Decision-support, Optimization, Traffic management, Mass transit, Real-time algorithms, Railway junctions

1 Introduction

L'exploitation ferroviaire en zone dense, où la fréquence des trains est élevée et le nombre de voyageurs important, est un vrai défi dont l'enjeu principal est la qualité du service offert aux voyageurs. L'optimisation de l'exploitation ferroviaire relève de plusieurs facteurs avec des échéances temporelles différentes. En amont de la production se situent, par exemple, la conception de plans de transport robustes et résilients, la maintenance des voies et du matériel, l'affectation des conducteurs et la gestion des rames. En gestion opérationnelle, la mise en œuvre de mesures de régulation efficaces est nécessaire afin d'optimiser la production : changements de la desserte ou des voies utilisées, adaptation d'offre, etc. La gestion des zones de convergence figure parmi ces mesures de régulation.

Les zones de convergences sont des endroits du réseau ferroviaire où deux ou plusieurs voies se réunissent. L'exemple le plus simple, qui sera étudié au cours de cet article, est montré en Figure 1. La gestion des zones de convergence consiste à choisir l'ordre de passage des trains lors que deux trains de deux branches différentes s'apprêtent à rentrer dans la zone de convergence. Bien que la thématique d'optimisation de l'ordre de passage existe pour toutes zones de convergence, en zone dense le choix est particulièrement impactant : en effet, dans un réseau dense les retards se propagent facilement sur les trains suivants. Chaque décision de régulation peut donc avoir un impact sur un grand nombre de trains, et de conséquence sur un nombre important de voyageurs. Ces décisions dépendent de nombreux paramètres : par exemple, les retards des trains, leur desserte, le nombre de passagers dans les trains et sur les quais. En plus, les choix doivent être pris rapidement par le régulateur : pour ces raisons il est nécessaire de se doter d'outils d'aide à la décision performants qui puissent venir en support du travail des régulateurs.

En Ile-de-France, SNCF Mobilités, via son activité Transilien, et RATP sont les principaux opérateurs de transport ferroviaire. Ils opèrent par le compte d'Île-de-France Mobilité, qui définit également les objectifs de performance des opérateurs de transport. Transilien fait circuler 6200 trains chaque jour, qui transportent au quotidien 3 millions de voyageurs, nombre qui connait une augmentation de 3% par an en moyenne (source Transilien). L'amélioration de la 'ponctualité des

 ©AfIA 2018

voyageurs', premier élément pour la satisfaction des clients, est l'objectif principal de Transilien et le critère utilisé par Île-de-France Mobilité pour évaluer la performance d'une ligne.

Dans ce contexte, la gestion des zones de convergence, est particulièrement complexe et impactant. En effet, ces zones sont spécialement sollicitées car les trains des deux branches doivent y passer : la fréquence de passage peut aller jusqu'à 32 trains par heure. L'ordonnancement des trains dans une zone de convergence peut être modifié par les opérateurs. Bien que les horaires théoriques indiquent l'ordre de passage des trains à une zone de convergence, cet ordre peut être remis en cause par les aléas : la décision optimale doit considérer l'impact de ce choix sur toute la ligne d'un point de vue des voyageurs concernés.

La Direction de l'Innovation et de la Recherche de SNCF a développé pour Transilien un outil d'aide à la gestion des zones de convergence en temps réel. Cet outil préconise l'ordre de passage des trains à la convergence permettant de maximiser la performance globale de la ligne, définie par le critère de 'ponctualité des voyageurs'.

L'article est organisé de la façon suivante : la Section 2 est consacrée à une présentation succincte des travaux déjà existants sur le sujet ; la Section 3 détaille le critère d'optimisation et les enjeux du projet, la Section 4 décrit la modélisation du problème et l'architecture de l'outil ; la Section 5 présente les résultats de l'expérimentation menée dans un contexte opérationnel sur un site pilote ; enfin, des conclusions et perspectives sont exposées en Section 6.

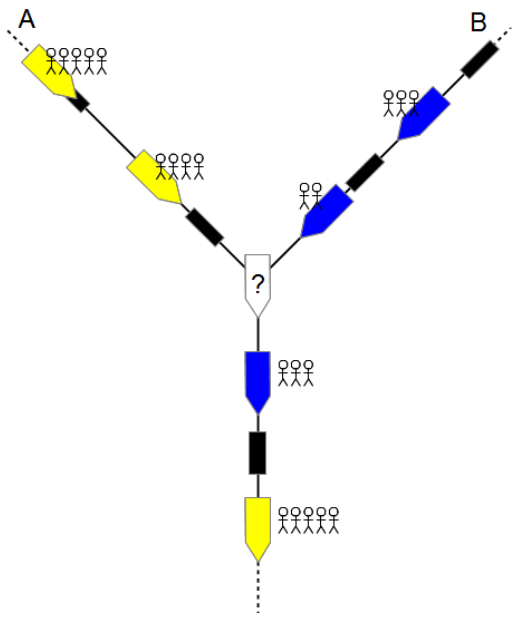


Figure 1 - Schéma d'une zone de convergence

2 Travaux antérieurs

La gestion du trafic en temps réel est un problème largement abordé dans la littérature. Une revue extensive des approches pour réduire les conséquences des incidents est présentée dans [1], en distinguant entre petits aléas et perturbations importantes. Les actions à mettre en place sont différentes dans les deux cas, et incluent la replanification d'itinéraires, la gestion des correspondances, la replanification des engins ou des conducteurs. Au niveau des adaptations du plan de transport en temps réel, les stratégies utilisées en zone dense sont la rétention d'un train en gare, l'ajustement de la vitesse, le changement de desserte, et l'ordre de passage en certains points du réseau. L'ordonnancement des trains à une convergence est un cas particulier de cette dernière stratégie de replanification.
L'article [2] résume les modèles existants dans la planification et replanification ferroviaire, dont les décisions d'ordonnancement à une convergence, ou de replanification d'itinéraires. Une autre revue des modèles de contrôle de trafic en temps réel est proposée par [3].

L'optimisation de l'exploitation du point de vue des voyageurs, et non pas que des retards des trains, est traitée dans plusieurs articles : par exemple dans [4], les horaires d'arrivée et de départ des métros sont choisis afin de minimiser les temps d'atteintes et de trajet, ainsi que la congestion; [5] proposent un algorithme de replanification qui minimise l'insatisfaction des voyageurs.

Plusieurs articles traitent spécifiquement le problème de gestion des convergences ferroviaires : [6] propose un modèle d'ordonnancement optimal des trains aux convergences afin de minimiser une somme pondérée des retards des trains aux gares desservies. Ce modèle est résolu grâce à un algorithme à évolution différentielle et, sur les instances de test, permet de réduire les retards par apport à une stratégie du type premier entré, premier sorti. Un modèle de replanification en temps réel qui inclut l'optimisation de la vitesse des trains afin de minimiser les retards est proposé dans [7]. Une comparaison entre différentes stratégies de résolution des conflits aux convergences est décrite dans [8]: premier arrivé premier servi, priorité selon le type de train, et minimisation du nombre et de la durée des retards grâce à une heuristique. Les trois scénarios sont évalués grâce à des simulations.

Les travaux présentés dans cet article proposent un approche diffèrent car centré sur une gestion optimale du point de vue de la performance globale de la ligne, en considérant le point de vue des voyageurs, et non pas que le retard des trains.

3 Critère d'optimisation

L'objectif de nos travaux est d'optimiser la gestion d'une zone de convergence afin de maximiser la performance globale de la ligne.
Aujourd'hui, en cas de conflit dans une zone de convergence, la décision d'ordonnancement des trains est prise au cas par cas grâce à l'expertise humaine du régulateur, et dépend principalement des retards de deux trains en conflit, ainsi que de leurs missions. En cas de perturbation, le régulateur a une charge de travail élevée et doit prendre des décisions rapides qui ont un impact majeur sur toute la ligne. L'idée est de venir en aide aux régulateurs, en leur fournissant en outil qui puisse à la fois les décharger d'une partie de leur charge cognitive, leur donner des informations pour mieux analyser la situation, et leur proposer le meilleur ordonnancement du point de vue de la performance globale de la ligne.

Un des principaux indicateurs de la performance globale d'une ligne est appelé 'ponctualité voyageurs'. Ce critère mesure le nombre de voyageurs arrivés à leur gare de destination moins de 5 minutes après leur heure d'arrivée théorique. Il est utilisé pour évaluer la qualité de service de SNCF par Île-de-France Mobilités, et c'est le critère d'optimisation choisi dans l'outil d'aide à la gestion des zones de convergence.
Pour calculer la ponctualité voyageurs, des données statistiques sur le nombre de voyageurs pour chaque couple origine-destination par heure sont utilisées afin d'affecter les voyageurs aux différents trains, en se basant sur les horaires théoriques. Les données sur le nombre de voyageurs sont issues de comptages manuels des voyageurs effectués régulièrement sur les lignes concernées. Grâce à ces données et aux horaires théoriques des trains, il est possible de calculer l'horaire théorique d'arrivée de chaque passager à la gare d'origine et à la gare de destination. Une fois les trains circulés, les horaires réels d'arrivée des voyageurs à destination sont calculés grâce aux horaires réels de passage des trains par les différentes gares. On considère que chaque voyageur arrive en gare à l'horaire théorique prévu, et monte ensuite dans le premier train qui dessert sa destination. Par la suite, on calcule l'indicateur de ponctualité voyageurs comme le pourcentage de voyageurs arrivant à la gare de destination avec moins de 5 minutes de retard par rapport à l'horaire théorique.

L'indicateur de ponctualité voyageurs a pour but de s'affranchir de la simple notion de retard du train pour représenter la perception du voyageur. En effet, en zone dense, où la fréquence des trains est élevée, l'horaire théorique du train réellement emprunté a peu d'importance pour les voyageurs par rapport à d'autres paramètres tels que le temps d'atteinte en gare et le temps de trajet.

Le critère de ponctualité voyageurs diffère de celui du simple retard des trains pour plusieurs aspects :

- Le retard d'un train transportant beaucoup de voyageurs a plus d'impact sur le critère de ponctualité voyageurs par rapport au même retard d'un train moins chargé.
- Il est possible d'avoir des voyageurs en retard dans un train à l'heure (par exemple, si un train est supprimé et que celui d'après et à l'heure, les voyageurs qui voulaient prendre le premier train seront en retard).
- Il est également possible d'avoir des voyageurs à l'heure dans un train en retard (notamment si un train roule à l'horaire du train suivant, donc en retard, les voyageurs souhaitant prendre le deuxième train pourront monter dans le premier est être à l'heure).

Ces aspects rendent difficile pour les opérateurs une prise de décision en fonction du critère de ponctualité voyageurs, ce qui justifie le besoin d'un outil d'aide à la décision. En effet, selon le choix d'ordonnancement les voyageurs pourront être affectés à des trains différents, et le choix peut impacter les retards des trains suivants (et donc l'horaire d'arrivée à destination des voyageurs qui voudrait monter dans ces trains). La présence de missions directes et de missions omnibus, avec des temps de parcours différents, complexifie d'avantage la compréhension des impacts du choix sur la ponctualité voyageurs.

4 Modélisation du problème et architecture

La Direction de l'Innovation et de la Recherche a développé un outil visant à aider le choix d'ordonnancement des trains du point de vue de la ponctualité voyageurs. L'outil calcule la ponctualité voyageurs associée aux différents scénarios d'ordonnancement des trains en se basant sur un flux de géolocalisation des trains en temps réel. Les résultats sont restitués via une interface homme-machine (IHM). L'architecture du projet est synthétisée en Figure 2.

En entrée nous avons des données structurelles et des données temps réel. Les *données structurelles* incluent la modélisation du réseau et des paramètres de circulations tels que les espacements minimaux entre les trains, le plan de transport théorique et les données statistiques sur le nombre des voyageurs par heure par couple origine/destination. Certaines de ces données sont fixes pour un calcul donné, comme la modélisation du réseau, d'autre sont journalières, comme le plan de transport théorique. La modélisation du réseau inclut des informations sur les zones aux alentours de la convergence (voir Figure 3). En particulier, il contient une définition de la zone de décision, c'est-à-dire la zone en amont de la convergence où les trains doivent se trouver pour qu'il soit possible et pertinent d'effectuer un réordonnancement.

Les *données temps réel* incluent les messages de géolocalisation des trains et les messages déclenchés par les actions de régulation telles que les adaptations de l'offre.

©AfIA 2018

La localisation des trains se fait grâce à un système de balises placées sur les rails : le passage d'un train sur une balise déclenche un message de géolocalisation. Cela signifie que la connaissance de la position des trains n'est pas connue en continue, mais de façon discrète : une fois le train passé sur une balise, il n'est pas possible de connaitre sa position ou sa vitesse jusqu'au passage par la balise suivante.

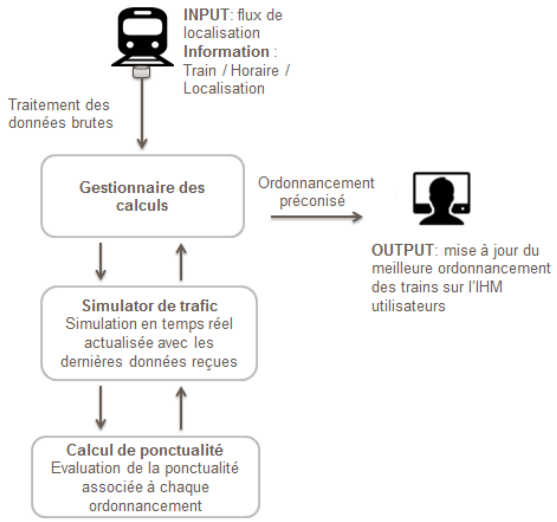


Figure 2 - Architecture du logiciel

Toutefois, une estimation de la position d'un train à un instant donné est possible en se basant sur l'horaire de passage à la dernière balise, sur les temps de trajet théorique et sur la position des autres trains sur le réseau.

À la réception de chaque nouveau message de localisation d'un train, l'outil met à jour la position des trains dans le réseau, identifie les différentes possibilités d'ordonnancement pour les trains concernés et simule l'impact de chaque possibilité sur les horaires des trains de l'instant présent jusqu'à la fin de la journée. Cela permet au module de ponctualité de calculer la ponctualité voyageurs pour chaque scénario d'ordonnancement, et d'identifier le meilleur choix.

Tout d'abord le module de *gestion des calculs* reçoit les données de localisation et identifie les trains présents dans les différentes zones du réseau montrée en Figure 3. Les trains dans la 'zone de décision' sont potentiellement concernés par la décision d'ordonnancement. Bien que plusieurs trains de chaque branche puissent être dans la zone de décision à un instant donné, nous considérons que le réordonnancement n'est possible qu'entre les deux les plus proches de la convergence. Deux situations sont donc à examiner (en se référant à la Figure 3) : le train A passe avant le train B, et le train B passe devant le train A.
Le module de gestion des calculs transmet au *simulateur de trafic* l'information des trains concernés par le réordonnancement, ainsi que l'information sur le dernier message de localisation reçu. Le simulateur met à jour les positions des trains grâce au message de localisation, et une simulation est ensuite lancée pour chacun de scénarios d'ordre de passage. Ces simulations permettent d'évaluer l'impact du choix d'ordonnancement sur tous les trains de la ligne. La simulation est menée grâce à un simulateur de trafic développé à la Direction de l'Innovation et de la Recherche de la SNCF [9]. Le simulateur prend en entrée les horaires théoriques et les horaires réels de localisation des trains pour calculer les circulations futures des trains du début jusqu'à la fin de la journée.

Le *simulateur* retourne l'information sur l'estimation des horaires de passage à chaque gare pour tous les trains de la ligne. Le module de *calcul de la ponctualité* utilise cette information pour évaluer la ponctualité voyageurs telle que décrite dans la Section 3. Il retourne la valeur de l'indicateur de ponctualité voyageurs pour les deux scénarios d'ordonnancement (A passe devant B et B passe devant A).
Les informations sont par la suite transmises au module de gestion des calculs qui les élabore et les transmette à

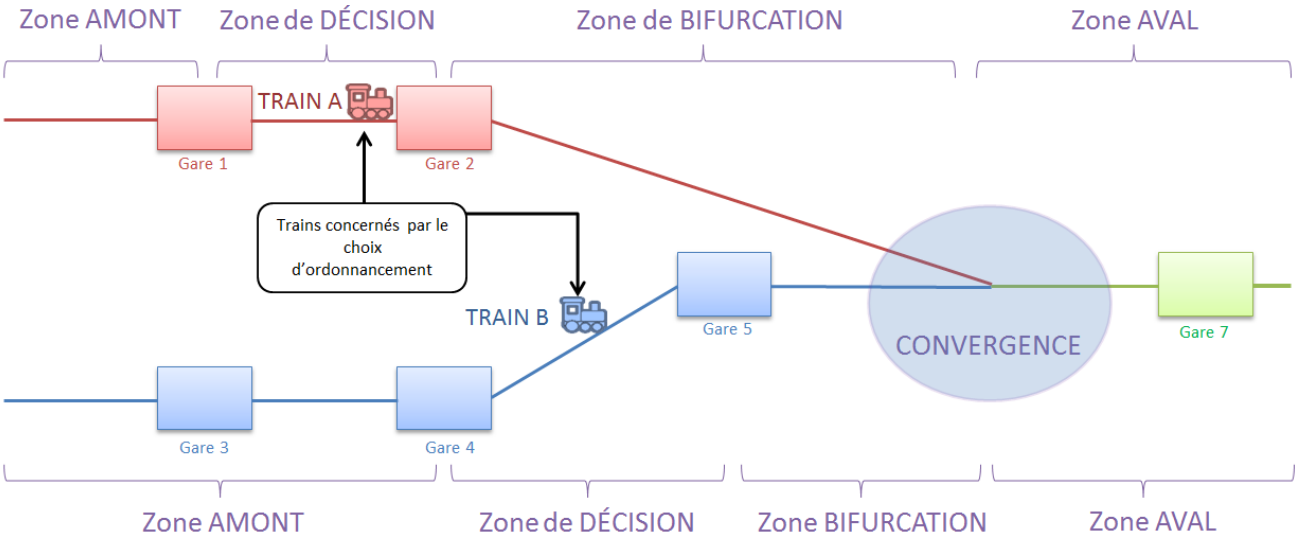


Figure 3 - Schéma des zones aux alentours de la convergence

l'IHM. L'interface a été élaborée en concertation avec des experts en design et en sciences cognitives, ainsi qu'avec les utilisateurs, afin de donner les indicateurs pertinents de façon simple et immédiate. Les enjeux principaux dans l'élaboration de l'interface ont été d'une part la nécessité d'expliquer le choix proposé par l'outil, et d'autre part de ne pas surcharger l'utilisateur avec trop d'informations. Cela est particulièrement crucial dans un contexte opérationnel où la charge de l'opérateur est élevée, notamment en cas de perturbations.

5 Cas d'étude et expérimentations

Afin d'éprouver la qualité des préconisations de l'outil pour l'amélioration de la ponctualité voyageurs, nous avons mené des expérimentations sur le périmètre de la ligne L et du RER A (voir Figure 4). L'outil propose en continu une préconisation d'ordre de passage à la convergence pour les deux trains dans la zone de prise de décision: le prochain train au départ de Poissy et le premier train en aval de la gare de Conflans Fin d'Oise. Ces trains sont marqués en rouge dans la Figure 4. Par rapport au schéma dans la Figure 3, il n'y a pas de zone amont pour les trains en provenance de Poissy, et le train est en zone de décision jusqu'à la détection de son départ de la gare d'origine.

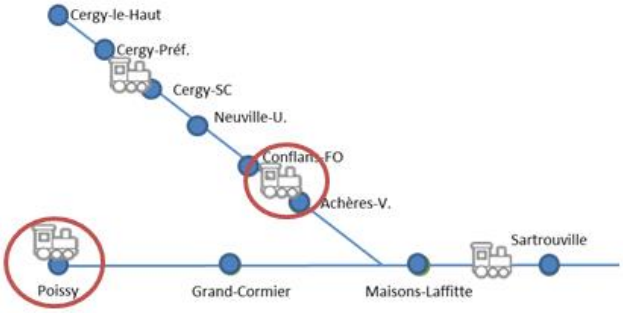


Figure 4 – Schéma de la zone de convergence étudiée

Les expérimentations se sont déroulées pendant trois semaines sur la pointe du matin, de 7:30 à 9:30, à la présence constante d'un membre de l'équipe projet. Dans une première phase, nous avons suivi le travail des régulateurs sans imposer la mise en place des préconisations de l'outil. Cela a permis d'évaluer la cohérence et l'apport de l'outil par rapport à la situation existante. Ensuite, une application systématique des préconisations a été demandée : cela a permis d'évaluer la qualité des préconisations de l'outil dans un contexte opérationnel.

Qualité des prévisions par rapport au réalisé

Nous nous sommes d'abord intéressés à l'étude de la qualité des prévisions des positions des trains par rapport au réalisé.

Pour cela, nous avons calculé les écarts entre les positions simulées et réalisées des trains sur des différentes plages horaires de 0 à 30 minutes à partir du moment de préconisation. La distribution de ces écarts, basée sur 1500 simulations, est montrée en Figure 5. Bien évidemment, nous ne pouvons à chaque fois que confronter le scénario correspondant au réordonnancement réalisé.

L'erreur moyenne sur la position de chaque train entre 0 et 5 minutes après la simulation est de 5 secondes. Cette erreur devient plus important le plus on s'éloigne du moment de la simulation, tout en restant acceptable pour nos applications. Par exemple, il est inférieur à 40 secondes en moyenne pour le calcul des positions des trains entre 25 et 30 minutes après la simulation, bien que l'écart type soit plus important. L'augmentation de l'écart-type est due aux aléas ou incidents qui peuvent apparaitre sur la ligne, et dont le simulateur ne peut pas avoir connaissance au moment du déclanchement de la simulation.

Nous pouvons noter que les sous-estimations des retards (partie droite du graphe) sont plus importantes que les surestimations. Cette différence est normale : le simulateur ne prévoit pas la possibilité qu'un train puisse rattraper son retard, car cela signifierait un temps de parcours plus court par rapport au temps de parcours théorique. Il y a tout de même des cas où le simulateur peut surestimer un retard : notamment si ce retard est dû à l'espacement avec un train qui sera au final supprimé.

La qualité de la simulation est très bonne par rapport à la précision requise pour nos usages, et nous conforte sur la qualité des préconisations données par l'outil.

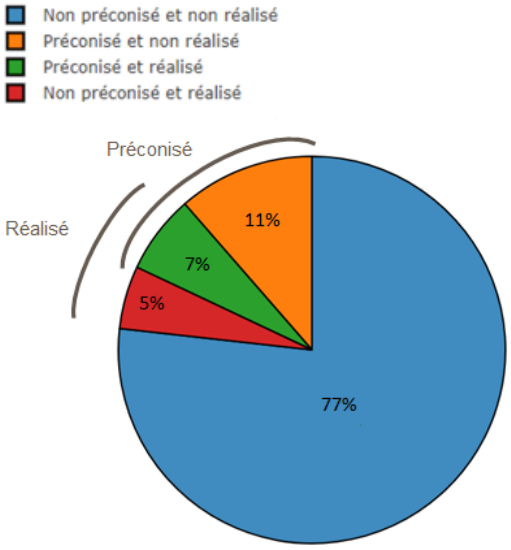


Figure 6 – Pourcentage de réordonnancements préconisés vs réalisés

Réordonnancements : effectués et préconisés

Nous avons par la suite analysé les réordonnancements effectués par le régulateur par rapport aux réordonnancements préconisés par l'outil. Les résultats sont montrés en Figure 6.

Nous retrouvons un accord entre la préconisation de l'outil et la décision de l'opérateur dans le 84% des cas (partie bleu et partie verte du graph). Cet accord montre que les préconisations de l'outil sont globalement cohérentes, bien qu'une marge

©AfIA 2018

d'optimisation importante reste possible : 11% des réordonnancements proposés ne sont pas effectués, et 5% des réordonnancements effectués n'étaient pas préconisés par l'outil. La cause principale de cette discordance est à rechercher dans la complexité du critère de ponctualité voyageurs : en effet, la plupart de réordonnancements non effectués (partie jaune) ne réduisait pas les retards des trains. Ce dernier critère est plus facile à estimer par l'opérateur par rapport au critère de ponctualité voyageurs. D'autres raisons de désaccord entre l'outil et la réalité incluent l'utilisation des voies supplémentaires afin de décharger la zone de convergences, ou la nécessité de faire passer un train en premier pour des sujets de réutilisation des rames ou de réaffectation des conducteurs. Certains de ces aspects pourront être intégrés par des versions futures de l'outil, bien que la disponibilité en temps réel des données nécessaires à l'implémentation de ces fonctionnalités reste pour l'instant un point bloquant.

Analyse de la phase d'application systématique des préconisations de l'outil

L'objectif est de comprendre si l'application des préconisations a amélioré la ponctualité voyageurs, comme prédit par l'outil. Pour cela, nous avons analysé dans le détail les situations où l'opérateur a suivi la préconisation de l'outil, mais où il aurait pris la décision opposée en dehors du cadre des expérimentations, sans l'outil à disposition. Un exemple est analysé en Figure 7, qui montre les Graphiques Espace-Temps (GET) pour les deux scénarios d'ordonnancement des trains. Les GET affichent en abscisse les horaires et en ordonnée les positions sur la ligne. Les deux trains A et B viennent de deux branches différentes, mais nous affichons sur le graphique la position par rapport à la convergence.

La décision a des impacts sur les trains suivants, mais nous ne montrons par simplicité que les deux trains réordonnancés.

Le scénario réalisé est celui du réordonnancement : B passe devant A. La simulation prévoit les positions des trains avec une bonne précision par rapport au réalisé. À noter qu'à chaque fois nous manquons d'informations sur la qualité de la simulation par rapport au scénario non retenu.

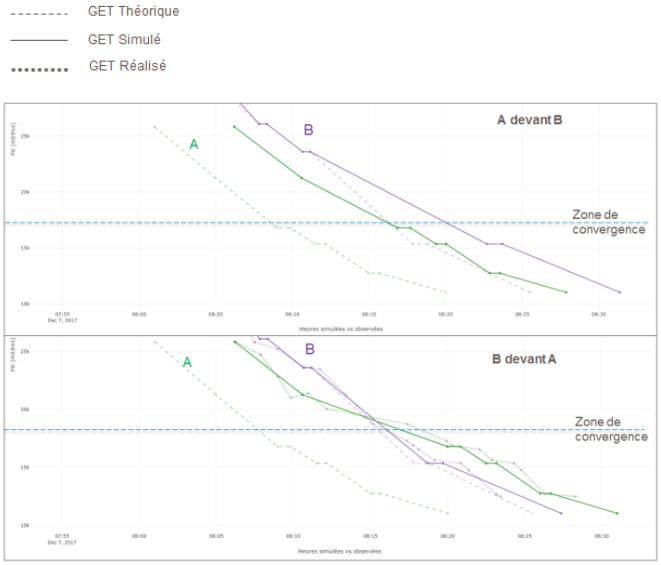


Figure 7 – GET pour le scénario A passe devant B et B passe devant A (scénario retenu)

Une analyse de chaque groupe de voyageurs impacté par la décision a confirmé l'amélioration de la ponctualité voyageurs par rapport au scénario non préconisé.

Estimation du gain en ponctualité voyageurs

Une estimation solide du gain en ponctualité voyageurs demanderait l'application systématique des préconisations de l'outil sur une période de plusieurs semaines, afin d'analyser si la ponctualité est

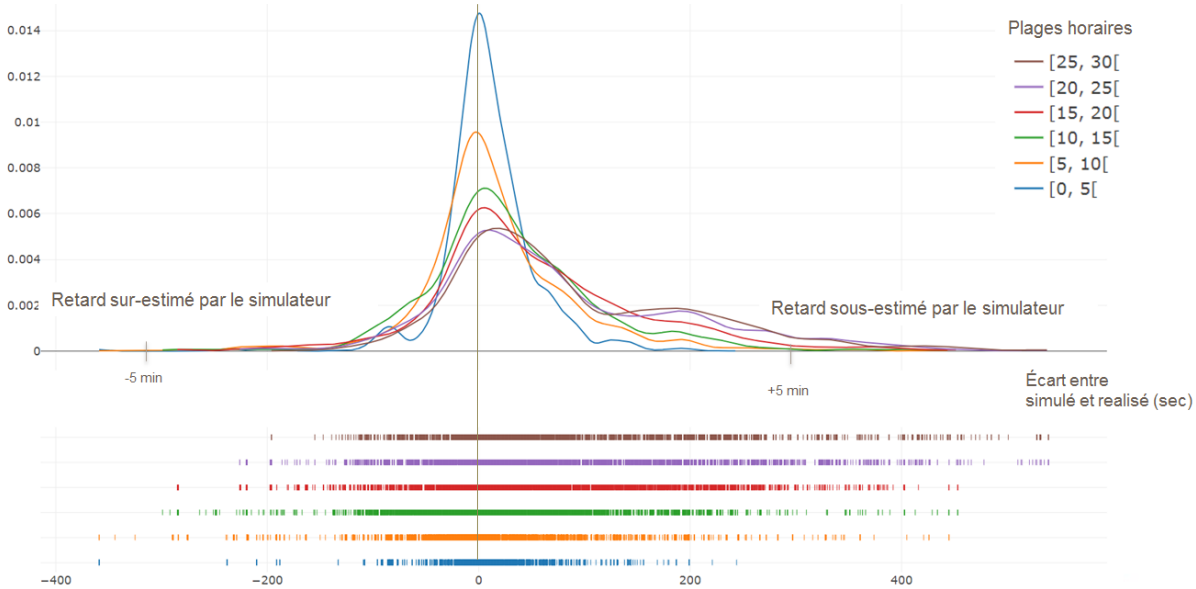


Figure 5 - Distribution des écarts entre réalisé et simulé sur des plages horaires de 5 minutes entre 0 et 30 minutes après la préconisation

significativement améliorée par rapport à une période de référence. Cela n'était pas possible dans le cadre de cette expérimentation. Nous nous sommes contentés d'effectuer une estimation moins fine en faisant une simple somme des gains en ponctualité générés par un suivi régulier de l'outil : grâce à une application systématique des préconisations, il serait possible d'éviter la perte en ponctualité liée aux réordonnancements non préconisés et réalisés (partie rouge du graphe en Figure 6) et de profiter du gain générés par l'application des réordonnancements préconisés mais non réalisés (partie jaune du graphe en Figure 6). L'estimation montre qu'un nombre important aujourd'hui en retard pourrait arriver à l'heure en suivant les ordonnancements préconisés par l'outil.

Retours des utilisateurs

Suite aux expérimentations, les utilisateurs ont complété un questionnaire de retour. Les questions visaient à comprendre le ressenti de leur travail quotidien, en particulier pour ce qui concerne la gestion de la convergence, en terme de difficulté et charge cognitive. Des avis ont été demandés sur les expérimentations et l'outil : confiance dans les préconisations, pertinence des informations présentée, etc. Les opérateurs ont donné un avis positif sur les préconisations de l'outil et les informations montrées au niveau de l'interface. La conduite du changement reste à travailler pour permettre l'utilisation de ce type d'algorithmes dans un contexte opérationnel. Malgré les nombreux échanges avec les utilisateurs pendant la conception du prototype, il reste difficile de changer des pratiques d'exploitations consolidées, notamment dans le cadre d'une expérimentation d'une durée de quelques semaines.

6 Conclusions et perspectives

Dans cet article, nous avons présenté un outil d'aide à la gestion des zones de convergence ferroviaire, et son application à un cas de test.

L'outil s'est montré pertinent sur plusieurs aspects : aider les décisions des opérateurs au quotidien, en leur fournissant des éléments objectifs pour une prise de décision rapide et optimale ; réduire la variabilité liée au facteur humain, en rendant l'exploitation plus homogène d'un jour à l'autre ; améliorer la ponctualité voyageurs, objectif primaire de Transilien.
Suite à ces résultats, le passage en industrialisation du prototype est prévu.

Du point de vue du projet de recherche, des améliorations sont envisageables pour traiter des cas plus complexes : par exemple, pour aider la gestion des lignes avec des convergences multiples, ou pour intégrer d'autres décisions en temps réel comme des adaptations de l'offre.

Bibliographie

[1] V. Cacchiani, D. Huisman, M. Kidd, L. Kroon, P. Toth, L. Veelenturf, et J. Wagenaar (2014). An overview of recovery models and algorithms for real-time railway rescheduling. Transportation Research Part B: Methodological, 63, pages 15-37.

[2] A. Toletti, M. Laumanns, P. Grossenbacher et U. Weidmann (2015). Meeting functional requirements for real-time railway traffic management with mathematical model. Advanced Systems in Public Transport Conference.

[3] F. Corman and L. Meng (2013). A review of online dynamic models and algorithms for railway traffic control. Intelligent Rail Transportation (ICIRT), 2013 IEEE International Conference, pages 128-133.

[4] C. J. Goodman et S. Murata (2001). Metro traffic regulation from the passenger perspective. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, Vol 215, Issue 2, pages 137 – 147.

[5] N. Tomii, Y. Tashiro, N. Tanabe, C. Hirai et K. Muraki (2005). Train rescheduling algorithm which minimizes passengers' dissatisfaction. Ali M., Esposito F., Innovations in Applied Artificial Intelligence. IEA/AIE. Lecture Notes in Computer Science, vol 3533. Springer, Berlin, Heidelberg

[6] L. Chen, F. Schmid, M. Dasigi, B. Ning, C. Roberts, and T. Tang. (2010). Real-time train rescheduling in junction areas. Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 224(6), pages 547-557.

[7] A. Galapitage, A. R.Albrechta, P. Pudneyab, X. Vua et P. Zhoua (2017). Optimal real-time junction scheduling for trains with connected driver advice systems. Journal of Rail Transport Planning & Management.

[8] S. Van Thielen, S. Burggraeve, and P. Vansteenwegen (2015). "Optimal train rescheduling after conflict detection". Conference on Advanced Systems in Public Transport (CASPT)

[9] J. Damay, A. Boillot, (2014). "An efficient macroscopic railway system simulator". 20th Conference of the International Federation of Operational Research Societies (IFORS).

Remerciements

Ces travaux ont été financés par SNCF Transilien et ont été menés par la Direction de l'Innovation de la SNCF. Nous tenons à exprimer nos remerciements à toutes personnes impliquées dans le projet : le comité de pilotage, le Lab Mass Transit, la DSI de SNCF Mobilité, les opérateurs et les responsables du centre opérationnel où nous avons testé l'outil, les équipes de développements du logiciel, les fournisseurs des données de géolocalisation et de plans de transport.

©AfIA 2018

Trois approches pour classifier les données du web des données

Justine Reynaud Yannick Toussaint Amedeo Napoli

Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France

prénom.nom@loria.fr

Résumé

Dans cet article, nous nous intéressons au processus de classification de données relationnelles issues du web des données. Nous disposons d'un ensemble d'objets entre lesquels il existe des relations. Ces objets appartiennent à une ou plusieurs classes. Celles-ci sont définies en extension*, et nous cherchons à construire une description en* intension *en s'appuyant sur les relations des objets qui les composent. Pour cela, nous employons trois approches : les règles d'association qui s'appuient sur l'Analyse de Concepts Formels (FCA), les redescriptions et les règles de traduction qui s'appuient sur la Longueur de Description Minimale (MDL). À partir d'expérimentations sur DBpedia, nous discutons les spécificités et la complémentarité de ces trois approches. Nous montrons que les règles d'association sont les plus exhaustives tandis que les règles de traduction ont une meilleure couverture des données. Les redescriptions pour leur part, sont les règles les plus faciles à appréhender et interpréter.*

Mots Clef

Données relationnelles, Analyse de Concepts Formels, Fouille de redescriptions, DBpedia.

Abstract

In this paper we study a classification process on relational data that can be applied to the web of data. We start with a set of objects and relations between objects, and extensional classes of objects. We then study how to provide a definition to classes, i.e. to build an intensional description of the class, w.r.t. the relations involving class objects. To this end, we propose three different approaches based on Formal Concept Analysis (FCA), redescription mining and Minimum Description Length (MDL). Relying on some experiments on RDF data from DBpedia, where objects correspond to resources, relations to predicates and classes to categories, we compare the capabilities and the complementarity of the three approaches. This research work is a contribution to understanding the connections existing between FCA and other data mining formalisms which are gaining importance in knowledge discovery, namely redescription mining and MDL.

Keywords

Relational data, Formal Concept Analysis, Redescription mining, DBpedia.

1 Introduction

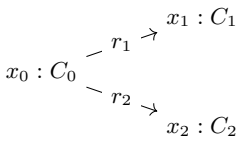
Dans cet article, nous nous intéressons à la possibilité de découvrir des définitions dans les données RDF issues du web des données. Ces définitions peuvent être réutilisées dans la conception de bases de connaissances (KBs) ou dans l'enrichissement de KBs préexistantes. Étant donné l'immense quantité de données du web des données, il s'agit d'un enjeu majeur.

Le problème est le suivant : nous disposons d'un ensemble d'objets connectés par des relations, comme une ABox en logique de descriptions (DL) (Baader *et al.*, 2003). L'objectif est de classifier ces objets en fonction des relations dans lesquelles ils sont impliqués.

Les objets qui partagent des éléments communs appartiennent à une même classe. Ce partage peut être exact — les éléments sont identiques — ou approximatif — les éléments sont similaires. Finalement, nous obtenons un ensemble de classes organisées selon un ordre partiel, et les descriptions associées à ces classes. Ces descriptions sont nécessaires afin de construire les définitions des différentes classes. Les définitions sont considérées comme des conditions nécessaires (NC) et suffisantes (SC) pour classifier de nouveaux objets. Si $\mathtt{x}$ est une instance de la classe `Rouge` alors $\mathtt{x}$ a la couleur `rouge` (NC), et inversement, si $\mathtt{x}$ à la couleur `rouge` alors $\mathtt{x}$ est une instance de la classe `Rouge` (SC).

Pour poursuivre l'analogie avec les DLs, l'idée dans cet article est de construire et d'appliquer des règles d'induction de la forme « $\mathtt{si}\ \mathtt{r}(\mathtt{x},\mathtt{y})\ \mathtt{et}\ \mathtt{y}{:}\mathtt{C}\ \mathtt{alors}\ \mathtt{x}{:}\exists\mathtt{r}.\mathtt{C}$ ». Cela signifie que, étant données $\mathtt{y}$ une instance de la classe $\mathtt{C}$ et $\mathtt{r}$ une relation telle que $\mathtt{r}(\mathtt{x},\mathtt{y})$, alors $\mathtt{x}$ appartient à une classe, disons $\mathtt{D}$, dont la description comprend $\exists\mathtt{r}.\mathtt{C}$. C'est-à-dire que les instances de $\mathtt{D}$ sont reliées à au moins une instance de $\mathtt{C}$ par la relation $\mathtt{r}$. Nous utilisons ce type de règles pour construire les définitions des classes. Ces définitions sont de la forme $\mathtt{C_i} \equiv \mathtt{e_1} \sqcap \mathtt{e_2} \sqcap \ldots \sqcap \mathtt{e_n}$ où $\mathtt{e_j}$ est une expression de la forme $\exists\mathtt{r}.\mathtt{C_j}$. Notre travail se divise en trois tâches principales, à savoir (i) préparer les données, (ii) découvrir les définitions, (iii) évaluer la qualité de ces

 ©AfIA 2018




```
ex : x0 ex : r1 ex : x1.
ex : x0 ex : r2 ex : x2.
ex : x0 rdf : type ex : C0.
ex : x1 rdf : type ex : C1.
ex : x2 rdf : type ex : C2.
```

FIGURE 1 – Exemple de données relationnelles et les triplets associés.

définitions. Nous nous appuyons ici sur trois approches, les règles d'association, la fouille de redescriptions et la découverte de règles de traduction.

Cet article est dans la continuité de travaux de recherche sur la découverte de définitions dans le web des données. Son originalité est de comparer trois approches qui ne s'appuient pas sur les mêmes principes mais qui s'avèrent être complémentaires au vu des résultats de nos expérimentations. À notre connaissance, il s'agit du premier article où une telle comparaison est proposée à la fois à un niveau théorique et à un niveau expérimental.

L'article est organisé comme suit : dans la section 2, nous présentons la nature des données sur lesquelles nous travaillons et les processus de classification dans le web des données. Dans la section 3, nous détaillons les trois approches de classification et leurs applications. La section 5 présente les expériences qui ont été menées ainsi que l'évaluation des règles obtenues. Enfin, la section 6 présente une discussion ainsi que les pistes de recherches futures avant la conclusion section 7.

2 Représentation des données

2.1 Web des Données

Dans cette section, nous présentons les données issues du Web des Données (LOD) que nous considérons. Le LOD peut être vu comme un ensemble de KBs interconnectées. Une KB est composée de deux éléments : une TBox qui définit son *schema* et une ABox qui introduit les individus et leurs relations. L'unité de base d'une KB est le triplet RDF, noté $\langle s, p, o\rangle$, qui encode une assertion sous la forme sujet–prédicat–objet. Les différentes composantes d'un triplet peuvent être des *ressources* U identifiées de manière unique, des *littéraux* L (chaîne de caractères, numérique, date, ...) ou des *nœuds anonymes* B (assimilables à une variable existentiellement quantifiée), de telle sorte que $\langle s, p, o\rangle \in (B \cup U) \times U \times (B \cup U \cup L)$. Dans cet article, nous nous restreignons aux triplets composés uniquement de ressources, c'est-à-dire tels que $\langle s, p, o\rangle \in U \times U \times U$. Les ressources peuvent faire référence à n'importe quel objet ou abstraction et sont identifiées par une URI (*Uniform Resource Identifier*). Une URI est une adresse composée de deux parties. La première partie est l'espace de nom (*namespace*) qui indique de quelle KB viennent les ressources. La seconde partie donne un nom à la ressource au sein de cette KB.

La figure 1, présente un exemple de données relationnelles et l'ensemble de triplets correspondant. Le préfixe `ex:` correspond à un *namespace* créé pour notre exemple, tandis que le préfix `rdf:` correspond à un *namespace* préexistant.`rdf:type` est la relation d'instanciation. Ainsi le triplet $\langle$`ex:x0 rdf:type ex : C0`$\rangle$ indique que $\mathtt{x_0}$ est une instance de la classe $\mathtt{C_0}$.

Le LOD peut être interrogé grâce aux requêtes SPARQL. Par exemple, la requête `SELECT ?x WHERE { ?x rdf:type ex:C0}` retourne toutes les instances de $\mathtt{C_0}$. Si l'on prend les données de la figure 1, seul `ex:x0` est retourné.

2.2 Analyse de Concepts Formels

Nous utilisons ici l'Analyse de Concepts Formels (FCA) de Ganter et Wille (1999) pour présenter et comparer les différentes approches. Étant donné un ensemble G d'entités [1], un ensemble M d'attributs et une relation binaire $I \subseteq G \times M$, (G, M, I) est un contexte formel. L'expression gIm s'interprète comme « l'entité g possède l'attribut m». Les correspondances de Galois (notées $.'$) pour un ensemble d'entités $X \subseteq G$ et un ensemble d'attributs $Y \subseteq M$ sont définies comme suit :

$$X' = \{m \in M \mid \forall x \in X, xIm\} \quad \text{et}$$
$$Y' = \{g \in G \mid \forall y \in Y, gIy\}.$$

À partir des données RDF, nous construisons un contexte formel dont les entités sont les sujets des triplets. Les attributs sont les paires (prédicat,objet) issues des triplets. Nous distinguons deux types d'attributs : des descriptions et des classes. Le premier ensemble d'attributs, dénoté $\mathcal{C}$, est composée des paires de la forme $(\mathtt{rdf:type}, \mathtt{C})$, qui font d'un sujet une instance de la classe `C`. Ce sont ces classes que nous cherchons à définir. Le second ensemble, dénoté $\mathcal{D}$ est composé de paires $(\mathtt{p}, \mathtt{o})$ telles que $\mathtt{p} \neq \mathtt{rdf:type}$. Les attributs du contexte sont donc $M = \mathcal{C} \cup \mathcal{D}$. Si l'on prend l'exemple de la figure 1, la figure **??** présente le contexte associé.

		Attributs				
		Descriptions		Classes		
		$\exists r1{:}C_1$	$\exists r2{:}C_2$	C_0	C_1	C_2
Objets	x_0	×	×	×		
	x_1				×	
	x_2					×

FIGURE 2 – Contexte formel associé aux données représentées figure 1.

À partir de là, il s'agit de trouver un ensemble de catégories et un ensemble de descriptions de manière à ce que leurs correspondances soient les mêmes. Sur la figure **??** par exemple, on a $\{C_0\}' = x_0$ et $\{\exists r_1{:}C_1, \exists r_2{:}C_2\}' = x_0$. On peut alors construire la définition suivante :

$$C_0 \equiv \exists r_1{:}C_1 \sqcap \exists r_2{:}C_2$$

1. En FCA, G est l'ensemble des *objets*, renommés *entités* afin de ne pas confondre avec les objets en RDF.

Comme les données peuvent être incomplètes, il se peut qu'on ne retrouve pas une égalité entre une classe $\mathtt{C_i}$ et sa description $\exists \mathtt{r}{:}\mathtt{C_j}$. Autrement dit $\{\mathtt{C_i}\}' \neq \{\exists \mathtt{r}{:}\mathtt{C_j}\}'$. Il nous faut donc des approches qui tolèrent une forme d'approximation. C'est ce que permettent les trois algorithmes utilisés, et nous allons les détailler plus précisément dans la section suivante.

3 Algorithmes de fouille de règles

3.1 Règles d'association

Le but de la fouille de règles d'association (Agrawal *et al.*, 1993; Klemettinen *et al.*, 1994) est de trouver des dépendances entre les attributs. Une règle d'association entre deux ensembles d'attributs X et Y, notée $X \to Y$, signifie « si un objet a tous les attributs de X, alors il a tous les attributs de Y ». À cette règle est associée une *confiance*, qui peut être vue comme une probabilité conditionnelle :

$$\mathrm{conf}(X \to Y) = \frac{|X' \cap Y'|}{|X'|},$$

où $.'$ correspond à la correspondance de Galois. La confiance est une mesure de qualité des règles d'association. Une règle d'association est valide si sa confiance est supérieure à un seuil θ défini par l'utilisateur. La confiance n'est pas symétrique : $X \to Y$ peut être valide sans que $Y \to X$ le soit. Si la confiance vaut 1, on dit qu'il s'agit d'une *implication* et l'on note $X \Rightarrow Y$. Si on a également $Y \Rightarrow X$, alors X et Y forment une définition, notée $X \equiv Y$.

Nous nous intéressons ici à des définitions. Nous considérons donc conjointement $X \to Y$ et sa réciproque $Y \to X$ et cherchons à estimer à quel point ces règles s'approchent d'une implication. Pour cela, nous introduisons la notion de *quasi-définition* qui est à la définition ce que la règle d'association est à l'implication.

Définition 1 (Quasi-définition)**.** Étant donné deux ensembles d'attributs X et Y, une quasi-définition $X \leftrightarrow Y$ est valide si, pour un seuil θ donné,

$$\min(\mathrm{conf}(X \to Y), \mathrm{conf}(Y \to X)) \geqslant \theta.$$

L'algorithme `Eclat` (Zaki, 2000) est l'un des nombreux algorithmes de fouille de règles d'association existant. Il énumère de manière exhaustive toutes les règles d'association valides pour un seuil donné. Nous utilisons cet algorithme pour notre comparaison.

Une règle d'association $r_0 : x_1, \ldots, x_n \to y_1, \ldots, y_m$ peut se décomposer sous la forme de plusieurs règles d'association $r_i : x_1, \ldots, x_n \to y_i$ pour $i \in \{1, \ldots, m\}$. Si r_0 est valide, l'ensemble des r_i est valide. Pour que les règles obtenues définissent des classes comme nous le souhaitons, il faut que les règles d'association soient de la forme «Classes $\to$ Descriptions» ou «Descriptions $\to$ Classes». C'est à dire que la règle $X \to Y$ est telle que $X \subseteq \mathcal{C}, Y \subseteq \mathcal{D}$ ou $X \subseteq \mathcal{D}, Y \subseteq \mathcal{C}$. Il nous faut donc utiliser un post-traitement afin de s'assurer que la règle ne contient que des catégories d'un côté et que des descriptions de l'autre. Pour cela, seules les règles $X \to Y$ qui satisfont l'une des deux contraintes alternatives suivantes sont conservées : (i) l'antécédent ne contient que des classes ($X \subseteq \mathcal{C}$) et il y a au moins une description dans la conséquence ($Y \cap \mathcal{D} \neq \emptyset$) ; (ii) l'antécédent ne contient que des descriptions ($X \subseteq \mathcal{D}$) et il y a au moins une classe dans la conséquence ($Y \cap \mathcal{C} \neq \emptyset$). Les règles conservées sont décomposées de manière à ne garder que des classes (resp. des descriptions) dans la conséquence si l'antécédent ne contient que des descriptions (resp. des classes).

Par exemple, $\exists r_1{:}C_1, C_0 \to \exists r2{:}C_2$ n'est pas conservée parce que l'antécédent contient à la fois une catégorie et une description. En revanche, la règle $\exists r_1{:}C_1 \to \exists r_2{:}C_2, C_0$ peut être décomposée en deux règles r_1: $\exists r_1{:}C_1 \to \exists r_2{:}C_2$ et r_2: $\exists r_1{:}C_1 \to C_0$. La règle r_2 est conservée. Si sa réciproque est valide, la quasi-définition obtenue est $C_0 \leftrightarrow \exists r_1{:}C_1$.

3.2 Redescriptions

La fouille de redescriptions, introduite par Ramakrishnan *et al.* (2004), a pour but de trouver des caractérisations multiples d'un même ensemble d'entités. À la différence des règles d'association, les redescriptions s'appuient sur une séparation de l'ensemble des attributs en *vues*. Une vue est un sous-ensemble d'attributs. L'ensemble des vues forme une partition des attributs. Nous utilisons ici deux vues, qui correspondent aux deux types d'attributs — classes et descriptions.

La similarité entre deux ensembles d'attributs, provenant de deux vues différentes, est mesurée avec le coefficient de Jaccard :

$$\mathrm{jacc}(A, B) = \frac{|\ A' \cap B'\ |}{|\ A' \cup B'\ |},$$

où $.'$ correspond à la correspondance de Galois. Une paire d'ensembles d'attributs (A, B) est une redescription si le coefficient de Jaccard $\mathrm{jacc}(A, B)$ est supérieur à un seuil donné. Le coefficient de Jaccard est symétrique, contrairement à la confiance. Une redescription dont le coefficient de Jaccard vaut 1 est une définition. Toutes les redescriptions sont nécessairement des quasi-définitions. En effet,

$$\min(\mathrm{conf}(A \to B), \mathrm{conf}(B \to A)) \geqslant \mathrm{jacc}(A, B).$$

L'algorithme `ReReMi` (Galbrun et Miettinen, 2012) est utilisé ici. En plus des données binaires, `ReReMi` permet de tenir compte de données numériques et catégorielles. Les redescriptions obtenues peuvent être des fonctions booléennes contenant des disjonctions et des négations d'attributs. Afin de comparer les trois algorithmes, `ReReMi` est ici restreint à des conjonctions d'attributs, raison pour laquelle nous ne parlons pas de fonctions booléennes.

Dans un premier temps, `ReReMi` cherche des paires d'attributs — un attribut de chaque vue — susceptibles de prendre part à une définition, c'est-à-dire celles qui ont le coefficient de Jaccard le plus élevé. Ces paires sont ensuite

©AfIA 2018

étendues tour à tour, en ajoutant à chaque fois un attribut à l'un des côtés de la redescription, jusqu'à ce que le nombre d'attributs maximum soit atteint ou que le coefficient de Jaccard n'augmente plus. Les redescriptions ainsi obtenues qui satisfont les critères de sélection sont retournées à l'utilisateur.

3.3 Règles de traduction

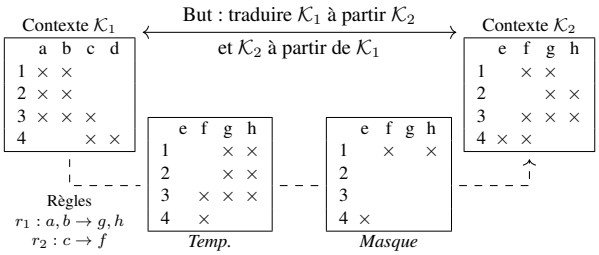


FIGURE 3 – Intuition du fonctionnement de `Translator`. Ici, seule la traduction de $\mathcal{K}_1$ vers $\mathcal{K}_2$ est représentée. À chaque étape, le but est de trouver la règle qui va minimiser le nombre de croix dans le masque.

L'algorithme `Translator` (van Leeuwen et Galbrun, 2015) s'appuie également sur une séparation du contexte en deux vues, et cherche un ensemble d'associations entre ces deux vues. Ces associations ont la même forme que des règles d'association, la différence se situant au niveau de la constitution de l'ensemble de règles.
Cet ensemble doit être compact et représentatif. D'une part, il doit couvrir la majorité des données. D'autre part, les règles doivent être aussi petites que possible en terme d'attributs. Afin de trouver un équilibre entre ces deux contraintes, `Translator` s'appuie sur le concept de Longueur de Description Minimum (MDL). Étant donné un contexte $\mathcal{K}$ et $X \subseteq M$ un ensemble d'attributs, la *longueur* de X est

$$L(X) = -\sum_{x \in X} \log_2 P(x \mid \mathcal{K}) \quad \text{où } P(x \mid \mathcal{K}) = \frac{\mid x' \mid}{\mid G \mid}.$$

Cette longueur correspond au nombre minimal de bits requis pour encoder X.
Pour construire l'ensemble des règles, les auteurs font l'analogie avec la notion de traduction. Une règle est une traduction d'une vue vers une autre. L'idée sous-jacente est la suivante : on souhaite construire un ensemble de règles qui permettent, connaissant une des deux vues, de reconstruire la seconde et *vice versa*. L'idée générale est représentée figure 2. Les erreurs entre le contexte cible et le contexte reconstruit sont corrigées à l'aide d'un masque. La taille de ce masque indique donc le nombre d'erreurs introduites.
L'ensemble de règles est construit de manière itérative, en prenant, à chaque étape, la règle $X \rightarrow Y$ qui maximise Δ

$$\Delta(X \rightarrow Y) = \underbrace{L(\mathit{Mask}^-) - L(\mathit{Mask}^+)}_{\text{Gain d'information}} - \underbrace{L(X \rightarrow Y)}_{\text{Longueur de la règle}}$$

où Mask^+ correspond aux éléments ajoutés au masque (erreurs introduites par la règle) et Mask^- correspond aux éléments retirés du masque (erreurs corrigées par la règle). Des règles sont ajoutées tant que Δ est positif.
L'algorithme `Translator` est le seul dont la mesure de qualité tient compte des règles déjà extraites. Le masque étant mis à jour à chaque étape, la qualité d'une règle dépend des règles déjà extraites. Ainsi, `Translator` favorise un *bon ensemble de règles* plutôt qu'un *ensemble de bonnes règles*. Chaque règle apporte une information qui n'est pas contenue dans les autres, ce qui limite la redondance d'information.

4 Travaux connexes

La FCA est un outil de conceptualisation qui permet de structurer une ontologie par une approche *bottom-up*. Sertkaya (2011) fait une synthèse des différentes contributions qui s'intéressent au lien entre la FCA et les ontologies. La plupart des ontologies sont construites avec une approche *top-down*, c'est à dire en construisant d'abord le schéma. L'une des façons de tirer parti de la FCA est donc de consuire des connaissances qui vont enrichir une ontologie préexistante. C'est ce que nous faisons dans notre approche, en partant de triplets pour trouver des définitions.
Dans (Ferré et Cellier, 2016), les auteurs proposent une extension de la FCA pour les graphes conceptuels. Contrairement aux graphes RDF qui reposent sur des relations binaires, les graphes conceptuels permettent de tenir compte de relation n-aires. Les concepts construits correspondent à une association entre une requête SPARQL et l'ensemble des solutions. Dans notre approche, nous cherchons à mettre en relation deux *graph patterns* et nous ne connaissons pas *a priori* la requête SPARQL qui correspond à la description d'une catégorie.
Notre travail poursuit un travail initié dans (Alam *et al.*, 2015). Les auteurs s'appuient sur la fouille de règles d'association pour fournir un espace de navigation sur des données RDF. Pour cela, les auteurs recherchent des implications et les ordonnent en fonction de la confiance de leur réciproque. Dans notre approche, nous introduisons la séparation des attributs et nous comparons les règles d'association avec les redescriptions et les règles de traduction.
L'algorithme AMIE, proposé par Galárraga *et al.* (2013), est une référence en matière de fouille de règles sur le web des données. Ces règles sont de la forme $B_1, B_2, \ldots B_{n-1} \rightarrow B_n$ où B_i correspond à une relation entre deux variables, que l'on peut exprimer sous forme de triplet $<?x \ \mathtt{r} \ ?y>$. Par exemple, $manufacturer(x, Samsung)$, $successor(y, x) \rightarrow manufacturer(y, Samsung)$ est une règle qui peut être trouvée par AMIE. Les auteurs ajoutent une condition, à savoir que toutes les variables doivent apparaître deux fois dans la règle. Si l'on considère les triplets sous forme de graphe, il en résulte qu'une règle est un motif cyclique entre des variables (indépendamment de l'orientation des arcs). Dans notre cas, les motifs recherchés

caractérisent une seule ressource identifiée (motif en étoile) et l'association est bidirectionnelle. On trouve par exemple la règle $Samsung_Mobile_Phone(x) \leftrightarrow manufacturer(x, Samsung)$.

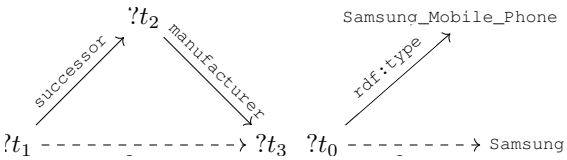


FIGURE 4 – Motifs recherchés dans la base de connaissances par AMIE et par notre approche. Les traits pleins concernent la partie gauche de la règle, et les pointillés la partie droite. Les $?t_i$ sont des variables existentielles.

5 Expérimentations

Nous avons réalisé nos expérimentations sur des données issues de *DBpedia*, qui est une KB construite à partir de *Wikipédia*. Nous cherchons à définir des *catégories*, c'est-à-dire les ressources qui sont dans le codomaine de la relation `dct:subject`. Pour cela, nous extrayons un sous-ensemble de *DBpedia* à l'aide d'une requête SPARQL avant de transformer les triplets obtenus en contexte comme mentionné en Section 2.2. Enfin, nous utilisons les trois algorithmes présentés pour comparer et évaluer les résultats obtenus.

TABLE 1 – Statistiques sur les jeux de données extraits.

	Triplets	Objets	Attributs	
			Cat.	Descr.
`Turing_Award`	2 642	65	503	857
`Smartphones`	8 418	598	359	1 730
`Sports_cars`	9 047	604	435	2 295
`French_films`	121 496	6 039	6 028	19 459

5.1 Méthodologie

Nous avons extrait quatre sous-ensembles de triplets de *DBpedia*, de différents domaines [2]. Afin d'isoler un sous-ensemble de triplets, nous avons récupéré tous les triplets dont le sujet est lié à une catégorie donnée. Pour cela, nous avons utilisé la requête SPARQL suivante :

```
SELECT DISTINCT * WHERE {
 ?s ?p ?o .
 ?s dct:subject dbc:X .
 ?p a owl:ObjectProperty .
}
```

Les quatres domaines utilisés (X) sont `French_films`, `Turing_Award_laureates`, `Smartphones` et `Sports_cars`. Le triplet `?p a owl:objectProperty` assure que $?o$ n'est pas un littéral.

Les triplets extraits sont divisés en deux groupes. L'ensemble des triplets qui ont pour prédicat `dct:subject` correspond aux attributs de classe ($\mathcal{C}$). L'ensemble des triplets dont le prédicat n'est pas `dct:subject` correspond aux attributs de description ($\mathcal{D}$).

5.2 Résultats et évaluation

TABLE 2 – Évaluation des résultats pour chaque jeu de données. $\overline{|C_i|}$ (resp. $\overline{|D_i|}$) désigne le nombre moyen de catégories (resp. de descriptions) dans une règle.

`Turing_Award_laureates`

X	`Eclat`	`ReReMi`	`Translator`
$\lvert\mathcal{B}^X_{cand}\rvert$	47	12	11
$\lvert\mathcal{B}^X_{def}\rvert$	30	9	9
Précision	.64	.75	.85
$\overline{\lvert C_i\rvert}$ — $\overline{\lvert D_i\rvert}$	2 — 4	1 — 1	3 — 5

`Sports_cars`

X	`Eclat`	`ReReMi`	`Translator`
$\lvert\mathcal{B}^X_{cand}\rvert$	132	52	31
$\lvert\mathcal{B}^X_{def}\rvert$	95	30	23
Précision	.72	.68	.74
$\overline{\lvert C_i\rvert}$ — $\overline{\lvert D_i\rvert}$	2.8 — 4.5	1.3 — 1.4	2.6 — 4.1

`Smartphones`

X	`Eclat`	`ReReMi`	`Translator`
$\lvert\mathcal{B}^X_{cand}\rvert$	810	98	41
$\lvert\mathcal{B}^X_{def}\rvert$	521	57	31
Précision	.64	.58	.76
$\overline{\lvert C_i\rvert}$ — $\overline{\lvert D_i\rvert}$	4.3 — 7.8	1.6 — 1.8	3.1 — 3.1

`French_films`

X	`Eclat`	`ReReMi`	`Translator`
$\lvert\mathcal{B}^X_{cand}\rvert$	546	36	93
$\lvert\mathcal{B}^X_{def}\rvert$	371	12	89
Précision	.68	.33	.96
$\overline{\lvert C_i\rvert}$ — $\overline{\lvert D_i\rvert}$	2.8 — 4.4	1.2 — 1.1	2.3 — 4.2

Chaque algorithme retourne un ensemble ordonné de quasi-définitions de la forme $C_0, \ldots, C_n \leftrightarrow D_0, \ldots, D_m$. Chacune des quasi-définitions est évaluée manuellement par trois personnes jouant le rôle d'experts. Étant donnée une quasi-définition $C_0, \ldots, C_n \leftrightarrow D_0, \ldots, D_m$ d'un jeu de données J, l'évaluateur répond à la question « Étant donné que nous parlons de *J*, est il correct de dire que *faire partie à la fois de* C_0*, de ... et de* C_n et *avoir les propriétés* $D_0, \ldots$*et* D_m est équivalent ? »

L'évaluation finale est l'évaluation majoritaire des trois experts. Si la règle est acceptée, elle rejoint l'ensemble des

2. Les jeux de données ainsi que les résultats obtenus sont disponibles à l'adresse `https://gitlab.inria.fr/jreynaud/DefinitionMiningComparison`.

©AfIA 2018

définitions obtenues, dont 20 exemples sont présentés figure 4. Les experts ont été du même avis dans 95.4% des cas. La règle R5 de la figure 4 est une règle qui n'a pas fait l'unanimité.

Les algorithmes sont comparés sur la base des définitions extraites et des catégories définies. La figure 5 présente un diagramme de Venn pour chaque corpus, qui représente le nombre de définitions extraites par chaque algorithme. Par exemple, pour le corpus `Turing_Award_laureates`, il y a 22 définitions trouvées uniquement par `Eclat`, et 8 trouvées à la fois par `Eclat` et `Translator`. Au total, `Eclat` a extrait 30 définitions. La figure 6 présente également un diagramme de Venn pour chaque corpus, qui correspond au nombre de catégories définies par chaque algorithme. Une catégorie est considérée comme définie dès lors qu'elle apparaît dans au moins une définition. Il peut donc y avoir une ou plusieurs catégories considérées comme définies pour une seule définition. C'est notament le cas des règles R2, R8 – R10, R12, R14, R17 – R20 de la figure 4.

6 Discussion

Ci-dessous, $\mathcal{B}_{cand}[X]$ désigne l'ensemble de toutes les quasi-définitions extraites par l'algorithme X et $\mathcal{B}_{def}[X]$ l'ensemble des définitions de X, c'est-à-dire, l'ensemble des quasi-définitions de $\mathcal{B}_{cand}[X]$ évaluées vraies par les experts. L'ensemble $\mathcal{B}_{cand}$ désigne la totalité des quasi-définitions,indépendamment de l'algorithme qui les a extraites. De la même façon, $\mathcal{B}_{def}$ désigne l'ensemble de toutes les définitions extraites.

6.1 Précision et rappel

La précision d'un algorithme X est $\frac{|\mathcal{B}^X_{def}|}{|\mathcal{B}^X_{cand}|}$. La précision de `ReReMi` a une très forte variabilité (entre 33 et 75%) et est globalement la plus faible, tout particulièrement sur le jeu de données `French_films`. La précision d'`Eclat` est stable (entre 64 et 72%). La meilleure précision est obtenue par `Translator` : quel que soit le jeu de données, elle toujours supérieure à 74%.

Le rappel, défini comme $\frac{|\mathcal{B}^X_{def}|}{|\mathcal{B}_{def}|}$ ne peut pas être utilisé comme une mesure de performance. En effet, certaines règles se recouvrent (ont des attributs en commun des deux côtés). C'est le cas des règles R6 à R10 (figure 4) qui définissent toutes la catégorie `McLaren_vehicles`. Tandis que `Translator` n'extrait qu'une seule règle (R8), `ReReMi` extrait 2 règles (R6 et R7) et `Eclat` en extrait 9 (seules R8 à R10 sont reportées ici).

La table **??** indique le nombre de catégories pour chaque jeu de données. Si l'on se fie à ces valeurs, le rappel est très faible : dans le corpus `Turing_Award_laureates` par exemple, pour 503 catégories dans les données de départ, seules 19 font partie d'une règle. Cela s'explique principalement par la très faible densité des contextes ; une grande quantité de catégories ne concerne qu'une seule entité du jeu de données. Si l'on ne compte que les catégories qui ont un support d'au moins 3, elles ne sont plus que 105, et 47 catégories seulement on un support d'au moins 5. Nous considérons donc le rappel non pas par rapport au nombre de catégories présentes dans le jeu de données, mais par rapport aux catégories retrouvées par l'ensemble des algorithmes, comme présenté figure 6.

6.2 Forme et interprétation des règles

D'après la figure 6, 70% des catégories définies par `Eclat` ou `Translator` sont définies par les deux algorithmes. `Translator` extrait considérablement moins de règles que `Eclat` (jusqu'à 16 fois moins pour le corpus `Smartphones`). Cela s'explique par la façon dont sont générées les règles d'association. Si dans les données, la règle $A \to B$ a le même support que la règle $A \to B, C$, alors seule la règle $A \to B, C$ est conservée. En revanche si le support de $A \to B$ est strictement supérieur au support de $A \to B, C$, `Eclat` génère deux règles. En conséquence, on obtient avec `Eclat` des règles qui ne diffèrent que d'un attribut, comme R9 et R10 par exemple, alors que `Translator` ne génère qu'une seule règle, R8 dans l'exemple.

`ReReMi` n'extrait aucune définition en commun avec `Eclat` et `Translator`. Cette différence s'explique par l'heuristique employée par `ReReMi`. Si C est une catégorie, et que D_1 et D_2 sont deux descriptions telles que $C' = D'_1 = D'_2$, alors `ReReMi` privilégie deux définitions $C \equiv D_1$ et $C \equiv D_2$ tandis que `Eclat` ne génère qu'une seule définition $C \equiv D_1 \sqcap D_2$. C'est par exemple le cas des définitions R6, R7 générées par `ReReMi` et R8, générée par `Eclat` (figure 4). Si $C' = D'_1$ et $D'_1 \subset D'_2$, `ReReMi` génère la définition $C \equiv D_1$, tandis que `Eclat` génère la définition $C \equiv D_1 \sqcap D_2$. Ce cas a également été retrouvé dans nos résultats, comme le montrent les définitions R12 et R13.

La taille des définitions générées par `ReReMi` est inférieure à celle des définitions de `Translator` et `Eclat`. Cela s'explique par l'heuristique de `ReReMi` déaillée dans le paragraphe précédent. `ReReMi` a en moyenne entre 1 et 2 attributs de chaque côté de la définition tandis que pour `Eclat` et `Translator`, il y a en moyenne 3 catégories et 4 descriptions par définition.

Les conjonctions dans les définitions extraites par `ReReMi` n'ont pas le même sens que celles extraites par `Eclat` et `Translator`. Par exemple, si l'on considère la définition R15, l'attribut `(a Device)` peut être enlevé sans altérer la définition : parce que toutes les entités considérées sont des instances de `Device`. À l'inverse, dans la définition R14, aucun attribut ne peut être retiré sans que la définition ne devienne fausse : tous les attributs font partie de la condition nécessaire. Dans notre approche, il nous semble plus pertinent de n'intégrer que des attributs qui contribuent pleinement à la définition dont ils font partie. Ainsi, la définition R14 nous parait préférable à la définition R15, parce qu'elle est apparue plus facile à interpréter.

Turing_Award_laureates

R	**Harvard_University_alumni** ↔ (almaMater Harvard_University)	R1
ET	**Harvard_University_alumni, Turing_Award_laureates** ↔ (a Agent), (a Person), (a Scientist), (almaMater Harvard_University)	R2
E	**Turing_Award_laureates** ↔ (a Agent), (a Person), (award Turing_Award)	R3
ET	**Turing_Award_laureates** ↔ (a Agent), (a Person), (a Scientist), (award Turing_Award)	R4
E	**Modern_cryptographers** ↔ (field Cryptography)	R5

Sports_cars

R	**McLaren_vehicles** ↔ (manufacturer McLaren_Automotive)	R6
R	**McLaren_vehicles** ↔ (assembly Surrey)	R7
ET	**McLaren_vehicles, Sports_cars** ↔ (a Automobile), (a MeanOfTransportation), (assembly Woking), (assembly Surrey), (assembly England), (bodyStyle Coupé), (manufacturer McLaren_Automotive)	R8
E	**McLaren_vehicles, Sports_cars** ↔ (a Automobile), (a MeanOfTransportation), (assembly England), (assembly Surrey), (bodyStyle Coupé)	R9
E	**McLaren_vehicles, Sports_cars** ↔ (a Automobile), (a MeanOfTransportation), (assembly Surrey), (bodyStyle Coupé)	R10

Smartphones

ET	**Firefox_OS_devices, Open-source_mobile_phones, Smartphones, Touchscreen_mobile_phones** ↔ (a Device), (operatingSystem Firefox_OS)	R11
R	**Nokia_mobile_phones** ↔ (manufacturer Nokia)	R12
ET	**Nokia_mobile_phones, Smartphones** ↔ (a Device), (manufacturer Nokia)	R13
R	**Samsung_Galaxy** ↔ (manufacturer Samsung_Electronics), (operatingSystem Android_(operating_system))	R14
ET	**Samsung_Galaxy, Samsung_mobile_phones, Smartphones** ↔ (a Device), (manufacturer Samsung_Electronics), (operatingSystem Android_(operating_system))	R15

French_films

R	**Pathé_films** ↔ (distributor Pathé)	R16
R	**Films_directed_by_Georges_Méliès** ↔ (director Georges_Méliès)	R17
ET	**Films_directed_by_Georges_Méliès, French_films, French_silent_short_films** ↔ (a Film), (a Wikidata :Q11424), (a Work), (director Georges_Méliès)	R18
ET	**Films_directed_by_Jean_Rollin, French_films** ↔ (a Film), (a Wikidata :Q11424), (a Work), (director Jean_Rollin)	R19
ET	**Film_scores_by_Gabriel_Yared, French_films** ↔ (a Film), (a Wikidata :Q11424), (a Work), (musicComposer Gabriel_Yared)	R20

FIGURE 5 – Exemples de quasi-définitions obtenues par `Eclat`, `ReReMi` et `Translator` pour chaque corpus. Lorsque le préfixe d'un attribut de droite n'est pas spécifié, il s'agit de `dbo`. Le préfixe des catégories (`dbc`) n'est pas spécifié.

7 Conclusion

Dans cet article, nous nous sommes intéressés à trois algorithmes pour trouver des définitions de classes dans le web des données. Nous avons vu que chaque algorithme avait ses spécificités et nous avons vérifié empiriquement que les résultats extraits reflètent ces spécificités. Nous avons aussi montré que malgré des approches très différentes, les algorithmes `Eclat` et `Translator` extraient de nombreuses règles communes. À l'inverse, `ReReMi`, malgré une mesure de qualité semblable à `Eclat`, extrait des règles plus courtes. L'intérêt de ces algorithmes dépend de l'objectif de l'utilisateur. Dans le cas de notre expérimentation, `Eclat` est l'algorithme qui a qualifié le plus de catégories, au prix d'un nombre de règles extraites très important. `Translator` extrait significativement moins de règles avec un nombre de catégories définies légèrement inférieur. Enfin, `ReReMi`, malgré un nombre de catégories définies plus faible, offre des définitions plus simples en n'incluant pas les attributs qui ne participent pas pleinement à la règle. Par la suite, plusieurs directions de recherche sont envisagées, au niveau de la préparation des données, de la fouille et de l'évaluation.

Comme mentionné précédemment, les contextes construits à partir des données RDF sont très creux. Dans le cas d'`Eclat` et de `ReReMi`, un seuil de support étant utilisé, l'existence d'attributs à très faible support n'a pas d'impact sur les règles retrouvées. Pour `Translator` en revanche, les règles doivent permettre de reconstruire intégralement le contexte. Dans ce cas, il peut être pertinent de simplifier le contexte avant de procéder à la fouille de règles. Cela peut notamment passer par la suppression des attributs à très faible support, ou au contraire, des attributs ayant un support quasi maximal, et donc qui concernent toutes les entités. Nous pourrons également travailler sur le type des données en entrée, en intégrant notamment des valeurs numériques, ce qui nous permettra de prendre en compte des triplets laissés de côté pour cette expérimentation, tels que les informations de date, d'âge, de taille, ...

Pour améliorer le processus de fouille, il est possible d'ajouter des contraintes sur les règles obtenues afin de restreindre l'espace de recherche, mais aussi d'obtenir des règles plus faciles à interpréter. Dans notre cas, n'autoriser qu'une catégorie par règle serait une contrainte pertinente. Nous pouvons également étendre l'expressivité des règles recherchées en permettant par exemple des opérateurs logiques tels que les disjonctions ou les négations. Si `ReReMi` permet déjà d'obtenir de telles formules lo-

©AfIA 2018

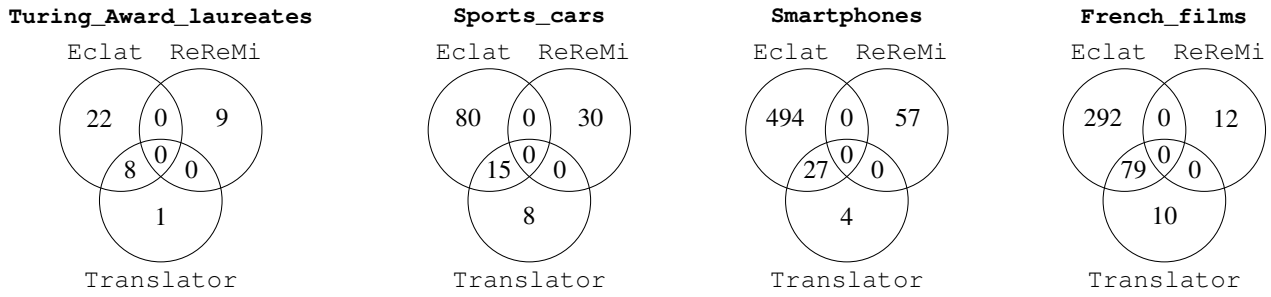


FIGURE 6 – Règles extraites par `Eclat`, `ReReMi` et `Translator` pour chaque jeu de données.

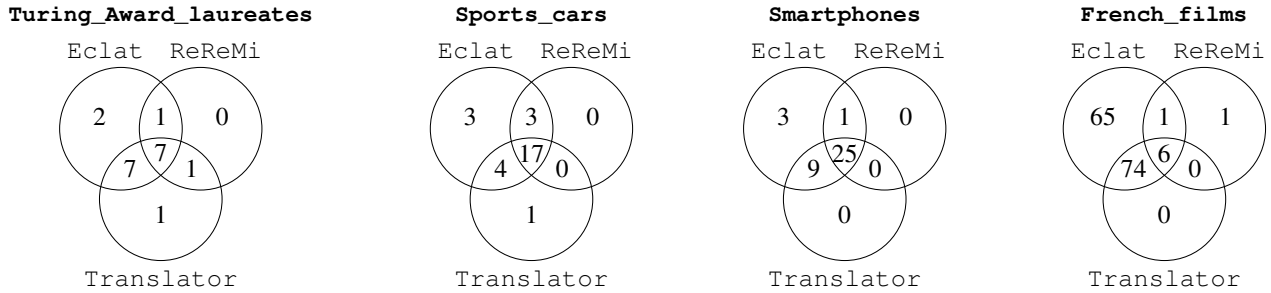


FIGURE 7 – Catégories définies par `Eclat`, `ReReMi` et `Translator` pour chaque jeu de données.

giques, et que des travaux dans ce sens existent pour les règles d'association, ce n'est pas le cas pour les règles de traduction. Cela implique de revoir le calcul de la longueur d'une règle.
Concernant l'évaluation, il nous faudra un moyen de comparer plus formellement les règles entre elles, peut-être *via* l'introduction d'une notion de redondance, qui permettrait de définir une base de règles.

Remerciements

Ce travail est financé avec le support de la Direction Générale de l'Armement et de la Région Lorraine. Merci à Esther Galbrun pour ses conseils ainsi que Quentin Brabant et Jérémie Nevin pour leur aide à l'évaluation des règles.

Références

R. AGRAWAL, T. IMIELIŃSKI et A. SWAMI : Mining association rules between sets of items in large databases. *In ACM SIGMOD Rec.*, vol. 22, p. 207–216. ACM, 1993.

M. ALAM, A. BUZMAKOV, V. CODOCEDO et A. NAPOLI : Mining definitions from rdf annotations using formal concept analysis. *In IJCAI*, p. 823–829, 2015.

F. BAADER, D. CALVANESE, D. MCGUINNESS, D. NARDI et P. PATEL-SCHNEIDER, éds. *The Description Logic Handbook*. Cambridge University Press, Cambridge, UK, 2003.

S. FERRÉ et P. CELLIER : Graph-FCA in Practice. *In Proceedings of 22nd ICCS*, p. 107–121, 2016.

L. A. GALÁRRAGA, C. TEFLIOUDI, K. HOSE et F. M. SUCHANEK : AMIE : association rule mining under incomplete evidence in ontological knowledge bases. *In WWW'13*, p. 413–422, 2013.

E. GALBRUN et P. MIETTINEN : From Black and White to Full Color : Extending Redescription Mining Outside the Boolean World. *Statistical Analysis and Data Mining*, 5(4):284–303, 2012.

B. GANTER et R. WILLE : *Formal concept analysis - mathematical foundations*. Springer, 1999.

M. KLEMETTINEN, H. MANNILA, P. RONKAINEN, H. TOIVONEN et A. I. VERKAMO : Finding interesting rules from large sets of discovered association rules. *In CIKM'94*, p. 401–407, 1994.

N. RAMAKRISHNAN, D. KUMAR, B. MISHRA, M. POTTS et R. F. HELM : Turning CARTwheels : an Alternating Algorithm for Mining Redescriptions. *In KDD'04*, p. 266–275, 2004.

B. SERTKAYA : A survey on how description logic ontologies benefit from formal concept analysis. *CoRR*, abs/1107.2822, 2011.

M. van LEEUWEN et E. GALBRUN : Association Discovery in Two-View Data. *TKDE*, 27(12):3190–3202, déc. 2015.

M. J. ZAKI : Scalable algorithms for association mining. *TKDE*, 12(3):372–390, May 2000.

Découverte de cardinalité maximale contextuelle dans les bases de connaissances

E. A. Sidi Aly[1,2] M. L. Diakité[1] A. Giacometti[2] B. Markhoff[2] A. Soulet[2]

[1] Département Mathématiques et Informatique - Université de Nouakchott Al Aasriya (Mauritanie)
[2] Laboratoire d'Informatique Fondamentale et Appliquée de Tours - Université de Tours (France)

arbi2fr@yahoo.fr, diakite@una.mr, prenom.nom@univ-tours.fr

3 juillet 2018

Résumé

Les bases de connaissances du web sémantique doivent être enrichies par des informations utiles aux applications de fouille, de recherche d'information, de question-réponse, etc. En effet, leur génération à partir de plateformes collaboratives ou d'intégration de sources diverses produit des manques d'information, d'une part, et des erreurs ou incohérences d'autre part. Heureusement, leur volume important permet d'en induire des contraintes vraisemblables. Tel est l'objet de l'algorithme présenté dans cet article, qui extrait des règles de cardinalité maximale à partir d'une base de connaissances. L'enrichissement de la base par ces nouveaux axiomes permet d'y trouver plus de faits, positifs ou négatifs, ce qui rend plus pertinentes les évaluations de la qualité des règles générées par des algorithmes de fouille classiques. Les expérimentations menées sur une partie de DBpedia et sur l'ensemble d'une base de connaissances numismatiques démontrent la faisabilité de l'approche et la pertinence des contraintes extraites.

Mots Clef

Découverte de cardinalité, base de connaissances.

Abstract

The big semantic web knowledge bases have to be enriched for applications in data mining, information retrieval, question answering, etc. Indeed, their generation from collaborative platforms or integration of various sources leads to lack of information on the one hand, and inconsistencies on the other hand. Fortunately, their volume makes it possible to induce probable constraints. This is the aim of the algorithm presented in this article, which extracts maximum cardinality rules from a knowledge base. Adding these new axioms to the knowledge base allows applications to find more facts, positive or negative, which makes more relevant the evaluations of the quality of the rules generated by traditional datamining algorithms. Experiments conducted on part of DBpedia and on an entire numismatic knowledge base demonstrate the feasibility of the approach and the relevance of the discovered contextual constraints.

Keywords

Cardinality discovery, knowledge base.

1 Introduction

Nous considérons de grandes bases de connaissances du web, construites par des algorithmes de recherche d'information à partir de plateformes collaboratives (e.g., DBpedia [2]) et/ou d'intégration de sources diverses. Pour en désigner les éléments, nous utilisons les termes *concept*, *rôle* et *individu* au sens des logiques de description.

Contexte et motivations En représentation des connaissances les restrictions numériques précisant le nombre d'occurrences d'un rôle sont particulièrement utiles [3]. Parmi elles, les contraintes de cardinalité maximale permettent de savoir quand toutes les assertions sur un rôle donné pour un individu donné existent dans la base. C'est utile pour qualifier les réponses aux requêtes sur une base de connaissances, c'est-à-dire les compléter par des informations précises sur leur qualité en terme de *rappel* par rapport à une réalité [13, 17].
Il est illusoire d'espérer des ajouts manuels de telles contraintes d'intégrité dans de grandes bases de connaissances [1], qui soient correctes et suffisantes. Aussi, des techniques de type rétro-ingénierie [14] applicables sur ces grandes bases doivent être considérées, afin de les rechercher systématiquement. Des propositions existent déjà pour trouver des contraintes de clés [1, 11, 15, 16] dans des données RDF. Mais à notre connaissance, il n'y a pas encore de travaux sur l'extraction de contraintes de cardinalité maximale dans les bases de connaissances.

Challenge L'extraction de contraintes de cardinalité à partir des données existantes est connue comme un problème important de la rétro-ingéniérie des bases de données relationnelles [14, 18]. Par rapport au cadre des bases de données traditionnelles, ce problème est bien plus complexe pour les bases de connaissances du web.
Tout d'abord, ces bases de connaissances contiennent généralement des **données incohérentes**, que ce soient

1. [5] présente néanmoins un outil pour le faire sur Wikidata.

 ©AfIA 2018

des assertions fausses ou des assertions dupliquées. De ce fait, la cardinalité maximale observée pour un rôle donné ne saurait être considérée comme sa cardinalité maximale la plus vraisemblable. Par exemple, il est vraisemblable qu'une personne ait au plus une année de naissance et deux parents. Pourtant dans DBpedia (voir les rôles `dbo:birthYear` et `dbo:parent` dans la table 1), certaines personnes ont 5 années de naissance ou 6 parents ! Ces quelques assertions incohérentes ne doivent pas influencer la caractérisation des cardinalités maximales.

Ensuite, ces bases de connaissances sont souvent **incomplètes pour un rôle donné**. Pour cette raison, la cardinalité la plus observée n'est pas forcément la cardinalité maximale. Typiquement, la plupart des personnes décrites dans DBpedia n'ont qu'un seul parent renseigné (voir le rôle `dbo:parent` dans la table 1). Toutefois, certaines en ont plus et ceci n'est pas une anomalie, il faut en tenir compte : la cardinalité maximale du rôle `dbo:parent` pour une personne ne doit pas être sous-estimée (ici à 1) au vu de l'ensemble des cardinalités observées.

Enfin, des travaux récents sur la détection de contraintes de clefs dans les bases de connaissances [16] ont montré que de nombreuses contraintes intéressantes ne sont **valides que sur une partie** d'une base de connaissances. Par exemple, s'il semble difficile de déterminer une cardinalité maximale pour le nombre de nationalités d'une personne en général, comme certains états limitent le nombre de nationalités à 1 il est possible de détecter cette limite pour les ressortissants de tels états. Il est donc essentiel non seulement de détecter des cardinalités maximales, mais aussi d'identifier *les contextes* dans lesquels de telles contraintes peuvent être détectées.

Contributions Etant donnée une distribution de cardinalités $(n_i)_{i\geq 1}$ observées dans une base de connaissances $\mathcal{K}$ pour un rôle R dans un contexte C, nous commençons par proposer **une méthode de calcul d'une cardinalité maximale vraisemblable**, en calculant une estimation du taux de cohérence *réel* que la cardinalité i soit maximale. Cette estimation, notée τ_i, est calculée en prenant en compte tous les individus pour lesquels le rôle R est complet. Son calcul est détaillé et justifié dans la section 4.2. Pour être statistiquement valide, une version corrigée de cette estimation du taux de cohérence, notée $\widetilde{\tau}_i$, est également introduite. Des exemples d'estimations de taux cohérence, corrigés ou non, sont représentés dans la table 1 pour les rôles `dbo:birthYear`, `dbo:parent` et `dbo:nationality` en considérant le concept `dbo:Person` comme contexte.

Etant donnée une arborescence de concepts constituant les contextes candidats, nous proposons ensuite **un algorithme d'exploration systématique d'un ensemble de contraintes contextuelles** pour les rôles desquels nous recherchons les cardinalités maximales. Cet algorithme, décrit en section 4.3, vise à limiter les calculs en élaguant un maximum des contraintes possibles.

Enfin nous présentons et analysons des résultats expérimentaux obtenus sur une base de connaissances résultant d'un processus d'intégration de 5 bases de données numismatiques [6].

`dbo:Person` / `dbo:birthYear`			
i	n_i	τ_i	$\widetilde{\tau}_i$
5	1	1.0	0.0
4	2	0.667	0.0
3	4	0.571	0.0
2	91	0.928	0.775
1	**159841**	**0.999**	**0.996**

`dbo:Person` / `dbo:parent`			
i	n_i	τ_i	$\widetilde{\tau}_i$
6	1	1.0	0.0
4	9	0.9	0.420
3	75	0.882	0.718
2	**9392**	**0.991**	**0.975**
1	10643	0.529	0.518

`dbo:Person` / `dbo:nationality`			
i	n_i	τ_i	$\widetilde{\tau}_i$
8	2	1,000	0,000
6	1	0,333	0,000
5	1	0,250	0,000
4	13	0,765	0,397
3	167	0,908	0,796
2	3 263	0,947	0,921
1	123 386	0,973	0,969

TABLE 1 – Distributions de cardinalités de rôles de personnes dans DBpedia (i est la cardinalité ; n_i le nombre d'individus étant i fois sujets du rôle considéré ; τ_i est une estimation fréquentielle du taux de cohérence réel ; $\widetilde{\tau}_i$ en est une version corrigée s'appuyant sur la borne de Hoeffding)

2 Etat de l'art

Notre algorithme vise à augmenter la connaissance sur les données contenues dans les grandes bases de connaissances du web, en termes de validité comme en termes de complétude par rapport à la réalité représentée. Il permet d'enrichir la partie schéma (TBox en logiques de description) de ces bases pour mieux utiliser leur partie données (ABox). Plusieurs travaux récents vont dans ce sens [1, 11, 15, 16, 10] et d'autres s'en rapprochent [7, 13, 17] mais ciblent des individus (assertions de la ABox) plutôt que des concepts (assertions de la TBox).

Dans [17], une technique de fouille de textes de Wikipedia pour ajouter des précisions sur le degré de complétude des informations dans Wikidata est décrite. Notre proposition est complémentaire puisque notre algorithme traite les données déjà contenues dans les bases de connaissances. Mais surtout, il ne caractérise pas les rôles par rapport à des *individus* précis mais à des *concepts* définis (au sens des logiques de description). Les auteurs de [7, 13] présentent également des propositions pour déterminer quand est-ce qu'un rôle particulier (comme `dbo:parent`) manque pour un individu particulier (comme *Obama*). Plus générale, notre proposition consiste à calculer les cardinalités maximales vraisemblables des rôles relativement à des concepts

définissant des contextes : elle enrichit donc la partie schéma.

Ce sont des clés au sens des bases de données, donc des contraintes au niveau du schéma, qui sont recherchées dans [1, 11, 15, 16]. L'idée est de trouver des axiomes indiquant que tout individu d'un certain concept doit posséder une valeur unique pour un rôlé donné R. Cela constitue donc une cardinalité maximale du rôle R pour le concept C. Egalement très proches de nos travaux, dans [10], les auteurs proposent de déterminer automatiquement quels rôles devraient être obligatoirement renseignés pour un concept donné de la base de connaissances. Pour cela ils comparent la densité du rôle pour les individus de ce concept par rapport à sa densité pour les individus d'autres concepts, qui lui sont liés dans la hiérarchie des concepts. Notre proposition s'appuie sur d'autres critères pour calculer la cardinalité maximale du rôle pour un contexte (notion plus générale que seulement les concepts de la base). Elle peut être adaptée au calcul de la cardinalité minimale, auquel cas elle trouverait, entre autres, quels rôles ont une cardinalité minimale au moins supérieure à 1 pour un concept donné, soit plus d'information que seulement savoir si le rôle devrait exister ou pas.

Ces différentes sortes d'information supplémentaire sur la qualité des données de la base de connaissances, en termes de validité comme en termes de complétude par rapport à la réalité représentée, permettent d'améliorer le fonctionnement des applications qui les utilisent, en réduisant le flou de l'hypothèse du monde ouvert. Ainsi pour améliorer la mesure de qualité de règles issues de processus de fouille dans les bases de connaissances du web sémantique, une *hypothèse de complétude partielle* est définie et utilisée dans [8, 7] : cette règle stipule que si un rôle est renseigné pour un individu, alors les informations concernant ce rôle pour cet individu sont considérées complètes. Si on peut noter que cette hypothèse est contredite par l'observation de DBpedia (voir l'extrait fourni dans la table 1), elle rend tout de même plus précis le calcul de la confiance associée aux résultats de fouille. Ces auteurs ont démontré le besoin pour la fouille de ce qu'ils appellent des *oracles de complétude*, et proposé un certain nombre d'heuristiques pour en définir, comme par exemple la popularité des individus (qui augmente les chances que les faits renseignés sur eux soient complets), etc.

La fouille de données est loin d'être le seul domaine qui bénéficie d'axiomes tels que ceux découverts par notre algorithme, par exemple, s'appuyant sur des travaux de référence en base de données, les auteurs de [4, 12] et plus récemment [9] proposent de caractériser les réponses obtenues par des requêtes, en fonction des informations connues concernant le degré de complétude de la base de connaissances interrogée, par rapport à la réalité représentée.

3 Préliminaires

3.1 Bases de connaissances

Dans ce papier, nous considérons des *bases de connaissances* $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ où $\mathcal{T}$ et $\mathcal{A}$ sont respectivement les TBox et ABox de $\mathcal{K}$. $\mathcal{T}$ désigne un ensemble d'axiomes terminologiques définis à partir des concepts et rôles atomiques de $\mathcal{K}$, alors que $\mathcal{A}$ désigne l'ensemble des assertions ou faits de $\mathcal{K}$. Plus précisément, $\mathcal{A}$ contient des expressions de la forme $C(a)$ et $R(a, b)$ où C est un *concept*, R est un *rôle*, et a, b sont des *individus*.
Dans le cas de la base de connaissances DBpedia, `dbo:Country` et `dbo:Person` sont des exemples de concepts atomiques et `dbo:nationality` est un exemple de rôle atomique de sa TBox. Par ailleurs, `dbo:Country`$(Mauritania)$ et `dbo:nationality`$(Arby, Mauritania)$ sont des exemples de faits ou assertions de sa ABox. Le premier indique que $Mauritania$ est un pays, alors que le second indique que $Arby$ est de nationalité mauritanienne.

Les logiques de description permettent de définir des axiomes pour enrichir la TBox d'une base de connaissances. Par exemple, la relation d'inclusion $\sqsubseteq$ permet d'indiquer qu'un concept C_1 est inclus dans un concept C_2, noté $C_1 \sqsubseteq C_2$. Plus précisément, une base de connaissances $\mathcal{K}$ implique l'axiome $C_1 \sqsubseteq C_2$ si pour toute interprétation $\mathcal{I}$ de $\mathcal{K}$, $C_1^{\mathcal{I}} \subseteq C_2^{\mathcal{I}}$. Par exemple, les axiomes $\exists$`dbo:nationality`$.\top \sqsubseteq$ `dbo:Person` et $\exists$`dbo:nationality`$^{-}.\top \sqsubseteq$ `dbo:Country` indiquent respectivement que le domaine du rôle `dbo:nationality` est inclus dans le concept `dbo:Person`, et que le codomaine du rôle `dbo:nationality` est inclus dans le concept `dbo:Country`.

3.2 Contraintes contextuelles de cardinalité maximale

Soit R un rôle d'une base de connaissances $\mathcal{K} = (\mathcal{T}, \mathcal{A})$. On considère généralement que ce rôle satisfait dans $\mathcal{K}$ une contrainte de cardinalité maximale M si pour tout sujet s, le nombre d'objets o tels que $R(s, o)$ soit présent dans $\mathcal{K}$ (directement présent dans sa ABox $\mathcal{A}$ ou inférable à partir de ses TBox $\mathcal{T}$ et ABox $\mathcal{A}$) est inférieur ou égal à M.
En logique de description, une telle contrainte peut se représenter par un axiome de la forme $sqsubseteq$ en utilisant le constructeur de restriction de cardinalité $(\leq MR)$. En effet, en terme logique, une base de connaissances $\mathcal{K}$ implique l'axiome $\exists R.\top \sqsubseteq (\leq MR)$ si pour toute interprétation $\mathcal{I}$ de $\mathcal{K}$, $\{x : (\exists y)((x, y) \in R^{\mathcal{I}})\} \subseteq \{x : \#\{y : (x, y) \in R^{\mathcal{I}}\} \leq M\}$ où $\#E$ représente la cardinalité d'un ensemble E.
Plus précisément, dans ce papier, nous cherchons à identifier des contraintes *contextuelles* de cardinalité maximale, à savoir des contraintes qui ne sont pas nécessairement vérifiées par tous les sujets s d'un rôle R, mais par tous les sujets instances d'un concept, qu'il soit atomique ou composé, déjà défini dans $\mathcal{K}$ ou pas. Cette notion est introduite

©AfIA 2018

formellement dans la définition suivante :

Définition 3.1 (Contrainte contextuelle)**.** *Etant donnés un rôle R, un concept atomique ou défini C et un entier M, une contrainte contextuelle de cardinalité maximale définie sur R est une expression γ de la forme : $C \sqsubseteq (\leq MR)$.*
Le concept C est appelé le contexte de la contrainte γ. La contrainte γ est satisfaite dans une base de connaissances $\mathcal{K}$ si et seulement si pour toute interprétation $\mathcal{I}$ de $\mathcal{K}$, $C^{\mathcal{I}} \subseteq \{x : \#\{y : (x, y) \in R^{\mathcal{I}}\} \leq M\}$.

Par exemple, la contrainte contextuelle $(\texttt{dbo:Person}) \sqsubseteq (\leq 5\ \texttt{dbo:nationality})$ indique que toutes les personnes ont au plus 5 nationalités, alors que la contrainte contextuelle $(\texttt{dbo:Person} \sqcap \exists \texttt{dbo:nationality}.\{China\}) \sqsubseteq (\leq 1\ \texttt{dbo:nationality})$ indique que toutes les personnes de nationalité chinoise ont au plus une nationalité.

Dans ce travail, on cherche à extraire des contraintes contextuelles de cardinalité maximale qui soient les plus générales possibles.

Définition 3.2 (Contrainte contextuelle minimale)**.** *Soient deux contraintes contextuelles de cardinalité maximale γ_1 : $C_1 \sqsubseteq (\leq M_1\ R)$ et $\gamma_2 : C_2 \sqsubseteq (\leq M_2\ R)$ définies sur R.*
La contrainte γ_1 est dite plus générale que la contrainte γ_2 si $C_2 \sqsubset C_1$ et $M_1 \leq M_2$. Etant donné un ensemble de contraintes Γ définies sur R, une contrainte $\gamma_1 \in \Gamma$ est dite minimale dans Γ s'il n'existe aucune contrainte γ_2 dans Γ plus générale que γ_1.

Par exemple, la contrainte contextuelle $(\texttt{dbo:Person}) \sqsubseteq (\leq 2\ \texttt{dbo:nationality})$ est plus générale que la contrainte contextuelle $(\texttt{dbo:Person} \sqcap \exists \texttt{dbo:nationality}.\{China\}) \sqsubseteq (\leq 5\ \texttt{dbo:nationality})$ car $(\texttt{dbo:Person} \sqcap \exists \texttt{dbo:nationality}.\{China\}) \sqsubseteq \texttt{dbo:Person})$ et $2 \leq 5$.

La notion de minimalité a pour objectif de ne pas extraire de contraintes contextuelles qui soient redondantes. Intuitivement, considérons les deux contraintes γ_1 et γ_2 introduites dans la définition précédente, et supposons que γ_1 soit plus générale que γ_2. Etant donnée une base de connaissances $\mathcal{K}$ dans laquelle les contraintes γ_1 et γ_2 sont satisfaites, soit une instance s de C_2 dans $\mathcal{K}$. D'après γ_2, nous savons que pour toute interprétation $\mathcal{I}$ de $\mathcal{K}$, $\#\{o : (s, o) \in R^{\mathcal{I}}\} \leq M_2$. Mais comme γ_1 est plus générale que γ_2, nous savons par définition que $C_2 \sqsubseteq C_1$. Il en découle que s est aussi une instance de C_1 dans $\mathcal{K}$, et d'après γ_1, que pour tout interprétation $\mathcal{I}$ de $\mathcal{K}$, $\#\{o : (s, o) \in R^{\mathcal{I}}\} \leq M_1$, ce qui est une contrainte plus forte que $\#\{o : (s, o) \in R^{\mathcal{I}}\} \leq M_2$. En effet, par définition de la minimalité, nous savons que $M_1 \leq M_2$. Par rapport à la contrainte γ_1, la contrainte γ_2 est donc inutile car redondante, i.e. elle ne permet pas de déduire d'information supplémentaire.

Le problème traité dans ce papier est alors le suivant : **étant donnés une base de connaissances $\mathcal{K}$, un rôle R et une hiérarchie de concepts $(\mathcal{C}, \sqsubseteq)$, nous cherchons à découvrir l'ensemble des contraintes contextuelles de cardinalité maximale de la forme $C \sqsubseteq (\leq M\ R)$ avec $C \in \mathcal{C}$, qui soient satisfaites sur $\mathcal{K}$ et minimales dans $\mathcal{C}$.**

En pratique, une base de connaissances telle que DBpedia est très incomplète (par exemple, de nombreuses personnes ont seulement un parent), et elle comporte de nombreuses incohérences (par exemple, des personnes peuvent avoir jusqu'à 5 parents). Pour ces raisons, étant donnée une base de connaissances $\mathcal{K}$, il n'est pas pertinent de chercher à extraire des contraintes de cardinalité qui soient *parfaitement* satisfaites dans $\mathcal{K}$, mais les contraintes :

— *les plus probables et suffisamment probables* par rapport à un seuil donné, de manière à prendre en compte et tolérer les incohérences, et
— *suffisamment certaines* par rapport à un degré de confiance, pour ne pas extraire des contraintes qui soient remises en cause régulièrement par l'ajout de nouveaux faits dans la base de connaissances.

Nous détaillons dans la section suivante comment évaluer la probabilité qu'une contrainte soit satisfaite dans une base de connaissances $\mathcal{K}$ et comment mesurer la certitude que cette contrainte soit réelle.

4 Extraction de contraintes contextuelles de cardinalité maximale

Pour résoudre le problème énoncé précédemment, nous commençons par le reformuler en introduisant la notion de taux de cohérence dans la section 4.1, puis nous décrivons dans la section 4.2 comment détecter une cardinalité maximale pour un rôle R dans un contexte C. Ensuite, étant donné un ensemble de contextes candidats $\mathcal{C}$, nous montrons dans la section 4.3 comment explorer efficacement l'ensemble des contraintes contextuelles possibles.

4.1 Taux de cohérence

Etant donnée une base de connaissances $\mathcal{K} = (\mathcal{T}, \mathcal{A})$, supposons que i soit la cardinalité maximale du rôle R dans le contexte C. Soit s un individu de C dans $\mathcal{K}$, complet pour le rôle R dans $\mathcal{K}$ (dans le sens où tous les faits $R(s, o)$ possibles représentant le monde réel sont dans $\mathcal{A}$ ou inférables). Dans le cas où il existe exactement i faits dans $\mathcal{K}$ de la forme $R(s, o)$, cela renforce l'hypothèse que i soit la cardinalité maximale de R dans le contexte C. Inversement, s'il existe plus de i faits dans $\mathcal{K}$ de la forme $R(s, o)$, cela affaiblit l'hypothèse que i soit la cardinalité maximale de R dans le contexte C. Ainsi dans le tableau 1, pour la classe `dbo:Person`, les individus comportant au moins 3 assertions pour le rôle `dbo:parent` affaiblissent l'hypothèse que la cardinalité maximale soit 2 mais ils restent peu nombreux au regard des 9 392 individus qui ont exactement 2 parents.
En suivant ce raisonnement, nous introduisons la notion de

taux de cohérence pour évaluer si une cardinalité i pour le rôle R dans le contexte C a des chances d'être maximale :

Définition 4.1 (Taux de cohérence)**.** *Etant donnée une base de connaissances $\mathcal{K}$, le taux de cohérence de la cardinalité i pour le rôle R dans le contexte C est le ratio :*

$$\tau_i^{C,R}(\mathcal{K}) = \frac{n_i^{C,R}}{n_{\geq i}^{C,R}}$$

où $n_i^{C,R}$ (resp. $n_{\geq i}^{C,R}$) représente le nombre de sujets s tels que i faits $R(s,o)$ (resp. i faits ou plus) appartiennent à $\mathcal{K}$ dans le contexte C.

Par exemple, dans le tableau 1, $n_{\geq 2}^{\texttt{dbo:Person},\texttt{dbo:parent}}$ est égal à 9477 ($9477 = 9392{+}75{+}9{+}1$). De cette manière, le taux de cohérence $\tau_2^{\texttt{dbo:Person},\texttt{dbo:parent}}(\mathcal{K})$ est de 0,991 (i.e., $9392/9477$). Par la suite, quand le contexte et la relation sont clairs, nous pouvons les omettre dans les notations. Dans ce cas, n_i et τ_i désignent respectivement les termes $n_i^{C,R}$ et $\tau_i^{C,R}$.
Maintenant nous allons formaliser le lien entre le taux de cohérence et la notion de contrainte maximale. Originellement introduit dans [13], $\mathcal{K}^* = (\mathcal{T}^*, \mathcal{A}^*)$ désigne une hypothétique base de connaissances idéale qui contiendrait tous les axiomes et toutes les assertions du monde réel. Comme $\mathcal{K}^*$est correcte et complète, **le taux de cohérence au sein de $\mathcal{K}^*$, noté $\tau_M^{C,R}(\mathcal{K}^*)$, est égal à 1 si et seulement si $C \sqsubseteq (\leq M\ R)$ appartient à $\mathcal{T}^*$.**
En pratique, le taux de cohérence mesuré dans une base de connaissances est différent du taux de cohérence réel : $\tau_i(\mathcal{K}) \neq \tau_i(\mathcal{K}^*)$. Par exemple, le taux de cohérence $\tau_2(\mathcal{K})$ pour le rôle `dbo:parent` du tableau 1 est égal à 0,991 alors que le taux de cohérence réel est égal à 1. Plus grave, on a $\tau_6^{\texttt{dbo:Person},\texttt{dbo:parent}}(\mathcal{K}) = 1$! Le taux de cohérence sur $\mathcal{K}$ est donc une estimation peu fiable du taux de cohérence réel sur $\mathcal{K}^*$.

4.2 Détection d'une contrainte

L'estimation $\tau_i(\mathcal{K})$ de $\tau_i(\mathcal{K}^*)$ doit être corrigée pour être statistiquement valide. Pour ce faire, nous proposons d'utiliser l'inégalité de Hoeffding qui a l'avantage d'être vraie pour toute distribution. En terme de probabilité, si X est une variable aléatoire indiquant pour un sujet s tiré aléatoirement, le nombre de faits $R(s,o)$ appartenant à $\mathcal{K}$, alors τ_i est une estimation fréquentielle de la probabilité conditionnelle $P(X = i \ / \ X \geq i)$. Etant donné un niveau de confiance $1-\delta$, l'inégalité de Hoeffding stipule que $\tau_i(\mathcal{K}^*)$ est compris entre $\tau_i(\mathcal{K}) - \epsilon_i$ et $\tau_i(\mathcal{K}) + \epsilon_i$ où $\epsilon_i = \sqrt{\frac{\log(1/\delta)}{2n_{\geq i}}}$. Dans ce contexte, afin de prendre des décisions les plus sûres, nous proposons d'utiliser la borne inférieure de l'intervalle de confiance $[\tau_i - \epsilon_i, \tau_i + \epsilon_i]$. Plus formellement, on a la propriété suivante :

Propriété 4.1 (Minoration)**.** *Etant données une base de connaissances $\mathcal{K}$ et une confiance $1-\delta$, le taux de cohérence réel $\tau_i(\mathcal{K}^*)$ de la cardinalité i pour le rôle R dans le contexte C est supérieur à $\widetilde{\tau}_i(\mathcal{K})$:*

$$\tau_i(\mathcal{K}^*) \geq \widetilde{\tau}_i(\mathcal{K})$$

où $\widetilde{\tau}_i(\mathcal{K})$ est le taux de cohérence pessimiste défini par :

$$\widetilde{\tau}_i(\mathcal{K}) = \max\left\{\frac{n_i}{n_{\geq i}} - \sqrt{\frac{\log(1/\delta)}{2n_{\geq i}}}; 0\right\}$$

Cette propriété nous munit d'un outil efficace pour approximer le taux de cohérence réel. Il survient alors la difficulté de choisir la cardinalité maximale une fois que l'on dispose pour chaque cardinalité i du taux de cohérence pessimiste $\widetilde{\tau}_i(\mathcal{K})$ (pour un rôle R dans le contexte C).
Plus précisément, **étant donnés un seuil minimal de cohérence min_τ et un niveau de confiance $1-\delta$, nous considérons que M est la cardinalité maximale de R dans le contexte C si et seulement si $\widetilde{\tau}_M \geq min_\tau$ et** $M = \arg\max_{i\geq 1} \widetilde{\tau}_i(\mathcal{K})$.
Quelques exemples d'estimations $\widetilde{\tau}_i$ et de détection de cardinalités maximales contextuelles sont donnés dans la table 1. Dans les 3 exemples, on a considéré `dbo:Person` comme contexte, et on a cherché à détecter la cardinalité maximale contextuelle de trois rôles : `dbo:birthYear`, `dbo:parent` et `dbo:nationality`. Intuitivement, pour les deux premiers rôles, on souhaiterait détecter des cardinalités maximales respectives de 1 et 2. Pour un niveau de confiance $1-\delta = 99\%$ et un seuil $min_\tau = 0.97$, on constate que les cardinalités maximales supposées sont effectivement détectées (cf. lignes en gras dans la table 1). De manière intéressante, avec ces mêmes seuils, aucune cardinalité n'est détectée pour `dbo:nationality`.

4.3 Exploration de l'espace de recherche

Etant donnés une base de connaissances $\mathcal{K}$, un rôle R, une arborescence de concepts $(\mathcal{C}, \sqsubseteq)$, un degré de confiance δ et un seuil minimal de cohérence min_τ, nous cherchons à découvrir l'ensemble des contraintes contextuelles de cardinalité maximale de la forme $C \sqsubseteq (\leq M\ R)$ avec $C \in \mathcal{C}$, qui soient minimales et suffisamment certaines sur $\mathcal{K}$. En pratique, notons que l'arborescence $(\mathcal{C}, \sqsubseteq)$ peut être une arborescence déjà existante dans la TBox de la base de connaissances, ou une arborescence construite dans une phase préalable de préparation des données (voir la section 5.1).
Dans un tel cadre, il y a potentiellement un très grand nombre de contraintes contextuelles à considérer, évaluer et comparer. Néanmoins, il est possible de réduire la taille de l'espace de recherche à explorer en se basant sur les propriétés 4.2 et 4.3 énoncées ci-après. Tout d'abord, la propriété 4.2 montre qu'une contrainte $C \sqsubseteq (\leq M\ R)$ ne peut pas être suffisamment certaine si le contexte C contient trop peu d'individus dans $\mathcal{K}$, car alors l'intervalle de confiance du taux de cohérence calculé grâce à l'inégalité de Hoeffding est très large et sa borne inférieure ne peut être supérieure au seuil min_τ imposé.

©AfIA 2018

Propriété 4.2 (Nombre minimal d'observations)**.** *Etant donnés une base de connaissances $\mathcal{K}$, une contrainte contextuelle de cardinalité maximale $C \sqsubseteq (\leq M\ R)$ et un seuil min_τ, le taux de cohérence $\widetilde{\tau}_M(\mathcal{K})$ que M soit la cardinalité maximale de R dans C ne peut être supérieur à min_τ que si $|C| \geq \frac{log(1/\delta)}{2(1-min_\tau)^2}$.*

Par ailleurs, supposons qu'une contrainte γ définie par $C \sqsubseteq (\leq M\ R)$ avec $M = 1$ ait été détectée comme suffisamment certaine au cours de l'exploration. Alors, d'après la propriété 4.3, il n'est pas nécessaire d'explorer les contraintes γ' définies par $C' \sqsubseteq (\leq M'\ R)$ où C' est plus spécifique que C. Cette propriété découle directement de la définition 3.2 de la minimalité.

Propriété 4.3 (Contrainte minimale)**.** *Soient une base de connaissances $\mathcal{K}$ et une contrainte contextuelle de cardinalité maximale γ définie par $C \sqsubseteq (\leq M\ R)$ avec $M = 1$. Toute contrainte γ' définie par $C' \sqsubseteq (\leq M'\ R)$ avec $C' \sqsubseteq C$ et $M' \geq 1$ ne peut être minimale.*

L'algorithme 1 détaille notre fonction récursive d'exploration, la fonction $C3M$ (pour *Contextual Cardinality Constraint Mining*). Cette fonction prend en entrée une base de connaissances, un rôle à explorer, un contexte courant, une cardinalité maximale courante ($M = \infty$ si aucune cardinalité maximale n'a encore pu être détectée), et enfin, des seuils δ et min_τ. **Le démarrage de l'exploration d'une arborescence de racine $\top$ se fait en exécutant la fonction** $C3M(\mathcal{K}, R, \top, \infty, \delta, min_\tau)$.
Pour commencer, la fonction $C3M$ détermine si le nombre d'individus est suffisant dans le contexte C. Si ce n'est pas le cas, elle arrête l'exploration à la ligne 2 conformément à la propriété 4.2. Sinon, le taux de cohérence $\widetilde{\tau}_i$ est calculé pour chaque cardinalité i (lignes 4 à 6) et la ligne 7 retient la cardinalité maximale la plus probable. Si le taux de cohérence correspondant n'est pas supérieur au seuil min_τ, alors cela signifie qu'aucune cardinalité maximale n'a pu être détectée à ce niveau et i_M est fixé ligne 8 à ∞. Ensuite, si la cardinalité maximale détectée i_M est strictement inférieure à M (la cardinalité maximale détectée au niveau précédent), alors on dispose d'une nouvelle contrainte minimale de cardinalité maximale i_M et on l'ajoute à Γ, l'ensemble des contraintes recherchées. Finalement, conformément à la propriété 4.3, si i_M est égale à 1, il n'est pas nécessaire de poursuivre l'exploration en parcourant les contextes plus spécifiques de C. Sinon, la fonction $C3M$ est appelée récursivement à la ligne 12 pour tous les C' qui sont des sous-concepts directs de C.
Dans notre implémentation de la fonction $C3M$, nous avons appliqué une approche client-serveur où les distributions de cardinalité $n_i^{C,R}$ sont calculées par interrogation en SPARQL d'une base de connaissances localisée sur un serveur. Dans un tel cadre, la complexité de notre méthode en nombre de requêtes sur le serveur est en $\mathcal{O}(|\mathcal{C}|)$ où $|\mathcal{C}|$ représente le nombre de concepts dans l'arborescence $\mathcal{C}$ explorée. Dans le pire des cas, côté client, la complexité en nombre d'opérations est en $\mathcal{O}(|\mathcal{C}| \times i_{max})$ où i_{max} représente l'entier maximal pour lequel il existe au moins un sujet s tel que i_{max} faits $R(s, o)$ appartiennent à la base de connaissances $\mathcal{K}$, i.e. $i_{max} = \arg\max_{i\in\mathbb{N}}\{n_i^{\top,R} > 0\}$.

Algorithm 1 C3M

Input: Une base de connaissances $\mathcal{K}$, un rôle R, un contexte C, un entier M, un niveau de confiance δ et un seuil minimal de support min_τ
Output: Un ensemble Γ de contraintes contextuelles de cardinalité maximale

1: $\alpha := \frac{log(1/\delta)}{2(1-min_\tau)^2}$ et $n_{\geq 0}^{C,R} := |C|$
2: **if** $(n_{\geq 0}^{C,R} < \alpha)$ **then return** $\emptyset$
3: $\Gamma := \emptyset$ et $i_{max} := \arg\max_{i\in\mathbb{N}}\{n_i^{C,R} > 0\}$
4: **for all** $i \in [1..min\{M, i_{max}\}]$ **do**
5: $\quad \widetilde{\tau}_i := \max\left\{\frac{n_i^{C,R}}{n_{\geq i}^{C,R}} - \sqrt{\frac{\log(1/\delta)}{2n_{\geq i}^{C,R}}}; 0\right\}$
6: **end for**
7: $i_M := \arg\max_{i\in[1..min\{M,i_{max}\}]}\{\widetilde{\tau}_i\}$
8: **if** $(\widetilde{\tau}_{i_M} < min_\tau)$ **then** $i_M = \infty$
9: **if** $(i_M < M)$ **then** $\Gamma := \{C \sqsubseteq (\leq i_M\ R)\}$
10: **if** $(i_M > 1)$ **then**
11: $\quad$ **for all** $C' \in subClassOf(C)$ **do**
12: $\quad\quad \Gamma := \Gamma \cup C3M(\mathcal{K}, R, C', i_M, \delta, min_\tau)$
13: $\quad$ **end for**
14: **end if**
15: **return** Γ

5 Expérimentations

Outre les requêtes sur DBpedia (dont nous montrons des échantillons de réponses en table 1), qui ont été utilisées pour mettre au point la définition du taux de cohérence, nous avons expérimenté l'algorithme 1 sur un jeu de données mis à notre disposition par les auteurs de [6].

5.1 Données et protocole

Le jeu de données utilisé porte sur le domaine numismatique, il est le résultat d'un processus d'intégration mené dans le cadre du projet européen ARIADNE [2]. Ses auteurs ont utilisé le CIDOC-CRM [3] pour intégrer les contenus de 5 ressources construites par des institutions de différents pays européens. Il contient 3 123 998 triplets, dont les définitions de 114 classes et 373 rôles ou propriétés du CIDOC-CRM et d'ARIADNE. Il est stocké et interrogé avec le triplestore Blazegraph (v2.1.4), sur une machine virtuelle sous Linux avec 32 GB de mémoire virtuelle, sur un serveur ayant pour processeur un Dual Intel Xeon E5620 4 coeurs. L'algorithme 1 est implémenté en Java et utilise la bibliothèque de programmation pour RDF Jena [4]. La base porte sur des pièces de monnaies mais, par choix des intégrateurs, il n'existe pas de classe $Coin$. Les individus correspondant à des pièces sont des instances de `E22_Man_Made_Object` caractérisées par certains URIs (ex. `<http://nomisma.org/id/coin>`) comme va-

2. http://ariadne-infrastructure.eu/
3. http://www.cidoc-crm.org/
4. http://jena.apache.org

leur objet de certains rôles (ex. `P2_has_type`). Plusieurs rôles et plusieurs URIs sont utilisés pour cela, aussi nous avons décidé de construire notre propre arborescence d'exploration de la façon suivante :
Au **premier niveau**, notre arborescence contient tous les concepts C_i de la base, soit 114 concepts ($i \in [1..114]$). Tous ces concepts sont des sous-concepts du **concept racine** $\top$ **au niveau zéro**, i.e. pour tout i, nous avons $C_i \sqsubseteq \top$.
Au **deuxième niveau** notre arborescence contient tous les concepts C_i^j définis par $C_i^j := C_i \sqcap (\exists R_j.\top)$ où C_i ($i \in [1..114]$) et R_j ($j \in [1..373]$) sont respectivement des concepts et rôles de la base. A ce niveau, 42 522 concepts C_i^j sont ainsi définis. Enfin, au **troisième niveau**, notre arborescence contient tous les concepts $C_i^{j,k}$ définis par $C_i^{j,k} := C_i \sqcap (\exists R_j.\{a_k\})$ où C_i ($i \in [1..114]$) et R_j ($j \in [1..373]$) sont respectivement des classes et rôles de la base, et a_k est un individu du co-domaine de R_j, i.e. $a_k \in (\exists R_j^{-1}.\top)$. Grâce à ce dernier niveau, il est possible de considérer des contextes à la manière de notre exemple jouet où `dbo:Person` $\sqcap \exists$`dbo:nationality`$.\{China\}$. Notons finalement que pour tout i, j, k, nous avons $C_i^{j,k} \sqsubseteq C_i^j \sqsubseteq C_i$. Globalement, cette arborescence comporte 3 160 357 concepts, donc pour les 373 rôles de la base de connaissances cela représente plus d'un milliard de contraintes contextuelles possibles (exactement 1 178 813 161 contraintes). Néanmoins, comme nous le verrons dans la section suivante, l'utilisation des propriétés 4.2 et 4.3 permet d'élaguer une grande partie de l'espace de recherche.

5.2 Résultats

Tous les résultats présentés dans cette section ont été obtenus avec un **seuil minimal de confiance** $1 - \delta = 0,99\%$ (pour des contraintes les plus certaines possibles) et un **seuil minimal de cohérence** $min_\tau = 0,95$ (pour des contraintes suffisamment probables). Ce seuil a été défini expérimentalement. Sur des bases de connaissances de plus grande taille comme DBpedia, un seuil plus élevé est préférable. Néanmoins, l'approche est relativement peu sensible aux seuils (i.e., l'ensemble des contraintes trouvées est stable).

Analyse quantitative. Avec ces paramètres, la propriété 4.2 nous indique qu'une contrainte $C \sqsubseteq (\leq M\ R)$ ne peut être suffisamment certaine si son contexte C contient moins de $\alpha = \frac{log(1/\delta)}{2(1-min_\tau)^2} = 922$ instances. Ainsi, l'utilisation de la propriété 4.2 permet de n'explorer que 16 641 contraintes, soit moins de $0,002\%$ des plus de 1 milliard de contraintes possibles. Qui plus est, notre expérience montre que la propriété 4.3 permet de réduire encore de $82,5\%$ la taille de l'espace de recherche à explorer. Au final, avec les seuils choisis notre algorithme ne cherche à détecter une cardinalité maximale que pour 2 909 contextes possibles, avec un temps de calcul complet de moins de 50 minutes.
La table 2 donne une vue globale et quantitative du résultat de l'exploration réalisée. Sur les 2 909 contraintes contextuelles possibles, notre algorithme a détecté au total 887 contraintes de cardinalité maximale, 595 d'entre elles étant des contraintes minimales. Sur cet exemple, le critère de minimalité permet donc de réduire de près de 67% le nombre de contraintes retournées. On constate que les contraintes les plus nombreuses sont des cardinalités maximales avec $M = 1$, ce qui correspond à des contraintes où pour un rôle donné R, tout sujet s est en relation avec au plus un objet o. Néanmoins de très nombreuses contraintes sont trouvées avec des cardinalités maximales $M \in \{2, 3\}$ (37% des contraintes minimales détectées). On note également que si des contraintes de cardinalités maximales sont détectées dès le niveau 0 (65 contraintes avec un contexte $C \equiv \top$), la recherche de contraintes contextuelles est particulièrement pertinente. Il faut en effet noter que les contraintes les plus nombreuses sont trouvées au niveau 3 (75% des contraintes détectées), sachant que par construction de notre arborescence, c'est à ce niveau que sont caractérisées les pièces de monnaie.

	Niveau dans l'arborescence				
M	**0** $\top$	**1** C_i	**2** C_i^j	**3** $C_i^{j,k}$	**Total**
1	60	28	10	222	320
2	3	6	9	90	108
3	0	7	14	92	113
4	1	0	8	20	29
5	1	0	0	16	17
6	0	0	0	8	8
Total	**65**	**41**	**41**	**448**	**595**

TABLE 2 – Répartition par niveau et cardinalité maximale M des contraintes minimales détectées

Analyse qualitative. Tout d'abord, dès le niveau 0, notre méthode permet de retrouver des contraintes fonctionnelles attendues, par exemple, pour les 3 rôles du CIDOC-CRM `P1_is_identified_by`, `P52_has_current_owner` et `P50_has_current_keeper`, indiquant que si un sujet décrit dans la base possède plus d'un identifiant, un propriétaire ou un conservateur, alors on peut en déduire que ces identifiants (respectivement, propriétaires et conservateurs) sont identiques. Concernant le rôle `P45_consists_of` du CIDOC-CRM (permettant de décrire les matériaux constitutifs d'un objet), il est intéressant de noter qu'une cardinalité maximale de 2 est détectée dès le niveau 1 pour la classe `E22_Man_Made_Object`. La base de connaissances décrit notamment des médailles constituées d'or et de pierre précieuse (telle l'agate). Pour ce même rôle, une cardinalité maximale de 1 est détectée au niveau 3 pour les pièces de monnaie. Cette information est notamment représentée par la contrainte `E22_Man_Made_Object` $\sqcap$ $\exists$`P2_has_type`$.\{$`<http://nomisma.org/id/coin>`$\} \sqsubseteq$ (≤ 1 `P45_consists_of`). Cette contrainte est détectée bien qu'à certaines pièces la relation `P45_consists_of` associe deux matériaux ; mais c'est rare (et le plus souvent il

©AfIA 2018

s'agit du même matériau dans deux langues différentes). Un même type de contrainte (avec $M = 1$) est trouvée au niveau 3 pour tous les contextes décrivant des pièces, concernant le rôle `P62_depicts` (ce qui est dépeint sur l'objet). C'est raisonnable car dans le cas d'une pièce de monnaie, on trouve le plus souvent une seule représentation figurative (sur une des deux faces de la pièce), alors qu'une telle contrainte n'est pas valide pour d'autres objets.

Au passage, l'étude de l'ensemble des contraintes extraites par notre méthode a mis en évidence des redondances dans la base, sans doute du fait des choix d'intégration. Dans une phase de post-traitement, la connaissance d'axiomes tel que $\exists \texttt{P2_has_type}.\{\ldots coin\} \sqsubseteq \exists \texttt{Thing_has_type_Concept}.\{\ldots moneta\}$ pourrait réduire encore le nombre de contraintes extraites.

6 Conclusion

Nos expérimentations démontrent la faisabilité d'une exploration systématique d'une base de connaissances, à la recherche de contraintes contextuelles de cardinalité maximale, grâce à l'algorithme que nous proposons dans cet article : dans le cas étudié, cela prend moins d'une heure pour une base de connaissances contenant plus de 3 millions de triplets, décrits par une centaine de concepts et plus de 300 rôles. Les propriétés utilisées par notre algorithme font que seules 595 contraintes ont été obtenues, ce qui reste analysable manuellement. Cela nous a permis de vérifier que ces contraintes sont bien pertinentes dans le contexte de la base étudiée. De plus, nos expérimentations démontrent l'importance du contexte dans cette découverte de contraintes.

Il s'agit à notre connaissance de la première proposition de calcul de contraintes contextuelles de cardinalité maximale dans une base de connaissances du web sémantique. Ces grandes bases de connaissances, reflet d'une intelligence collective, sont générées à partir de l'expertise limitée de nombreux contributeurs et souffrent encore, tantôt de lacunes dans les informations, tantôt d'incohérences. Utiliser leurs contenus courants afin de mieux caractériser les connaissances représentées est donc très utile, comme montré dans l'état de l'art : cela permet aux applications qui exploitent ces grandes bases de connaissances de produire des résultats plus fiables.

Nous avons donc pour perspective d'exploiter les contraintes extraites pour calculer la confiance associée à des règles découvertes dans la base de connaissances ainsi enrichie. Mais avant cela, nous travaillons sur des post-traitements pour réduire encore le nombre de contraintes présentées en résultat. Pour cela, nous explorons le potentiel des raisonnements possibles sur la TBox, en particulier comment les relations de subsomption entre classes peuvent éliminer des redondances dans les ensembles de contraintes extraites.

Références

[1] Atencia, M., David, J., Scharffe, F. : Keys and pseudo-keys detection for web datasets cleansing and interlinking. In : EKAW. pp. 144–153. Springer (2012)

[2] Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., Ives, Z. : Dbpedia : A nucleus for a web of open data. In : The semantic web, pp. 722–735. Springer (2007)

[3] Baader, F., Sattler, U. : Expressive number restrictions in description logics. Journal of Logic and Computation 9(3), 319–350 (1999)

[4] Darari, F., Nutt, W., Pirrò, G., Razniewski, S. : Completeness statements about rdf data sources and their use for query answering. In : ISWC. pp. 66–83. Springer Berlin Heidelberg (2013)

[5] Darari, F., Razniewski, S., Prasojo, R.E., Nutt, W. : Enabling Fine-Grained RDF Data Completeness Assessment. In : Web Engineering. pp. 170–187. Springer International Publishing, Cham (2016)

[6] Felicetti, A., Gerth, P., Meghini, C., Theodoridou, M. : Integrating heterogeneous coin datasets in the context of archaeological research. In : EMF-CRM@ICTPDL. pp. 13–27. CEUR-WS.org (2015)

[7] Galárraga, L., Razniewski, S., Amarilli, A., Suchanek, F.M. : Predicting completeness in knowledge bases. In : WSDM. pp. 375–383. ACM (2017)

[8] Galárraga, L.A., Teflioudi, C., Hose, K., Suchanek, F. : Amie : Association rule mining under incomplete evidence in ontological knowledge bases. In : WWW. pp. 413–422. ACM (2013)

[9] Galárraga, L., Hose, K., Razniewski, S. : Enabling Completeness-aware Querying in SPARQL. In : Proceedings of WebDB. pp. 19–22. ACM (2017)

[10] Lajus, J., Suchanek, F.M. : Are All People Married ? Determining Obligatory Attributes in Knowledge Bases . In : WWW (2018)

[11] Pernelle, N., Saïs, F., Symeonidou, D. : An automatic key discovery approach for data linking. Web Semantics : Science, Services and Agents on the World Wide Web 23, 16–30 (2013)

[12] Razniewski, S., Korn, F., Nutt, W., Srivastava, D. : Identifying the extent of completeness of query answers over partially complete databases. In : SIGMOD. pp. 561–576. ACM (2015)

[13] Razniewski, S., Suchanek, F., Nutt, W. : But what do we actually know ? In : 5th Workshop on Automated Knowledge Base Construction. pp. 40–44 (2016)

[14] Soutou, C. : Relational database reverse engineering : algorithms to extract cardinality constraints. Data & Knowledge Engineering 28(2), 161–207 (1998)

[15] Symeonidou, D., Armant, V., Pernelle, N., Saïs, F. : Sakey : Scalable almost key discovery in RDF data. In : ISWC. pp. 33–49. Springer (2014)

[16] Symeonidou, D., Galárraga, L., Pernelle, N., Saïs, F., Suchanek, F. : Vickey : Mining conditional keys on knowledge bases. In : ISWC. pp. 661–677. Springer (2017)

[17] Tanon, T.P., Stepanova, D., Razniewski, S., Mirza, P., Weikum, G. : Completeness-aware rule learning from knowledge graphs. In : ISWC. pp. 507–525. Springer (2017)

[18] Yeh, D., Li, Y., Chu, W. : Extracting entity-relationship diagram from a table-based legacy database. Journal of Systems and Software 81(5), 764–771 (2008)

Activité AfIA

1er août 2016 – 31 juillet 2018

Éditeurs : Conseil d'Administration de l'AfIA – Année 2018

Table des matières

Flyer AfIA 2018.
À propos de l'AfIA v

Conseil d'Administration de l'AfIA.
CA AfIA vii

Collèges AfIA.
Collèges RR, SIC et SMAA ix

Prix de Thèse IA.
2017 et 2018 xv

Bulletins Trimestriels.
Du n°094 au n°101 xvii

Journées propres.
PDIA et FIIA xix

Compétitions.
Compétitions xxv

Journées communes.
Journées communes xxix

Plates-formes IA.
PFIA 2017, PFIA 2018 xlvii

A propos de l'AFIA

L'objet de l'AFIA, Association Loi 1901 sans but lucratif, est de promouvoir et de favoriser le développement de l'Intelligence Artificielle (IA) sous ses différentes formes, de regrouper et de faire croître la communauté française en IA, et, à la hauteur des forces de ses membres, d'en assurer la visibilité.

L'AFIA anime la communauté par l'organisation de grands rendez-vous annuels. L'AFIA organise ainsi chaque année une semaine de l'IA, la « Plate-forme IA » (PFIA 2017 Caen, PFIA 2018 Nancy) au sein de laquelle se tiennent la « Conférence Nationale d'Intelligence Artificielle » (CNIA), les « Rencontres des Jeunes Chercheurs en IA » (RJCIA) et la « Conférence sur les Applications Pratiques de l'IA » (APIA) ainsi que des conférences thématiques qui peuvent évoluer d'une année à l'autre, sans récurrence obligée.

C'est ainsi que du 2 au 6 juillet 2018 à Nancy, lors de son édition 2018, la Plate-Forme IA hébergera, outre les 21ème CNIA, 16ème RJCIA et 4ème APIA, les 29èmes IC et les 13èmes JFPDA. Y sont également programmées cinq journées communes organisées avec d'autres associations ou institutions. La Plate-Forme IA 2018, quant à elle, se tiendra du 1au 5 juillet 2019 à Toulouse, dates à réserver sur vos agendas !

Forte du soutien de ses 514 adhérents à jour de leur cotisation en 2017, l'AFIA assure :

- Le maintien d'un Site Web dédié à l'IA, reproduisant également les Brèves de l'IA,
- Une journée recherche annuelle sur les Perspectives et Défis en IA (PDIA 2017),
- Une journée industrielle annuelle ou Forum Industriel en IA (FIIA 2018),
- La remise annuelle d'un Prix de Thèse de Doctorat en IA,
- Le soutien à plusieurs Collèges, actuellement au nombre de 4, ayant leur propre activité :
    - Collège Industriel (depuis janvier 2016),
    - Collège Science de l'Ingénierie des Connaissances (depuis avril 2016),
    - Collège Systèmes Multi-Agents et Agents Artificiels (depuis octobre 2016),
    - Collège Représentation et Raisonnement (depuis avril 2017)
- La parution trimestrielle des Bulletins de l'AFIA, en accès libre à tous depuis le site web,
- Un lien entre adhérents sur les réseaux sociaux LinkedIn, Facebook et Twitter,
- Le parrainage, scientifique, mais aussi éventuellement financier, aux événements en IA,
- La diffusion mensuelle de Brèves sur les actualités de l'IA en France (inscription aux Brèves)
- La réponse aux consultations officielles ou officieuses (Ministères, Missions, Organismes)
- La réponse aux questions de la presse écrite et de la presse orale, également sur internet.

L'AFIA organise aussi mensuellement des Journées communes avec d'autres Associations (en 2018 : EGC & IA avec EGC, Santé & IA avec AIM, TAL & IA avec ATALA, Statistiques et IA avec la SFdS) et avec d'autres institutions (en 2018 : Ethique & IA avec INRIA/COERLE, IA pour l'Education avec la MEN/DNE, Rob & IA avec le CNRS / GdR Robotique).

Enfin, l'AFIA contribue à la participation de ses membres aux grands événements de l'IA. Ainsi, les membres de l'AFIA, pour leur inscription à PFIA, bénéficient-ils d'une réduction équivalente à deux fois le coût de leur adhésion à l'AFIA, leur permettant d'assister à PFIA 2018 sur 5 jours au tarif de 95,00 € !

Rejoignez-nous vous aussi et adhérez à l'AFIA pour contribuer au développement de l'IA en France. L'adhésion peut être individuelle ou, à partir de cinq adhérents, être réalisée au titre de personne morale (institution, laboratoire, entreprise). Pour adhérer, il suffit de vous rendre sur le site des adhésions de l'AFIA.

Merci également de susciter de telles adhésions en diffusant ce document autour de vous !

AfIA
Association française
pour l'Intelligence Artificielle

Conseil d'Administration 2017

Président	Yves DEMAZEAU	GT Collèges, GT Communication, GT Conférences, PFIA
V.-Président	Pierre ZWEIGENBAUM	Parrainages, Fiches, Collèges, Livre Blanc, PDIA 2017
Trésorier	Catherine FARON-ZUCKER	GT Finances, Livre Blanc, PFIA
Secrétaire	Olivier BOISSIER	GT Livre Blanc, Prix de Thèse, PFIA, E&IA 2017
Webmestre	Audrey BANEYX	GT Web & Mail, Site Web, Cloud
Rédacteur	Florence BANNAY	GT Bulletins, Nuit Info, IA Dossier n°097
Membres	Carole ADAM	GT Compétitions, Brèves, Votes, Ouvrage 2017
	Emmanuel ADAM	Fiches, Web & Mail, MACS&IA 2017
	Patrick ALBERT	Communication, Livre Blanc
	Olivier AMI	Bulletins, Dossier n°096, Réseaux sociaux
	Sandra BRINGAY	GT Cartographie, Web & Mail
	Frédéric MARIS	Démonstrations PFIA
	Arnaud MARTIN	Postmaster, Site Web, Web & Mail
	Engelbert MEPHU NGUIFO	GT Journées, Collèges, Journées PFIA, Dossier n°098
	Davy MONTICOLO	Web & Mail, EIAH&IA 2017
	Philippe MORIGNOT	Bulletins, Compétitions, Collèges
	Jean-Denis MULLER	Parrainages, Fiches, Collèges, FIIA 2017
	Philippe MULLER	Cartographie, Collèges, Livre Blanc, Dossier n°095
	Bruno PATIN	Collège Industriel, FIIA 2017
	Serena VILLATA	Adhésions, Trésorerie, Livre Blanc, FR@IJCAI 2017

Conseil d'Administration 2018

Président	Yves DEMAZEAU	GT Collèges, GT Communication, GT PFIA, EF&IA 2018
V.-Président	Elise BONZON	Parrainages, Collèges, Communication
Trésorier	Catherine FARON-ZUCKER	GT Finances, Collèges, Enseignement, PFIA
Secrétaire	Sandra BRINGAY	Cartographie, Enseignement, Prix Thèse, PFIA
Webmestre	Arnaud MARTIN	GT Web & Mail, Postmaster, Dossier n°103
Rédacteur	Dominique LONGIN	GT Bulletins, Communication, Ouvrage 2018
Membres	Carole ADAM	Brèves, Compétitions, Votes
	Emmanuel ADAM	Brèves, Bulletins, Twitter, Web & Mail
	Patrick ALBERT	Communication
	Florence BANNAY	Nuit Info, Dossier n°100, Compétitions
	Alain BERGER	Dossier n°099, Collège Industriel, Communication
	Florence BANNAY	Nuit Info, Dossier n°100, Compétitions, Enseignement
	Frédéric MARIS	Démonstrations PFIA, FIIA 2018
	Juliette MATTIOLI	FIIA 2018, LinkedIn, Web & Mail, FIIA 2018
	Engelbert MEPHU NGUIFO	GT Journées, Dossier n°102, Sta & IA 2018, PFIA
	Davy MONTICOLO	Facebook, Web & Mail, PFIA
	Philippe MORIGNOT	Bulletins, Compétitions
	Bruno PATIN	Collège Industriel, Dossier n°101
	Gauthier PICARD	GT Enseignement, Plateforme
	Olivier SIMONIN	Compétitions, Dossier n°104, PDIA 2018
	Serena VILLATA	Adhésions, Finances, Journées, FR@IJCAI 2018

Collège Représentation et Raisonnement

Objectif du Collège

L'objectif du collège est d'animer les communautés de recherche françaises en Représentation et Raisonnement. Les thématiques de recherche sont relatives aux méthodes et outils fondamentaux de l'Intelligence Artificielle. Elles portent sur :

- La définition de modèles de représentation des informations (croyances, connaissances, préférences, obligations et permissions, actions, incertitude, confiance, réputation) comme les langages des logiques classiques ou non classiques, les modèles possibilistes, les ontologies, les langages à base de contraintes, les représentations graphiques, etc.,

- La définition et l'automatisation de raisonnements sur ces informations : raisonnement spatio-temporel, dynamique des informations, révision de croyances, fusion d'informations symboliques, raisonnement par argumentation, raisonnement causal, raisonnement abductif, raisonnement à partir de cas, etc.,

- La perspective algorithmique et de représentation pour des concepts utilisés dans des théories connexes comme la théorie des jeux ou la théorie du choix social (équilibre, stratégie gagnante, manipulation, etc.) : théorie des jeux algorithmique et choix social computationnel,

- La mise au point de méthodes de codage des informations et d'algorithmes de traitement efficaces : compilation de connaissances, SAT, ASP, etc.,

- La modélisation formelle de l'interaction : entre utilisateurs et systèmes informatiques, entre entités informatiques autonomes (agents),

- Et généralement le lien avec différentes techniques liées à la décision, la planification, l'ordonnancement, le diagnostic, l'apprentissage, les sciences des données, etc.

Ces thématiques couvrent de très nombreux contextes d'application, comme par exemple le Web sémantique, le Web des données, les systèmes de recommandation ou d'aide à la décision, les agents conversationnels et assistants personnels, la programmation des jeux, la robotique, etc.

Programme de travail

Le collège R&R accompagne l'organisation régulière des événements suivants :

- IAF. Journées d'Intelligence Artificielle Fondamentales. Ces journées ont lieu tous les ans, à l'initiative du comité IAF. Elles sont articulées autour d'exposés de synthèse invités, ainsi que d'un programme constitué après appel à communication.

- JFPDA. Journées Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes. Ces journées ont lieu tous les ans, à l'initiative des anciens présidents de comités de programme et des participants à la liste de diffusion jfpda@loria.fr. Elles sont articulées autour d'exposés de synthèse invités, ainsi que d'un programme constitué après appel à communication.

- LFA. Rencontres Francophones sur la Logique Floue et ses Applications. Ces journées ont lieu tous les ans, à l'initiative du comité de pilotage LFA. Elles sont articulées autour de 3 ou 4 conférences invitées, ainsi que d'un programme constitué après appel à communication.

Le collège contribution également à l'organisation de journées communes de l'AFIA, (en 2017, Jeux Informatisés et IA avec le pré-GDR AFAIA, et RO&IA avec la ROADEF ; en 2018, journée commune avec le pré-GdR AFAIA, thème à préciser).

Le collège sert également, pour l'AFIA, de structure d'interaction avec le Pré-GdR AFAIA (Aspects Formels et Algorithmiques de l'Intelligence Artificielle).

Comité de Pilotage

Le comité de pilotage du collège est constitué des personnes suivantes, représentatives en particulier des conférences IAF, JFPDA, et LFA, et / ou sont également engagées dans le pré-GdR AFAIA.

- Elise Bonzon
  Université Paris-Descartes, LIPADE, Paris

- Laurent Garcia
  Université d'Angers, LERIA, Angers

- Sébastien Konieczny
  CNRS, CRIL, Lens

- Tiago de Lima
  Université d'Artois, CRIL, Lens

- Jean Lieber
  Université de Lorraine, LORIA, Nancy

- Fréderic Maris
  Université Paul Sabatier, IRIT, Toulouse

- Nicolas Maudet
  Université Pierre et Marie Curie, LIP6, Paris

- Amedeo Napoli
  CNRS, LORIA, Nancy

- Henri Prade
  CNRS, IRIT, Toulouse

- Bruno Zanuttini
  Université Caen Normandie, GREYC, Caen

Contacts

Coordinateur du Collège : frederic.maris@irit.fr

Listes de diffusion : bull-i3@irit.fr, bull-ia@gdria.fr, jfpda@loria.fr

Deux membres du comité de pilotage du collège sont membres du Conseil d'Administration de l'AFIA :

- Elise Bonzon, elise.bonzon@mi.parisdescartes.fr

- Fréderic Maris rederic.maris@irit.fr

Collège Science de l'Ingénierie des Connaissances

Objectif du Collège

À l'heure numérique, les données et les outils se multiplient mais assurer un accès intelligent aux données reste un défi et ce malgré les langages et les technologies qui sont maintenant à disposition des informaticiens et des experts de domaine. Partager des données et des connaissances au sein d'une communauté, d'une entreprise ou sur le web suppose leur explicitation, leur représentation, leur mise en relation, leur diffusion et leur maintenance. L'ingénierie des connaissances est au cœur de ces problématiques.

L'essor des sciences et technologies de l'information et de la communication, et notamment des technologies du web, dans l'ensemble de la société engendre des mutations dans les pratiques individuelles et collectives. L'ingénierie des connaissances accompagne cette évolution, en inventant les modèles, les méthodes, les algorithmes et les outils permettant l'intégration de connaissances et de raisonnements automatiques sur ces connaissances dans des environnements informatiques.

De fait, on constate aujourd'hui une omniprésence des connaissances, en même temps qu'un décloisonnement des communautés de recherche en Intelligence Artificielle pour répondre aux problèmes qui se posent. Ainsi, la construction des systèmes à base de connaissances fait appel à l'ingénierie des connaissances, au traitement automatique des langues, à l'apprentissage automatique et à la fouille de données ; le traitement des connaissances associe ingénierie des connaissances avec recherche d'information, recherche opérationnelle, analyse de données ; la modélisation des environnements informatiques, de ses acteurs humains ou logiciels, allie ingénierie des connaissances et systèmes multi-agents ; enfin l'ingénierie des connaissances touche à des thématiques classiques dans les communautés de l'interaction homme-machine et des environnements informatiques pour l'apprentissage humain.

En plaçant ainsi l'utilisateur au cœur des systèmes informatiques, l'ingénierie des connaissances est sollicitée pour assister l'utilisateur dans le traitement de la masse de données disponibles, pour les visualiser, contextualiser, classer, prédire, approximer, sécuriser, fiabiliser, recommander, expliquer, valider, pour pallier leur imprécision voire incohérence.

Programme de Travail

Le collège Ingénierie des Connaissances a d'abord un rôle de diffusion de l'information. Il gère une liste de diffusion sur l'ingénierie des connaissances (qui compte près de 800 inscrits), et maintient une page web du collège IC sur le site Web de l'AFIA.

Le collège contribue également à l'animation et au dynamisme de la communauté de recherche en Ingénierie des Connaissances. Il est moteur dans l'organisation de la conférence nationale d'Ingénierie des Connaissances (IC), des Journées Francophones des Ontologies (JFO) et des Rencontres du Web des Données. Ses membres contribuent également à la fabrication de certaines « Journées Communes » de l'AFIA avec d'autres sociétés savantes, notamment les journées « Informatique Médicale et Intelligence Artificielle » (IM&IA), celles sur « Extraction et Gestion des Connaissances et Intelligence Artificielle » (EGC&IA), à l'organisation d'ateliers sur la Plate-Forme Intelligence Artificielle (PFIA) ou à l'occasion d'autres événements tels RISE.

Comité de Pilotage

Le comité de pilotage du collège est constitué de 10 chercheurs spécialisés dans le domaine de l'Ingénierie des Connaissances, tous membres des comités de programme de la conférence nationale

d'ingénierie des Connaissances (IC) et/ou de la conférence Journées Francophones des Ontologies (JFO), et la plupart des présidents d'éditions passées de la conférence nationale d'Ingénierie des Connaissances :

- Marie-Hélène Abel, membre du comité de programme de IC
  Université de technologie de Compiègne, HEUDIASYC

- Djamal Benslimane, membre du CP JFO
  Université Claude Bernard Lyon 1, LIRIS

- Sandra Bringay, membre des comités de programme de IC et de JFO
  Université Paul Valéry Montpellier 3, LIRMM

- Jean Charlet, membre des comités de programme de IC
  Assistance Publique-Hôpitaux de Paris, INSERM, LIMICS

- Sylvie Després, membre des comités de programme de IC et de JFO
  Université Paris 13, LIMICS

- Catherine Faron-Zucker, membre du comité de programme de IC
  Université Nice Sophia Antipolis, I3S et Inria SAM

- Nathalie Hernandez, membre des comités de programme de IC et de JFO
  Université Jean Jaurès Toulouse 2, IRIT

- Nathalie Pernelle, membre du comité de programme de IC
  Université Paris Sud, LRI

- Raphaël Troncy, membre du comité de programme de IC
  EURECOM

- Haifa Zargayouna, membre des comités de programme de IC et de JFO
  Université Paris 13, LIPN

Contacts

Coordinatrice du Collège SIC : sylvie.despres@univ-paris13.fr

Liste de diffusion Ingénierie des Connaissances : info-ic@inria.fr

Deux membres du comité de pilotage du collège sont membres du conseil d'administration de l'AFIA :

- Catherine Faron Zucker, faron@unice.fr

- Sandra Bringay, sandra.bringay@lirmm.fr

Collège Systèmes Multi-Agents et Agents Autonomes

Objectif du Collège

Le collège SMAA a pour mission de contribuer aux activités menées par l'AFIA, par des actions relatives au domaine de l'IA, émanant de la communauté des Agents Autonomes et des Systèmes Multi-Agents (SMAA). Les thématiques de recherche couvertes par ce collège sont celles relevant des Journées Francophones en Systèmes Multi-Agents (JFSMA) et du Workshop Affects, Compagnon Artificiel, Interactions (WACAI). Elles portent sur la modélisation à base d'agents logiciels mis en interaction au sein d'environnements partagés pour :

- La résolution collective de problème (modèles de coordination ; apprentissage, adaptation et planification distribuées ; décisions sous incertitudes, etc.),
- La simulation de phénomènes complexes où la modélisation multi-agents apporte un cadre conceptuel permettant la représentation et la simulation de systèmes faisant intervenir différentes entités en interaction
- Le développement de systèmes informatiques décentralisés où l'approche SMA permet l'intégration flexible et la coopération de logiciels et de services autonome
- Le développement de systèmes médiatisés dans lesquels utilisateurs humains et agents artificiels interagissent directement ou indirectement, dans le cadre d'activités collectives de type éducatif, culturel ou social
- L'étude, la modélisation, le développement et l'évaluation de l'interaction de systèmes interactifs avec leurs partenaires (humains ou artefacts).

Les recherches menées sur ces thèmes adoptent fréquemment une démarche pluridisciplinaire (biologie, linguistique, psychologie, sociologie…) dans la modélisation de sociétés d'agents artificiels ou pour la prise en compte de l'activité et des interactions avec des utilisateurs humains au sein de communautés mixtes.

Programme de Travail

Les missions du collège SMAA concernent le soutien à l'organisation de manifestations scientifiques (conférences, ateliers), l'animation de groupes de travail, l'édition de dossiers techniques ou de numéros spéciaux de journaux sur des thématiques d'intérêt pour la communauté et la diffusion et la communication autour des recherches des communautés françaises sur les disciplines ciblées.

Le collège SMAA accompagne notamment l'organisation régulière des JFSMA et de WACAI en sollicitant un parrainage de l'AFIA sous la forme de remise de prix, réduction des tarifs d'inscription, annonces d'événements sur ses relais de communication… Il propose également un soutien similaire à d'autres événements, ponctuels ou récurrents, relevant de son périmètre scientifique.

Il contribue aux actions initiées par l'AFIA sur la mise en place de journées bilatérales, notamment les journées communes « Réalité Virtuelle & Intelligence Artificielle » (RV & IA), « Interaction Homme-Machine & Intelligence Artificielle » (IHM & IA) ainsi que « Modélisation, Analyse et Conduite des

Systèmes dynamiques & Intelligence Artificielle » (MACS & IA) ; par une participation aux actions du collège industriel de l'AFIA ; par des contributions à la rubrique « Collèges » du bulletin de l'AFIA.

Les actualités du collège SMAA et plus largement des communautés françaises de recherche associées aux conférences JFSMA et WACAI sont publiées régulièrement sur son site web : http://smaa.lip6.fr/

Comité de Pilotage

Le comité de pilotage du collège SMAA est constitué de chercheurs spécialisés dans le domaine des systèmes multi-agents et des agents autonomes, tous membres des comités de programme ou du comité consultatif des journées francophones en systèmes multi-agents (JFSMA) et/ou du workshop Affect, Compagnon Artificiel, Interaction (WACAI).

- Emmanuel Adam
  Université de Valenciennes et du Hainaut-Cambrésis, LAMIH, Valenciennes

- Rémy Courdier
  Université de la Réunion, LIM, Saint Denis.

- Catherine Garbay
  CNRS, LIG, Grenoble

- Zahia Guessoum
  Université de Reims, CReSTIC/LIP6, Reims

- Salima Hassas
  Université Claude Bernard Lyon 1, LIRIS, Lyon

- Fabien Michel
  Université de Montpellier, LIRMM, Montpellier

- Frédéric Migeon
  Université Paul Sabatier, IRIT, Toulouse

- Maxime Morge
  Université de Lille 1, CRIStAL, Lille

- Nicolas Sabouret
  Université Paris-Sud, LIMSI, Saclay

- Laurent Vercouter
  INSA Rouen Normandie, LITIS, Rouen

Contacts

Coordinateur du Collège : laurent.vercouter@insa-rouen.fr

Listes de diffusion : sma@loria.fr, acai@poleia.lip6.fr

Un membre du comité de pilotage du collège est membre du conseil d'administration de l'AFIA :

- Emmanuel Adam emmanuel.adam@univ-valenciennes.fr

AfIA
Association française
pour l'Intelligence Artificielle

Prix de Thèse IA 2017

Le Jury

- Président : Jean-Gabriel Ganascia (LIP6, UMPC)

- Membres : Leila Amgoud, Nicolas Bredeche, Laurence Devillers, Amal Elfallah-Seghrouchni, Jérôme Euzenat, Jean-Gabriel Ganascia, Brigitte Grau, Jérôme Lang, Pierre Marquis, Henri Prade, Guillaume Raschia, Marie Christine Rousset, Julie Velcin.

Les Lauréats

- Prix de Thèse IA (ex-aequo) : Michaël Perrot "Théorie et Algorithmes pour l'Apprentissage de Métriques à Comportement Contrôlé". 13/12/16. Dir. : Amaury Habrard (LHC, Univ. St. Etienne).

- Prix de Thèse IA (ex-aequo) : Eric Piette "Une nouvelle approche au *General Game Playing* dirigée par les contraintes". 09/12/16. Dir. : Sylvain Lagrue et Frédéric Koriche (CRIL, Univ. Artois).

- Accessit : Fabrice Mayran de Chamisso "Navigation exploratoire au long de la vie : une approche intégrant planification, navigation, cartographie et localisation pour des robots mobiles disposant de ressources finies". 18/11/16. Dir. : Michaël Aupetit et Laurent Soulier (LIST, Univ. Paris 11).

Prix de Thèse IA 2018

Le Jury

- Président : Jean-Gabriel Ganascia (LIP6, UMPC)

- Membres : Rachid Alami, Gauvain Bourgne, Nicola Bredeche, Stephane Canu, Raja Chatila, Mohamed Chetouani, Hélène Fargier, Jean-Gabriel Ganascia, Andreas Herzig, Frederic Jurie, Jérôme Lang, Myriam Lewkowicz, Pierre Marquis, Benjamin Piwowarski, Marie Christine Rousset, Salem Benferhat, Thomas SChiex, Julien Velcin.

Les Lauréats

- Prix de Thèse IA (ex-aequo) : Nawal Benabbou "Procédures de décision par élicitation incrémentale de préférences en optimisation multicritère, multi-agents et dans l'incertain". 05/05/2017. Dir. : Patrice Perny (LIP6, Univ. Paris 06).

- Prix de Thèse IA (ex-aequo) : Daniele Calandriello "Efficient Sequential Learning". 18/12/2017. Dir. : Michal Valko (CRISTAL, Univ. Lille 1) et Alessandro Lazaric (CRISTAL, Univ. Lille 1).

- Accessit : Théo Trouillon "Complex-Valued Embeddings Models for Knowledge Graphs". 29/07/2017. Dir. : Christopher Dance (Naver Labs, Grenoble) et Eric Gaussier (LIG, Univ. Grenoble Alpes).

Bulletins

Le bulletin de l'AFIA fournit un cadre de discussions et d'échanges au sein de la communauté. Toutes les contributions, pour peu qu'elles aient un intérêt général, sont les bienvenues. Le bulletin contient des rubriques régulières de compte-rendu des conférences, journées et autres événements que l'AFIA organise ou parraine, les résumés d'Habilitations à Diriger des Recherches et de Thèses de Doctorat, et un Dossier qui dresse un état de l'art sur un domaine particulier de l'IA, ou présente des équipes françaises de recherche en IA, académiques ou industrielles. Les bulletins de l'AFIA sont accessibles librement depuis le site de l'AFIA.

Le Comité de Rédaction

Le comité de rédaction 2017, comprenait Florence Bannay comme rédactrice en chef, aidée de Claire Lefèvre, Dominique Longin et Philippe Morignot Le comité de rédaction 2018, désormais sous la responsabilité de Dominique Longin comprend Emmanuel Adam, Claire Lefèvre, et Philippe Morignot.

Numéro 94 Dossier « RFIA 2016 » Octobre 2016

- Le numéro contient un dossier sur RFIA 2016 (avec un éditorial de Yves Demazeau, Président de notre association), la conférence française sur notre domaine.
- Le numéro contient les compte-rendu des journées Francophones sur les Ontologies, de la journée Rob&IA organisée avec le GdR Robotique, de RF&IA organisée avec l'AFRIF. On y trouve aussi un compte-rendu des JFRB 2016. Le groupe de travail "Compétition" de l'AFIA nous fait un compte rendu de la première compétition IA sur Robots. On y trouve aussi les prix de Thèse IA 2016.

Bulletin 95 Dossier « Apprentissage » janvier 2017

- Le dossier de ce bulletin porte sur l'apprentissage. Il est composé d'une réflexion sur l'apprentissage, et, dans la continuité de PDIA 2015, par la présentation des travaux menés dans deux équipes : Equipe Interaction Algorithme Silicium (Télécom Bretagne) sur le thème de l'acquisition et la mémorisation d'informations mentales, Equipe Sequel (INRIA Lille-Nord Europe) sur le thème transfert et apprentissage multitâches dans l'apprentissage basé sur le renforcement.
- Ce bulletin inclut un compte-rendu de la journée Classification et IA organisée avec la SFC, de celle organisée avec le GdR MaDICS, de la journée PDIA 2016 sur le thème de l'impact social de l'IA. On y trouve aussi un compte-rendu d'une journée du groupe Représentation et Algorithmes en Pratique du pré-GdR IA, et un compte-rendu du défi de l'AFIA pour la Nuit de l'Info 2016.

Bulletin 96 Dossier « Entreprises de France » avril 2017

- Ce dossier contient les compte-rendu de la journée EF&IA organisée avec le MEDEF, de la journée « Philosophie des sciences et IA » organisée avec la SPS et le DEC, de l'intervention de l'AFIA à EduSpot, et celui de la 3e journée IHM&IA organisée avec l'AFIHM.
- Le dossier de ce bulletin contient le verbatim des discours tenus lors de la journée EF&IA. Plus de 510 inscrits à cette journée, qui a été honorée par l'intervention du Secrétaire d'Etat à la Recherche et à l'Enseignement Supérieur, M. Thierry Mandon.

Bulletin 97 Dossier « Equipes de recherche » juillet 2017

- De nombreuses équipes ont répondu à l'appel, le dossier spécial contient la description d'équipes de différents laboratoires français dont LGI2P, IRISA, IRIT, LIRMM, LIMSI, LIS, et LORIA.

- Ce bulletin présente les compte-rendu de la journée EIAH&IA organisé avec l'ATIEF, du deuxième Forum Industriel de l'AFIA, FIIA 2017, d'une journée commune entre les GdR IM et le pré-GdR AFAIA, ainsi que de la journée Bioss-IA. Il contient enfin deux hommages à deux chercheurs disparus le trimestre dernier : Alain Colmerauer et Daniel Kayser.

Bulletin 98 Dossier « PFIA 2017 » octobre 2017

- Le dossier est consacré à PFIA 2017 qui s'est tenue en juillet à Caen. Elle a hébergé cette année six conférences : APIA, IAF, IC, JFPDA, JFSMA et RJCIA ; mais aussi les journées Éthique et IA, Jeux Informatisés et IA, France@IJCAI qui a permis de présenter, en avant-première, des articles de chercheurs français acceptés à IJCAI 2017, et la compétition IA et Jeux Vidéo : BotContest.
- Ce bulletin retrace les deux journées « Recherche opérationnelle, apprentissage et fouille de données » (RO&IA), et « Apprentissage et Interactions ». Deux articles sont consacrés aux lauréats *ex-æquo* du prix de thèse IA : Michaël PERROT pour sa thèse intitulée « Théorie et Algorithmes pour l'Apprentissage de Métriques à Comportement Contrôlé » et Eric PIETTE pour sa thèse intitulée « Une nouvelle approche au General Game Playing dirigée par les contraintes ».

Bulletin 99 Dossier « État des lieux de la Gestion des Connaissance en France » janvier 2018

- Ce bulletin contient un dossier industriel sur l'« État des lieux de la Gestion des Connaissance en France ». Ce dossier est issu de réflexions menées lors de FIIA 2017. Il s'inspire des diverses contributions réalisées lors du forum, son objet est d'ouvrir les perspectives et de synthétiser une partie des questionnements qui sont associés à la problématique de la Gestion des connaissances.
- Ce bulletin contient les compte-rendu de la Journée PDIA 2017, de la Journée MACS&IA sur la « Conduite des Systèmes Dynamiques et l'IA », du défi de l'AFIA pour la Nuit de l'Info 2017, ainsi que de la seconde journée RI&IA sur « Connaissances ». Il présente en outre les thèses et habilitations en IA soutenues lors du dernier trimestre 2017 qui nous ont été signalées.

Bulletin 100 Dossier « ECCAI/EurAI Fellows » avril 2018

- Ce bulletin est un peu spécial puisqu'il est le 100e que l'AFIA édite ! Le dossier, porte sur la vision qu'ils ont de l'IA d'un certain nombre d'ECCAI/EurAI Fellows français. En particulier, il leur a été demandé de décrire un ou plusieurs articles les ayant particulièrement marqués, un bref état de l'art actuel de leur domaine, ainsi que leur vision de l'évolution de ce domaine dans les dix prochaines années. Un tiré à part de ce dossier a été imprimé et distribué lors de PFIA 2018.
- Ce bulletin contient le compte-rendu de la 3e journée Extraction et Gestion des Connaissances & IA (EGC&IA) organisée conjointement par l'AFIA et EGC. Cette journée visait à offrir un panorama des approches liées à la collecte, le stockage et la gestion des données personnelles en lien étroit avec les problématiques de vie privée et d'éthique. Enfin, vous trouvez également l'ensemble des soutenances de thèses et HDR dont nous avons eu connaissance et soutenues lors du premier trimestre 2018.

Bulletin 101 Dossier « Présentation d'équipes industrielles » juillet 2018

- Le dossier de ce bulletin présente plusieurs équipes de recherche industrielles.
- Ce bulletin contient le compte-rendu de WACAI 2018, Colloque sur les « Affects, Compagnons Artificiels et Interactions » qui s'est déroulé du 13 au 15 juin sur l'Île de Porquerolles, celui de FIIA 2018 qui s'est déroulé le 12 avril 2018 à l'Université Paris-Descartes, ainsi que celui des JFPC 2018 (Journées Francophones de Programmation par Contraintes) qui se sont déroulées du 13 au 15 juin à Amiens. En fin du bulletin, la rubrique « Thèses et HDR du trimestre » comporte la liste des soutenances du 2e trimestre 2018 dont nous avons eu connaissance.

PDIA 2016

L'Association Française pour l'Intelligence Artificielle (AFIA) organise sa troisième journée PERSPECTIVES ET DEFIS DE l'IA sur le thème : IMPACT SOCIAL DE L'IA. Les médias rendent abondamment compte des applications de l'IA, de ses promesses et des craintes qu'elle suscite. D'importantes mutations sont en gestation, la vie sociale va connaître des bouleversements qu'il est déjà possible d'esquisser, leur ampleur et leur calendrier faisant débat. En organisant cette journée, l'AFIA veut contribuer à éclaircir ces horizons nouveaux.

Cette journée se tient sous le patronage du MESR Direction Générale de la Recherche et de l'Innovation et de l'ISCC Institut des Sciences de la Communication avec le soutien de la Fondation Telecom.

Date et Lieu

- Date : 07 octobre 2016
- Lieu : Amphi Thévenin, Télécom ParisTech, 46 rue Barrault, 75013 Paris

Programme

- 09h30 Introduction de la journée par Yves Demazeau (Président de l'AFIA)
- 09h40 Ouverture de Pascal Estraillier *(*Chargé de mission STIC à la MESR DGRI, Direction Générale de la Recherche et de l'Innovation*)*
- 09h50 Ouverture de Pascal Griset *(Directeur de l'ISCC, Institut des Sciences de la Communication)*
- 10h00 « La société du calcul. Enjeux et promesses de la massification des données » par Dominique Cardon (Orange Labs, Paris)
- 10h45 Pause
- 11h00 « Blockchain, Registres distribués & IA », Laurent Leloup (Président de France Blocktech)
- 11h45 « Comment développer l'éthique des algorithmes et la loyauté des plateformes ? Peut-on
- 12h30 Buffet
- 13h45 « Apprendre à vivre avec les systèmes d'IA » par Aymeric Poulain Maubant (PDG Nereÿs)
- 14h00 « l'IA, levier de la profonde mutation de l'industrie financière » par François Leclerc
- 14h45 « L'IA et la métamorphose de la médecine, nouveaux business et intervenants » par Florence Pinaud (La Tribune Hebdo)
- 15h30 Pause
- 15h45 « L'information revisitée » par Eric Scherer (Directeur de la Prospective France Télévisions)
- 16h30 Clôture

Organisation

Cet événement est organisé par Patrick Albert et Olivier Boissier pour l'Association Française d'Intelligence Artificielle (AFIA), avec l'aide de François Leclerc.

Inscriptions

Les inscriptions à la journée (gratuites pour les membres de l'AFIA ou de la Fondation Telecom, 30 € sinon) sont obligatoires et à effectuer sur le site : http://inscriptions.afia.asso.fr/ . Le déjeuner est offert à toutes les personnes qui se seront inscrites à PDIA 2016 avant le 04/10/16, et seulement à celles-ci.

AfIA
Association française
pour l'Intelligence Artificielle

FIIA 2017

Lors de FIIA 2016, il est apparu que la gestion / l'ingénierie des CONNAISSANCES était essentielle pour soutenir tout système d'IA et pouvait de toute façon participer au déploiement d'applications. Ce deuxième FORUM INDUSTRIEL DE L'INTELLIGENCE ARTIFICIELLE est donc consacré à ce sujet, à travers différents thèmes. Chaque thème est composé de courtes présentations suivies d'une mini-table ronde dont l'objectif est d'aboutir à des éléments de feuille de route et de permettre des échanges accrus entre académiques et industriels. Des appels plus généraux concernant l'IA seront également annoncés. Une dernière session est réservée à des présentations rapides par des industriels invités l'ayant souhaité.

Date et Lieu

- Date 27 avril 2017
- Lieu Université Paris-Descartes, Amphithéâtre Fourier, 45 rue des Saints Pères, 75006 Paris

Programme

- 08h30 Accueil
- 08h45 Ouverture de Frédéric Dardel *(Président de l'Université Paris Descartes)* et Pavlos Moraitis *(Directeur du LIPADE),* Introduction par Catherine Faron-Zucker (*Coordinatrice du Collège Science de l'Ingénierie des Connaissances de l'AFIA*) et Bruno Patin *(Coordinateur du Collège Industriel de l'AFIA)*.
- 09h00 Présentation des travaux de l'OPECST sur l'IA par Dominique Gillot *(Sénatrice).*
- 09h15 Thème « Gestion vs. Ingénierie » de(s, la) connaissance(s), avec Jean Charlet *(APH Paris & INSERM),* Alain Berger *(Ardans),* et Nicolas Dubuc *(Michelin).*
- 10h00 Thème « Agents conversationnels », avec Frédérique Segond *(Viseo),* Henri Sanson *(Orange),* Cyril Texier *(DoYouDreamUp)* et Nicolas Sabouret (*Univ. Paris 11*).
- 11h00 Pause-café
- 11h30 Thème « Prise de décision », avec Bruno Patin *(Dassault Aviation),* Ph. Bonnard *(IBM),* J.-Marc David *(Renault),* Laurent Gouzènes *(Pacte Novation)* et J.-Fabrice Lebraty (*Univ. Lyon 3*).
- 12h30 Buffet
- 14h00 Thème « Transparence et confiance », avec Patrick Constant *(Owant-Pertimm),* J.-Pierre Cotton *(Ardans),* Victor Dermiaux *(CNIL)* et J.-Gabriel Ganascia *(Univ. Paris 6).*
- 15h00 « Les projets d'IA s'annoncent » avec Yves Demazeau *(AFIA),* Fabienne Réveillac *(SNCF)* et Pascal Ollivier *(SOGET).*
- 15h30 Pause-café
- 16h00 Présentations rapides de sociétés concernées par l'IA en « Trois planches »
- 17h00 Conclusion par Yves Demazeau *(Président de l'AFIA)*
- 17h30 Clôture

Organisation

Cet événement est organisé par Alain Berger pour le Collège Industriel de l'AFIA et Catherine-Faron-Zucker pour le Collège Science de l'Ingénierie des Connaissances de l'AFIA.

Inscriptions

Les inscriptions au Forum (gratuites pour les membres de l'AFIA, 30 € sinon) sont obligatoires et à effectuer sur le site des inscriptions : http://inscriptions.afia.asso.fr/. Le déjeuner est offert à toutes les personnes qui se seront inscrites à FIIA 2017 avant le 24/04/17, et seulement à celles-ci.

France @ IJCAI 2017

Cette journée propose un avant-goût des recherches françaises qui seront présentées à IJCAI 2017 à Melbourne. Les 42 chercheurs de France qui ont un article accepté à IJCAI 2017 ont été invités à venir le présenter lors de cette journée. Tous n'ont malheureusement pas pu se libérer, mais 7 ont pu répondre positivement à cette invitation.

Cette journée est inscrite au sein de la Plate-Forme Intelligence Artificielle 2017.

Date et Lieu

- Date 07 juillet 2017
- Lieu Université de Caen Normandie, Campus 2, Côte de Nacre, bd. Maréchal Juin, 14032 Caen

Programme

- 10h15 Présentation de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau *(Président de l'AFIA)*
- 10h30 « A Recursive Shortcut for CEGAR : Application to The Modal Logic K Satisfiability Problem » par Jean-Marie Lagniez, Daniel Le Berre, Tiago de Lima, et Valentin Montmirail *(Université d'Artois)*
- 10h45 « Constraint Games Revisited » par Anthony Palmieri et Arnaud Lallouet *(Huawei Boulogne-Billancourt)*
- 11h00 « An Improved Decision-DNNF Compiler » par Pierre Marquis *(Université d'Artois)* et Jean-Marie Lagniez.
- 11h15 « Temporal Sequences of Qualitative Information : Reasoning about the Topology of Constant-Size Moving Regions » par Quentin Cohen-Solal *(Université de Caen Normandie)*, Maroua Bouzid et Alexandre Niveau
- 11h30 « Strategically Knowing How » par Andreas Herzig *(CNRS Toulouse)*
- 11h45 « End-to-end optimization of goal-driven and multi-modal dialogue systems » par Florian Strub, Jeremie Mary *(Université Lille 3)*, Bilal Piot et Olivier Pietquin
- 12h00 « Belief Manipulation Through Propositional Announcements » par Aaron Hunter, Francois Schwarzentruber *(ENS Rennes)* et Eric Tsang
- 12h15 Clôture

Organisation

Cet événement est organisé par Serena Villata pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Christian Bessière pour l'International Joint Conference on Artificial Intelligence (IJCAI) 2017.

Inscriptions

Les inscriptions se font sur le site de PFIA 2017 : https://pfia2017.greyc.fr/inscriptions L'inscription à la Journée inclut la participation au repas et aux pause-café.

PDIA 2017

L'Association Française pour l'Intelligence Artificielle (AFIA) organise sa quatrième journée PERSPECTIVES ET DEFIS DE l'IA sur le thème : OUTILS LOGICIELS POUR L'IA. De nombreux outils, logiciels, bibliothèques ont été créés par et pour la recherche et le développement en IA. Ils participent à ses progrès et à son expansion. La journée vise à en dresser un panorama, à travers exposés et démonstrations.

Cette journée se tient en partenariat avec la Direction Générale de la Recherche et de l'Innovation (DGRI) du MESRI et avec la Direction Générale des Entreprises (DGE) du MINEF.

Date et Lieu

- Date : 06 octobre 2017
- Lieu : Université Paris-Descartes, Amphithéâtre Lavoisier, 45 rue des Saints Pères, 75006 Paris

Programme

- 08h30 Accueil
- 08h45 Ouverture de Pavlos Moraitis (*Directeur du LIPADE*), Benoît Leblanc (*Chargé de mission IA DGRI MESRI*) et Cédric Nozet (*Référent IA DGE MINEF*). Introduction par Pierre Zweigenbaum (*Vice-Président de l'AFIA*)
- 09h00 « Hommage à Alain Colmerauer » par Vincent Risch (*Université Aix-Marseille*)
- 09h45 « Représentation des problèmes combinatoires sous contraintes » par Christophe Lecoutre (*Université d'Artois*)
- 10h30 Pause
- 11h00 « Langages du Web de données et moteurs de recherche sémantique » par Olivier Corby (*Inria Sophia Antipolis*)
- 11h45 « Existe-t-il un couteau suisse pour la gestion des ontologies » par Sylvie Desprès (*Université Paris 13*)
- 12h30 Buffet
- 14h00 « Stratégie Nationale sur l'IA » par Marc Schoenauer (*Inria Saclay, expert auprès de Cédric Villani, chargé de mission par le Premier Ministre pour une Stratégie Nationale sur l'IA*). « Activités et Perspectives de l'AFIA » par Yves Demazeau (*Président de l'AFIA*)
- 14h45 « Outils et langages de programmation d'applications multi-agents » par Olivier Boissier (*Mines Saint-Étienne*)
- 15h30 Pause
- 16h00 « Des outils libres pour l'apprentissage statistique » par Gaël Varoquaux (*Inria Saclay*)
- 16h45 « Outils et plateformes pour le TAL » par Thierry Hamon (*Université Paris 13*)
- 17h30 « SAT : À l'assaut des problèmes difficiles » par Laurent Simon (*Université de Bordeaux*)
- 18h15 Clôture

Organisation

Cet événement est organisé par Pierre Zweigenbaum (*CNRS Orsay*), Catherine Faron-Zucker (*Université Nice-Sophia Antipolis*) et Engelbert Mephu Nguifo (*Université Clermont Auvergne*) pour l'Association Française d'Intelligence Artificielle (AFIA).

Inscriptions

Les inscriptions à la journée (gratuites pour les membres de l'AFIA, 30 € sinon) sont obligatoires et à effectuer sur le site : http://inscriptions.afia.asso.fr/ . Le déjeuner est offert à toutes les personnes qui se seront inscrites à PDIA 2017 avant le 03/10/17, et seulement à celles-ci.

FIIA 2018

Ce 3ème FORUM INDUSTRIEL DE L'INTELLIGENCE ARTIFICIELLE est consacré au « raisonnement », l'un des fondements de l'IA symbolique permettant la manipulation des connaissances d'un domaine, à travers différents thèmes. Chaque thème est composé de courtes présentations suivies d'une mini-table ronde dont l'objectif est d'aboutir à des éléments de feuille de route et de permettre des échanges accrus entre académiques et industriels. Une dernière session est réservée à des présentations rapides par des industriels invités.

Cette journée se tient sous le patronage de la Direction Générale de l'Armement (DGA) avec le partenariat de la Direction Générale de la Recherche et de l'Innovation (DGRI).

Date et Lieu

- Date : 12 avril 2018
- Lieu : Université Paris-Descartes, Amphithéâtre Fourier, 45 rue des Saints Pères, 75006 Paris

Programme

- 08h30 Accueil
- 08h45 Ouverture de Frédéric Dardel *(Président de l'Université Paris Descartes)* et Pavlos Moraitis *(Directeur du LIPADE)*. Introduction de la journée par Frédéric Maris (Coordinateur du Collège Représentation et Raisonnement de l'AFIA) et Bruno Patin (Coordinateur du Collège Industriel de l'AFIA).
- 09h00 « Man Machine Teaming » par Bruno Patin (Dassault Aviation)
- 09h30 Thème « Traitement d'informations symboliques et sémantiques » avec Sylvain Gatepaille (Airbus), Claire Laudy (Thales), Marie-Jeanne Lesot (Univ. Paris 6) et Sabri Bayoudh (ARCURE).
- 11h00 Pause
- 11h15 Thème « Résolution de problèmes complexes et optimisation combinatoire » avec Thierry Benoit (LocalSolver), Pierre Savéant (Thales), Gérald Petitjean (Eurodécision) et Eric-Olivier Leindecker (Sopra-Stéria).
- 12h45 Buffet
- 14h00 « Donner un sens à l'Intelligence Artificielle » par Marc Schoenauer (*Inria Saclay, membre de la Mission Villani)*.
- 14h30 Thème « Raisonnement, aide à la décision » avec Nicolas Museux (Thales), Thierno Kante (Univ. Nantes), Nedra Mellouli-Nauwynck (Univ. Paris 8), Sébastien Ecault (E-Cobot) et Isabelle Fantoni (CNRS) et Guillaume Lozenguez (IMT Lille Douai).
- 16h00 Pause
- 16h15 Présentations rapides de sociétés concernées par l'IA en « Trois planches »
- 17h30 Clôture

Organisation

Cet événement est organisé par Lamia Belaouer pour le Collège Représentation et Raisonnement de l'AFIA et Juliette Mattioli pour le Collège Industriel de l'AFIA.

Inscriptions

Les inscriptions à la journée (gratuites pour les membres de l'AFIA, 30 € sinon) sont obligatoires et à effectuer sur le site : http://inscriptions.afia.asso.fr/ . Le déjeuner est offert à toutes les personnes qui se seront inscrites à FIIA 2018 avant le 09/04/18, et seulement à celles-ci.

France@IJCAI 2018

Cette journée propose un avant-goût des recherches françaises qui seront présentées à IJCAI 2018 à Stockholm. Les 28 chercheurs de France qui ont un article accepté à IJCAI 2018 ont été invités à venir le présenter lors de cette journée. Tous n'ont malheureusement pas pu se libérer, mais 19 ont pu répondre positivement à cette invitation.

Cette journée est inscrite au sein de la Plate-Forme Intelligence Artificielle 2018.

Date et Lieu

- Date 02-06 juillet 2018
- Lieu Université de Lorraine, Campus Scientifique, Bâtiment C, 54506 Vandoeuvre-lès-Nancy

Programme

- 02 14h00 « Ceteris paribus majority for social ranking » par Adrian Haret et al.
- 02 14h20 « Open Loop Execution of Tree-Search Algorithms » par Erwan Lecarpentier et al.
- 02 14h40 « Efficient computation of approximate equilibria in discrete Colonel Blotto games » par Dong Quan Vu et al.
- 02 17h00 « Behavior of analogical inference w.r.t. Boolean functions » par Miguel Couceiro et al.
- 03 15h30 « A Decentralised Approach to Intersection Traffic Management » par Huan Vu et al.
- 03 15h50 « Conflict Directed Clause Learning for Maximum Weighted Clique Problem » par Emmanuel Hébrard et al.
- 03 16h10 « Boosting MCS Enumeration » par Éric Grégoire et al.
- 03 16h30 « An Information Theory based Approach to Multisource Clustering » par Pierre Alexandre Murena et al.
- 03 16h50 « Pseudo-Boolean Constraints from a Knowledge Representation Perspective » par Daniel Le Berre et al.
- 03 17h10 « A Distributed Model Counter » par Jean-Marie Lagniez et al.
- 04 10h35 « Accountable approval sorting » par Khaled Belahcene et al.
- 04 10h55 « Explaining Multi-Criteria Decision Aiding Models with an Extended Shapley Value » par Christophe Labreuche et al.
- 04 13h40 « Accelerated Difference of Convex Functions Algorithm and its Application to Sparse Binary Logistic Regression » par Duy Nhat Phan et al.
- 04 14h00 « Alternating-time Temporal Logic on Finite Traces » par Francesco Belardinelli et al.
- 04 14h20 « Small undecidable problems in epistemic planning » par Sébastien Lê Cong et al.
- 04 16h20 « Possibilistic ASP Base Revision by Certain Input » par Laurent Garcia et al.
- 04 16h40 « Belief Update in the Horn Fragment » par Nadia Creignou et al.
- 04 17h00 « Service Exchange Problem » par Julien Lesca et al.
- 05 14h30 « Learning Conceptual Space Representations of Interrelated Concepts » par Zied Bouraoui et al.

Organisation

Cet événement est organisé par Serena Villata pour l'Association Française pour l'AFIA et Pierre Marquis pour IJCAI 2018.

Inscriptions

Les inscriptions, incluant déjeuner et pauses-café, se font sur le site http://pfia2018.loria.fr/inscriptions/.

Nuit de l'Info 2016

La Nuit de l'Info 2016 a eu lieu dans la nuit du 1er au 2 décembre sur le thème « Venir en aide aux réfugiés ». Le principe de la Nuit de l'Info est très simple : Le jeudi 2, au coucher du Soleil, 16h40, en séance plénière : les organisateurs remettent un sujet (le même pour toute la France) aux participants. Les étudiants s'organisent en groupes (sur un ou plusieurs sites) : ils développent un projet (informatique, marketing, rédactionnel, ...) tout en ciblant un ou plusieurs défis. Le vendredi 3, au lever du Soleil, 08h00, les développements sont figés. Le vendredi matin, pendant que les étudiants dorment des jurys se réunissent (un jury par défi) et examinent les travaux réalisés par les différentes équipes durant la nuit.

La Compétition

Le défi qu'a proposé l'AFIA s'intitulait « Mettez de l'intelligence dans votre moteur » : « Vous mettez en œuvre une ou plusieurs méthodes d'Intelligence Artificielle (IA) dans votre projet et vous indiquez en quoi ces méthodes rendent votre réalisation plus performante ou pertinente. Vous pouvez faire appel à des techniques classiques, en cours de développement ou futuristes. L'équipe qui aura mis le plus en avant les avantages de l'utilisation de l'IA dans son projet remportera ce défi ». Les réponses devaient inclure :

- Une description des problématiques IA rencontrées dans le projet,
- Une explication de l'intégration de l'IA (effective ou imaginaire) dans l'application
- Une analyse des avantages et inconvénients de cette intégration

La récompense était un chèque de 1000 € assorti d'une invitation à présenter une démonstration lors de la compétition "IA et Jeux Vidée" organisée par l'AFIA au sein de la Plate-Forme IA 2017 à Caen.

Les Equipes

Parmi les 19 équipes inscrites pour relever ce défi, 4 ont réussi à remettre à temps un projet : « root@pls:~> sudo nano & », « ASpico », « Les fortran chocolat », et « Le Coin C'Net ».

Le Jury

Le jury était composé en 2017 de Carole Adam, Cindy Even, Thomas Guyet, et Florence Bannay (coordinatrice), s'est réuni par visio-conférence le 4 décembre de 9h à 11h. Les projets ont été jugés selon les critères: applicabilité / mise en œuvre, innovation, IA (évaluation qualitative et quantitative : richesse de la solution proposée, spectre couvert), et qualité des explications

Les Résultats

La première place et la récompense ont été attribuées à l'équipe « root@pls:~> sudo nano & » dont l'application « Botcher » est bien adaptée au sujet national, elle fournit un chatbot capable de donner des informations aux réfugiés dans plusieurs langues. Elle fonctionne sur mobile, exploite des techniques d'IA existantes (mais sans apporter d'innovation). Les explications de cette solution sont bien rédigées et claires.

Le jury a classé en deuxième position l'équipe « Le Coin C'est net » qui propose l'utilisation d'un robot Nao pour répondre à des questions, mémoriser le visage, le nom et la langue de l'utilisateur afin de personnaliser les interactions. Nous avons noté également que cette application n'est pas bien adaptée au thème de cette année, un robot Nao parait trop cher pour des réfugiés. De plus nous ne pouvons pas tester le code proposé sans disposer d'un robot Nao.

La troisième place a été attribuée à l'équipe « Les fortran chocolat » pour leur application qui contient un jeu vidéo destiné à amuser les réfugiés. Notons que ce jeu vidéo utilise des techniques d'IA mais ne répond pas très bien au sujet national de la nuit de l'informatique.

BOT
CONTEST
2017
COMPÉTITION
& Intelligence Artificielle
Jeux Vidéo
Lors de PFIA'17
à Caen, entre le 3 et le 7 Juillet 2017
Pour gagner, DÉVELOPPEZ LE BOT le plus HUMANLIKE*!
Plus d'info : http://goo.gl/rVZODq
Contact : botcontest2017@enib.fr
AfIA
Ne pas jeter sur la voie publique. Imprimé par l'atelier de reprographie de l'UBO. *Semblable à un humain.

Nuit de l'Info 2017

La Nuit de l'Info 2017 a eu lieu de 16h40 à 8h00 dans la nuit du 30 novembre au 1er décembre sur le thème de La sécurité routière. L'AFIA a proposé pour la 3e année un défi Mettez de l'IA dans votre moteur. Le principe de la Nuit de l'Info est très simple : Le jeudi 30, au coucher du Soleil, 16h40, en séance plénière : les organisateurs remettent un sujet (le même pour toute la France) aux participants. Les étudiants s'organisent en groupes (sur un ou plusieurs sites) : ils développent un projet (informatique, marketing, rédactionnel, ...) tout en ciblant un ou plusieurs défis. Le vendredi 1er, au lever du Soleil, 08h00, les développements sont figés. Le vendredi matin, pendant que les étudiants dorment des jurys se réunissent (un jury par défi) et examinent les travaux réalisés par les différentes équipes durant la nuit.

La Compétition

Le défi qu'a proposé l'AFIA s'intitulait « Mettez de l'intelligence dans votre moteur ». Voici sa description : « Vous mettez en œuvre une ou plusieurs méthodes d'Intelligence Artificielle (IA) dans votre projet et vous indiquez en quoi ces méthodes rendent votre réalisation plus performante ou pertinente. Vous pouvez faire appel à des techniques classiques, en cours de développement ou futuristes. L'équipe qui aura mis le plus en avant les avantages de l'utilisation de l'IA dans son projet remportera ce défi ». Les réponses devaient inclure :

- Une description des problématiques IA rencontrées dans le projet,
- Une explication de l'intégration de l'IA dans l'application (description de ce qui a été réalisé qui relève de l'IA ou de ce qui aurait pu être fait avec des outils d'IA)
- Une analyse des avantages et inconvénients de cette intégration.

La récompense était un chèque de 1000 € assorti d'une invitation à présenter une démonstration lors de la compétition organisée par l'AFIA au sein de la Plate-Forme IA 2018 à Nancy.

Les Equipes

Parmi les 14 équipes inscrites pour relever ce défi, 4 ont réussi à remettre à temps un projet : « ISIT'COM Talking Heads », « Canaille Space Advisors », « RobotSiffredi », et « Spooky Boys ».

Jury

Le jury était composé de Cindy Even, Thomas Guyet, Sylvie Saget et Florence Bannay (coordinatrice). Les projets ont été jugés sur cinq critères : Fonctionnement (est-ce que la proposition peut être mise en œuvre ?), Innovation, Qualité de l'IA, Spectre IA couvert, Qualité des explications.

Les Résultats

La première place et une récompense de 750 € ont été attribuées à l'équipe ISIT'COM Talking Heads : l'étude proposée est bien adaptée au sujet national, et suggère d'exploiter des techniques d'IA existantes. La réponse est bien expliquée. Plus précisément, l'équipe propose la reconnaissance de visage, la détection du clignotement de paupières, la reconnaissance d'émotions et de l'âge (en vue d'adapter les conseils) ainsi que la détection de piétons. Cette proposition est très novatrice. Le jury déplore l'absence d'une démonstration mais les images fournies sur la reconnaissance de fermeture de paupières permettent de juger la faisabilité.

Le jury a classé en deuxième position et a attribué un prix de 250 € à l'équipe Canaille Space Advisors [CSA]. L'équipe a utilisé des techniques d'apprentissage automatique, basées sur les réseaux de neurones, pour la construction d'un chatbot pouvant répondre automatiquement à des demandes d'aide d'un automobiliste. L'équipe a proposé une ébauche de solution technique à partir de librairies existantes. La solution présentée a été jugée originale et adaptée au sujet national.

Cla&IA 2016

La classification est au cœur des activités humaines. Lorsque l'on pense le monde, on l'organise et l'on classe les éléments qui le composent. Et, en tant que discipline s'étant initialement fixée pour objet les phénomènes constitutifs de l'intelligence humaine, l'IA s'est naturellement emparée de cette problématique et a produit de très nombreuses approches pour des problèmes variés. Selon les communautés, on distingue la classification de la discrimination, la classification non supervisée de la classification supervisée. La première s'ancre dans la longue tradition de la taxonomie. La seconde plus récente est associée à l'essor de l'apprentissage automatique. Mais aujourd'hui, avec le semi-supervisé, l'interactivité et les passages à l'échelle successifs, les questions posées s'entrecroisent et le dialogue entre les différentes approches semble plus que jamais nécessaire.

Cette journée CLASSIFICATION et INTELLIGENCE ARTIFICIELLE organisée conjointement par l'Association Française pour l'Intelligence Artificielle et la Société Francophone de Classification vise à faire le point sur des avancées récentes aux frontières des différentes communautés.

Date et Lieu

- Date 16 septembre 2016
- Lieu Salle du Conseil Université Paris Descartes, 12 rue de l'Ecole de médecine, 75006 Paris

Programme

- 00h00 Présentations de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau *(Président de l'AFIA) et* de la Société Francophone de Classification par Mohamed Nadif *(Président de la SFC).*
- 09h50 Présentation de la journée par Khalid Benabdeslem *(Université de Lyon)* et Pascale Kuntz *(Université de Nantes).*
- 10h00 « Classification non supervisée sous contraintes et programmation par contraintes » par Christel Vrain *(Université d'Orléans).*
- 11h00 « Classification et découverte de connaissances guidées sur l'analyse formelle de concepts » par Amedeo Napoli *(CNRS Nancy).*
- 12h00 Déjeuner libre
- 13h30 « Apprentissage statistique et optimisation à grande échelle » par Francis Bach *(Inria Paris).*
- 14h30 « Apprentissage de métriques : des algorithmes aux garanties théoriques » par Marc Sebban *(Université Jean Monnet, Saint Etienne).*
- 15h30 Pause
- 16h00 « Classification croisée : les modèles des blocs latents » par Mohamed Nadif *(Université Paris Descartes).*
- 17h00 Discussion
- 17h30 Clôture

Organisation

Cet événement est organisé par Khalid Benabdeslem pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Pascale Kuntz pour la Société Francophone de Classification (SFC).

Inscriptions

La participation à la journée est gratuite et ouverte à tous. Pour des raisons de logistique, nous vous demandons néanmoins de vous inscrire auprès de khalid.benabdeslem@univ-lyon1.fr .

MD&IA 2016

L'AFIA et le GdR MaDICS du CNRS organisent une journée d'étude sur l'articulation entre les problématiques posées par l'Intelligence Artificielle (IA) et celles posées par l'arrivée des données massives et le développement de la Science des Données (SD). L'IA ne constitue-t-elle qu'un terrain d'application aux techniques développées en SD ? Les réflexions menées en IA peuvent-elles impacter la manière dont les recherches en big data sont menées ? Les dernières avancées en SD sont-elles promises à résoudre la question du sens traitée en IA depuis de nombreuses années ? La notion de modèle va-t-elle disparaître au profit de systèmes ascendants fondés sur l'accumulation des données ? Pour le moins, les deux champs de recherche partagent des interrogations communes, comme la place de l'humain et le degré d'autonomie des systèmes mis en place. La journée constituera un lieu naturel de discussions et d'échanges sur ces questions.

La journée se déroulera au travers d'exposés et de posters portant sur des réflexions générales ou des applications particulières, et deux invités apporteront un éclairage complémentaire sur ces questions

Date et Lieu

- Date 2 décembre 2016
- Lieu Maison International des Langues et des Cultures, 35 rue Raulin, 75007 Lyon

Programme

- 10h00 Présentations de l'AFIA par Yves Demazeau *(Président de l'AFIA)*, du GdR MaDICS par Christine Collet *(Directrice du GdR)* et de la journée par Vincent Claveau et Julien Velcin.
- 10h30 « Big data vs. the right data » par Boi Faltings *(EPFL Lausanne)*
- 11h30 « Modélisation en Intelligence Artificielle au temps du *big data* : où est passée l'intelligence ? » par Fabrice Muhlenbach *(Univ. Saint Etienne)*
- 12h00 « Les machines ont elles perdu leur esprit ? » par Jean-Sébastien Vayre *(Univ. Toulouse 2)*
- 12h30 Pause Repas
- 14h00 « Learning deep representations in NLP » par Holger Schwenk *(FAIR Paris)*
- 15h00 « Combiner RDF, Datalog et SPARQL pour intégrer et analyser à la demande des données hétérogènes : retour d'expérience » par Marie-Christine Rousset et al *(Univ. Grenoble Alpes)*
- 15h30 Posters
  - « Big data for transport and logistics » par Mourad Abed et al *(Univ. Valenciennes)*
  - « What can I do ? Guiding users in a world of automated decisions » par Matthias Gallé *(XRCE Meylan)*
  - « Clustering sur de grands volumes de séquences d'origine biologique » par Alexandre Bazin et al *(Univ. Clermont-Ferrand)*
  - « From Mobility to Adaptive Services » par Nicolas Gutowski et al *(ESEO Angers)*
  - « Conception d'un modèle généraliste pour l'évaluation d'un test A/B » par Emmanuelle Claeys et al *(MC2I Groupe Paris)*
- 17h30 Discussion et clôture

Organisation

Cet événement est organisé par Julien Velcin pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Vincent Claveau pour le GdR Masse de Données, Informations et Connaissances en Sciences du CNRS (GdR MaDICS). La journée est organisée avec le soutien du laboratoire ERIC.

Inscriptions

Les inscriptions se font sur le site http://bigia2016.irisa.fr/#inscription. L'inscription est gratuite mais obligatoire. L'organisation et la prise en charge du déjeuner restent à la discrétion des participants.

EF&IA 2017

Cet événement commun organisé par le MEDEF et l'AFIA a pour ambition de promouvoir et de favoriser le développement de l'Intelligence Artificielle au sein des Entreprises de France. Des entreprises de toutes tailles et secteurs présenteront les applications et technologies d'IA qu'elles développent et les bénéfices qu'elles en tirent pour leur clientèle et leur propre croissance. L'objectif de la journée est double : informer de l'impact de l'Intelligence Artificielle (IA) pour la compétitivité des entreprises et échanger sur les perspectives de ces technologies pour la création de marchés d'avenir.

Date et Lieu

- Date 23 janvier 2017
- Lieu 55 avenue Bosquet 75007 Paris

Programme

- 09h00 Introduction par Pierre Gattaz, Président du MEDEF et Laurent Gouzènes, Président du Comité Recherche-Innovation au MEDEF *(MEDEF 75-Paris)*.
- 09h10 Intervention de Thierry Mandon, Secrétaire d'Etat à l'Enseignement Supérieur et à la Recherche *(MENESR 75-Paris)*
- 09h30 « Intelligence Artificielle de France » par Yves Demazeau, Président de l'AFIA.
- 09h50 « Le renouveau de l'Intelligence Artificielle et les Technologies Financières » par Yves Caseau, Directeur du Digital *(AXA Groupe Paris)*
- 10h10 « Ingénierie de la connaissance : mémoire de la société, transfert des savoirs, tremplin de l'innovation » par Alain Berger, Directeur Général *(Ardans 78-Montigny-le-Bretonneux)*
- 10h30 « L'Intelligence Artificielle, levier de progrès pour le développement des produits et services innovants de l'industrie automobile » par Patrick Bastard, Directeur de l'Ingénierie des Systèmes d'Aide à la conduite *(Renault 78-Guyancourt)*
- 10h50 « *We didn't do anything wrong, but somehow we failed*. Que faire pour réussir ? » par Patrick Albert, Président *(SuxccessionWeb 75-Paris)*
- 11h10 Pause-café
- 11h30 « Relation homme-système en aéronautique : l'apport de l'Intelligence Artificielle » par Bruno Patin, Responsable Etudes Amont Autonomie *(Dassault Aviation 92-Saint-Cloud)*
- 11h50 « Réseaux Bayésiens : modélisation des connaissances et apprentissage à partir des données pour l'aide à la décision » par Lionel Jouffe, Président *(Bayesia 53-Laval)*
- 12h10 « L'IA Appliquée aux Infrastructures Télécom et à la Relation Clients » par Henri Sanson, Directeur Recherche Données et Connaissances *(Orange Labs 35-Cesson-Sévigné)*
- 12h30 « Gestion optimisée des ressources : garantir la performance, intégrer l'humain, réduire l'empreinte environnementale des villes et des industries » par Jean-Denis Muller, Directeur Innovation Numérique *(Veolia Recherche & Innovation 78-Limay et Maisons-Laffitte)*
- 12h50 Discussion animée par Laurent Gouzènes, Directeur Scientifique *(Pacte Novation et Taj 92-Neuilly) et* Yves Demazeau, Directeur de Recherche (*CNRS 38-Grenoble*).
- 13h00 Cocktail

Organisation

Cet événement est organisé par Yves Demazeau et Bruno Patin pour l'Association Française pour l'Intelligence Artificielle (AFIA), et par Laurent Gouzènes et Patrick Schmitt pour le Mouvement des Entreprises DE France (MEDEF).

Inscriptions

Les inscriptions, gratuites, se font sur le site http://communication.medef.fr/LP/IhHybQRcPiOO.

PS&IA 2017

Dès son apparition dans les années 1950, l'Intelligence Artificielle (IA) a suscité une intense curiosité chez les philosophes. Soixante ans plus tard, l'IA semble en passe de remplir ses promesses et d'introduire une série de bouleversements économiques, culturels voire anthropologiques. Mais l'IA d'aujourd'hui est-elle celle d'hier ? Qu'annoncent ses succès présents ? Pour aborder ces questions, une réévaluation des fondements de l'IA et de sa place dans le système technoscientifique semble nécessaire.

À cette tâche, l'AFIA et la SPS organisent une journée PHILOSOPHIE DES SCIENCES et INTELLIGENCE ARTIFICIELLE, focalisée sur trois thèmes complémentaires : philosophie de l'IA et IA comme philosophie expérimentale ; impact du tournant pragmatique pour l'IA, en particulier approches 4EA (*Embodied*, *Embedded*, *Enacted*, *Extended*, *Affective*) ; apport des concepts d'émergence et de complexité.

Date et Lieu

- Date 02 février 2017
- Lieu Salle Jaurès, Ecole Normale Supérieure, 29 rue d'Ulm, 75005 Paris

Programme

- 08h30 Accueil des participants
- 09h00 Présentations de l'AFIA, par Yves Demazeau *(Président de l'AFIA)*, de la SPS par Stéphanie Ruphy *(Présidente de la SPS)*, du DEC par Daniel Andler. Présentation de la journée par Daniel Andler *(ENS Paris)* et Robin Lamarche-Perrin *(CNRS Paris)*
- 09h30 « L'IA comme épistémologie expérimentale : regards sur 40 ans d'histoire » par Joël Quinqueton *(U. Montpellier 3)*
- 10h20 Pause
- 10h40 « GOFAI, NEWFAI et le problème de l'intentionnalisme » par Jean-Michel Roy *(ENS Lyon)*
- 11h30 « Le rôle de la complexité et de la simplicité dans l'émergence des capacités cognitives » par Jean-Louis Dessalles *(Telecom ParisTech)*
- 12h20 Repas (offert)
- 13h40 « D'une intelligence artificielle à une intelligence incarnée dans l'activité humaine » par Catherine Garbay *(CNRS Grenoble)*
- 14h30 « Intelligence artificielle et *cognition 4E* : un changement de paradigme ? » par Pierre Steiner *(UTC Compiègne)*
- 15h20 Pause
- 15h40 « Questionnements éthiques sur la Robotique et l'IA » par Raja Chatila *(U. Paris 6)*
- 16h30 Discussion
- 17h30 Clôture

Organisation

Cet événement est organisé par Robin Lamarche-Perrin pour l'Association Française pour Intelligence Artificielle (AFIA) et Daniel Andler pour la Société de Philosophie des Sciences (SPS), avec le soutien du Département d'Etudes Cognitives de l'ENS Ulm.

Inscriptions

L'inscription est gratuite mais obligatoire : https://www.eventbrite.fr/e/inscription-journee-philosophie-des-sciences-et-intelligence-artificielle-30888523405. Elle inclut la participation au repas et aux pauses-café.

IHM&IA 2017

L'AFIHM et l'AFIA organisent leur troisième journée commune INTERACTION HOMME-MACHINE & INTELLIGENCEARTIFICIELLE.

Cette journée s'organisera autour de présentations, posters, démos, échanges sur des sujets à l'intersection des deux champs ou mettant en évidence les liens possibles entre les deux disciplines.

Date et Lieu

- Date 17 mars 2017
- Lieu UPMC, Amphithéâtre 25, 4 place Jussieu, 75005 Paris

Programme

- 08h30 Accueil
- 09h00 Présentations de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau *(Président de l'AFIA) et* de *l'AFIHM* par Géry Casiez *(Président de l'AFIHM).*
- 09h20 « Interface multimodale sur une table graphique multi-touch multi-utilisateurs » par Jean-Paul Barthès *(Université de Technologie de Compiègne)*
- 09h50 « Evolutionary visual exploration using EvoGraphDice » par Nadia Boukhelifa *(Telecom ParisTech)*
- 10h20 Pause
- 10h40 « Co-apprentissage pour les interfaces cerveau-ordinateur : de la science-fiction à l'interaction homme-machine » par Nataliya Kosmyna *(Inria Rennes)*
- 11h10 « Apport des systèmes multi-agents à l'interaction tangible sur table interactive RFID » par Christophe Kolski *(Université de Valenciennes et du Hainaut-Cambrésis)*
- 11h40 « Reconnaître un objet à partir de points de contact sur une surface tactile » par Caroline Appert *(CNRS Saclay)*
- 12h10 Déjeuner
- 14h00 « Vers une meilleure appréciation des algorithmes qui nous entourent » par Fanny Chevalier *(Inria Lille)*
- 14h30 « Conception participative d'un e-learning intelligent » par Alexandra Delmas *(Université de Bordeaux 2)*
- 15h00 « Les systèmes autonomes sont des outils informatiques comme les autres » par Nicolas Roussel *(Inria Lille)*
- 15h30 Pause
- 16h00 Démonstrations, posters, et échanges
- 18h00 Clôture

Organisation

Cet événement est organisé par Nicolas Sabouret pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Gilles Bailly pour L'Association Francophone d'Interaction Homme-Machine (AFIHM) avec le soutien du Labex SMART.

Inscriptions

Les inscriptions se font sur le site http://ihmia.afihm.org/inscription/index.html. L'inscription est gratuite et inclut la participation aux pause-café. L'organisation et la prise en charge du déjeuner restent à la discrétion des participants.

EIAH&IA 2017

L'AFIA et l'ATIEF organisent une troisième journée ENVIRONNEMENTS INFORMATIQUES POUR L'APPRENTISSAGE HUMAIN & INTELLIGENCE ARTIFICIELLE juste avant la conférence EIAH 2017.

L'objectif de cette journée est de rassembler dans un même espace-temps deux communautés, les chercheurs en Intelligence Artificielle et ceux travaillant dans le domaine des Environnements Informatiques pour l'Apprentissage Humain, afin de dialoguer et échanger autour de problématiques, de techniques, de concepts communs et de partager des retours d'expériences.

En 2013 il n'y avait pas de thème pour cette journée alors qu'en 2015 l'ensemble des présentations portaient sur l'utilisation de la fouille de données dans les EIAH. Cette année nous mixons ces deux approches. Tout d'abord, grâce à des présentations de papiers retenus suite à un appel à communications, nos échanges seront multi thématiques. Ensuite nous aborderons la problématique de la conception, la réalisation et l'expérimentation d'EIAH utilisant la réalité virtuelle et la réalité augmentée. Un invité nous fera part de son expertise et de ses retours d'expérience sur ce thème et nous terminerons avec une table ronde.

Date et Lieu

- Date 06 Juin 2017
- Lieu Hôtel du Département – Place du Quartier Blanc – 67000 Strasbourg

Programme

- 14h00 Présentations de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau *(Président de l'AFIA)* et de l'Association des Technologies de l'Information pour l'Education et la Formation par Serge Garlatti *(Président de l'ATIEF)*.
- 14h15 « Une architecture multicouche pour l'analyse des compétences non techniques en situation critique » par Yannick Bourrier *(Université Paris 6)*, Francis Jambon, Catherine Garbay et Vanda Luengo.
- 14h45 « Le contexte pour personnaliser les systèmes tutoriels » par Monique Grandbastien *(Université de Lorraine)*
- 15h15 « Apprentissage : corriger et visualiser » par Gabriel Illouz *(Université Paris 11)*, Anne-Laure Ligozat et Frédéric Vernier.
- 15h45 Pause
- 16h00 « Développer l'apprentissage et les compétences-piétons par la réalité virtuelle : enjeux, perspectives et limites » par Jérôme Dinet *(Université de Lorraine)*.
- 17h00 Table Ronde
- 17h45 Clôture

Organisation

Cet événement est organisé par Davy Monticolo pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Nicolas Delestre pour l'Association des Technologies de l'Information pour l'Education et la Formation (ATIEF).

Inscriptions

Les inscriptions se font via le site : http://eiah2017.unistra.fr/

JI&IA 2017

Cette deuxième journée JEUX INFORMATISES et INTELLIGENCE ARTIFICIELLE fait suite au succès rencontré lors de la première journée tenue lors de PFIA 2015 à Rennes. Tous les jeux sont concernés, aussi bien les jeux de plateau que les jeux vidéo ou les jeux à information incomplète.

L'objectif de cette deuxième journée est de rassembler la communauté française en IA des jeux (ou *Game AI*) et de présenter à ceux de nos collègues qui s'y connaissent moins ce qu'est le *Game AI* : ce qu'on y fait, et en quoi les jeux sont des terrains applicatifs idéaux pour l'intelligence artificielle.

Cette journée est inscrite au sein de la Plate-Forme Intelligence Artificielle 2017.

Date et Lieu

- Date 03 juillet 2017
- Lieu Université de Caen Normandie, Campus 2, Côte de Nacre, bd. Maréchal Juin, 14032 Caen

Programme

- 10h15 Présentations de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau *(Président de l'AFIA) et* du pré-GdR IA par Sébastien Konieczny *(Porteur du pré-GdR IA).*
- 10h30 « Stockfish, la meilleure IA d'Echecs au monde » par Stéphane Nicolet *(Université Paris 2)*
- 11h30 Série de présentations courtes
- 12h30 Déjeuner
- 14h00 « SPRITE : un jeu sérieux pour l'enseignement de la gestion des risques naturels » par Carole Adam *(Université Grenoble Alpes)*
- 14h30 Discussion autour des jeux sérieux. Pourquoi les jeux sérieux ? Quels enjeux ? Quelles possibilités de partenariat ?
- 15h30 Pause
- 16h00 « L'industrie du jeu vidéo aujourd'hui, ses besoins et ses attentes » par Eddy Célestine *(Altangames Nantes)*
- 16h30 Réunion du groupe de Travail « IA des Jeux » du pré-GdR IA.
- 18h00 Clôture

Organisation

Cet événement est organisé par Tristan Cazenave pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Florian Richoux pour le Groupe de Travail « IA des Jeux » du pré-GdR IA.

Inscriptions

Les inscriptions se font sur le site de PFIA 2017 : https://pfia2017.greyc.fr/inscriptions L'inscription à la Journée inclut la participation au repas et aux pause-café.

E&IA 2017

L'AFIA et le COMETS, en association avec la CNIL, organisent une deuxième Journée ÉTHIQUE et INTELLIGENCE ARTIFICIELLE avec pour vocation de croiser les questionnements et les approches, et de créer une communauté de réflexion sur ces sujets.

La question d'une régulation ou d'un contrôle éthique des agents logiciels ou robotiques ainsi que des systèmes constitués d'humains et d'agents autonomes se pose avec de plus en plus de force. Ces questions sont par nature transverses à de multiples disciplines. Le programme de la journée est constitué de présentations invitées et de présentations courtes relatifs à cette problématique.

Cette journée est inscrite au sein de la Plate-Forme Intelligence Artificielle 2017.

Date et Lieu

- Date 05 juillet 2017
- Lieu Université de Caen Normandie, Campus 2, Côte de Nacre, bd. Maréchal Juin, 14032 Caen

Programme

- 10h15 Présentations de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau *(Président de l'AFIA)* et du Comité d'Ethique du CNRS par Jean-Gabriel Ganascia *(Président du Comité d'Ethique)*.
- 10h30 « Contre la transparence. De la valeur du hasard pour un système apprenant » par Alexei Grinbaum *(CEA Saclay)*
- 11h15 « Une théorie formelle des agents moraux » par Emiliano Lorini *(CNRS Toulouse)*
- 11h40 « Approches logiques de la modélisation du raisonnement éthique » par Gauvain Bourgne *(Université Paris 6)*
- 12h30 Déjeuner
- 14h00 « Robots conversationnels empathiques : enjeux technologiques et éthiques » par Laurence Devillers *(Université Paris 4)*
- 14h45 « Pistes de réflexion sur une modélisation de l'éthique pour des agents intelligents autonomes » par Léa Guizol *(Ardans Montigny-le-Bretonneux)*
- 15h05 « ETHICAAA - Ethique et Agents Autonomes » par Grégory Bonnet *(Université de Caen)*
- 15h30 Pause
- 16h00 « Les problèmes juridiques posés par la robotique et l'IA » par Nathalie Nevejans *(Université d'Artois)*
- 16h45 Table ronde « *IEEE Standard Ethics* » avec L. Devillers, A. Grinbaum, N. Nevejans, et C. Tessier.
- 18h00 Clôture

Organisation

Cet événement est organisé par Olivier Boissier pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Jean-Paul Delahaye pour le Comité d'Ethique (COMETS) du CNRS.

Inscriptions

Les inscriptions se font sur le site de PFIA 2017 : https://pfia2017.greyc.fr/inscriptions L'inscription à la Journée inclut la participation au repas et aux pause-café.

RO&IA 2017

Cette quatrième journée RECHERCHE OPERATIONNELLE et INTELLIGENCE ARTIFICIELLE se déroule le 18 Septembre 2017 à Paris. La journée est organisée conjointement par le Collège Représentation et Raisonnement de l'Association Française pour l'Intelligence Artificielle (AFIA) et la Société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF).

L'objectif de ces journées est de rassembler les communautés de l'IA et de la RO et d'échanger autour de problématiques, de techniques et de concepts communs. La journée se focalisera cette année sur l'enrichissement mutuel entre RECHERCHE OPERATIONNELLE, L'APPRENTISSAGE ET LA FOUILLE DE DONNEES.

Date et Lieu

- Date 18 septembre 2017
- Lieu Salle A709, Université Paris-Dauphine, Pl. du M. de Lattre de Tassigny, 75016 Paris

Programme

- 10h00 Présentations de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau *(Président de l'AFIA)* et de la Société Française de Recherche Opérationnelle et d'Aide à la Décision par *Sourour Elloumi (Présidente de la ROADEF)*.
- 10h20 « Learning to Decompose : supervised learning for column generation » par Axel Parmentier *(ENPC Paris)*
- 11h05 « *30 ans d'intelligence dans les données* » par Gérald Petitjean, Sarah Bugalho et Toufic Zaraket *(EURODECISION Versailles)*
- 11h50 Pause
- 12h00 « Apprentissage des modèles de scores » par Yann Chevaleyre *(Univ. Paris 13 Villetaneuse)*
- 12h45 Déjeuner
- 14h00 « Apprentissage et fouille de données pour l'optimisation combinatoire : quelques cas d'études » par Jin Kao Hao *(Univ. d'Angers)*
- 14h45 « Towards cross-fertilization between Data Mining and Constraints » par Lakdar Sais *(Université d'Artois Lens)*
- 15h30 Pause
- 15h40 Discussion
- 17h00 Clôture

Organisation

Cet événement est organisé par Nicolas Maudet (Univ. Paris 6) pour le Collège Représentation et Raisonnement de l'Association Française pour l'Intelligence Artificielle (AFIA) et Meltem Ozturk pour la Société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF).

Inscriptions

Les inscriptions se font sur https://docs.google.com/forms/d/e/1FAIpQLSfsM5Cir_LDxWov32a4eE-wcfavIlXVA0Y18-EhCqObG5v3HQ/viewform . L'inscription est gratuite et inclut la participation aux pause-café. L'organisation et la prise en charge du déjeuner restent à la discrétion des participants.

MACS&IA 2017

Cette troisième journée entre le GdR MACS et l'AFIA est consacrée à la CONDUITE DES SYSTEMES DYNAMIQUES et INTELLIGENCE ARTIFICIELLE.

Le contrôle de systèmes complexes tels que les systèmes de production nécessite l'assistance d'une IA plus en plus répartie au niveau des ressources (machines, …) et des produits – intelligents -. L'intégration de ces entités intelligentes au sein d'un système réel (intégrant des opérateurs humains) avec lequel elles communiquent forme un système cyber-physique.

La journée est inscrite au sein des 24èmes journées Sciences et Technique de la Production du GdR MACS.

Date et Lieu

- Date 9 Novembre 2017
- Lieu AIP Lorrain - 745 Rue du Jardin Botanique - 54600 Villers-lès-Nancy

Programme

- 09h00 « Utilisation de l'internet des objets dans les systèmes de production » par Detlef Zühlke *(DFKI Aix-la-Chapelle)* (Conférence invitée du STP-MACS)
- 10h00 Pause
- 10h15 Présentations de l'AFIA par Yves Demazeau *(Président de l'AFIA)* et du GdR Modélisation, Analyse et Conduite de Systèmes, par André Thomas *(Membre du Comité de Direction)*
- 10h30 « L'autonomie dans la prise de décision » par Olivier Cardin *(Université de Nantes)* et Damien Trentesaux
- 11h15 « Modèle multi-agent d'inspiration physique pour la conduite de véhicules autonomes » par Franck Gechter *(Université de Technologie de Belfort-Montbéliard)*
- 12h00 « Les mécanismes de prise de décision en contexte distribué » par Hind Bril *(Université de Lorraine Nancy)* et Fouzia Ounnar
- 12h45 Déjeuner
- 13h50 « MECSYCO : une plateforme pour la co-simulation de systèmes cyber-physiques » par Vincent Chevrier *(Université de Lorraine Nancy)*
- 14h35 « La gestion de la myopie en architecture distribuée » par William Deringent *(Université de Lorraine Nancy)* et Emmanuel Adam
- 15h20 « De l'ingénierie des collectifs cyber-physiques pour les systèmes de production distribués » par Jean-Paul Jamont *(Université Grenoble Alpes Valence)*
- 16h05 Clôture

Organisation

Cet événement est organisé par Emmanuel Adam pour le Collège Systèmes Multi-Agents et Agents Autonomes de l'Association Française pour l'Intelligence Artificielle (AFIA), Olivier Cardin et William Derigent pour le Groupe de Travail Intelligent Manufacturing & Systems Services du GdR Modélisation, Analyse et Conduite de Systèmes (MACS).

Inscriptions

L'inscription est gratuite mais obligatoire, à prendre auprès de emmanuel.adam@univ-valenciennes.fr. Elle inclut la participation au déjeuner (dans la limite des places disponibles).

RI&IA 2017

Cette deuxième journée RECHERCHE D'INFORMATION et INTELLIGENCE ARTIFICIELLE se focalise sur thème des connaissances et porte sur l'extraction d'information à partir de documents en vue de construire ou peupler des bases de connaissances, sur l'annotation sémantique de documents, le raisonnement et leur apprentissage.

Date et Lieu

- Date 12 décembre 2017
- Lieu Auditorium de la MSH Paris Nord, 20 avenue George Sand 93210 Saint-Denis

Programme

- 09h20 Présentations de l'AFIA, par Yves Demazeau *(Président de l'AFIA) et* de l'ARIA par Max Chevalier *(Président de ARIA).*
- 09h40 « Learning with Knowledge Graphs: From Medical Decision Support to Human Perception and Memory » par Volker Tresp *(Ludwig Maximilian Univ. Munich).*
- 10h40 « Fouille de publications scientifiques pour une analyse bibliométrique de l'activité de recherche sur la déforestation » par Nathalie Neptune *(Univ. Toulouse 2)* et al (Poster)
- 10h40 « Les plateformes sémantiques : une étude empirique » par Fatma Chamek *(IAE Lyon 3)*
- 10h40 « Classification d'images pour la catégorisation de produits sur un site de e-commerce » par Elie Guardia Sebaoun *(CDiscount Bordeaux)*
- 10h40 « Défis de l'enrichissement du peuplement multilingue d'une ontologie à partir de corpus » par Yuliya Korenchuk *(Univ. Strasbourg)*
- 11h20 « RI et ontologies en médecine » par Jean Charlet *(Université Paris 6 et AH-HP).*
- 12h20 « Test de l'apport de Word2Vec sur des comptes rendus de maintenance pour la constitution semi-automatique de ressources lexicales » par Meryl Bothua (*EDF R&D Paris*) et al.
- 12h40 « Construction de thésaurus assisté par apprentissage machine » par Thomas Perrais *(Proxem R&D Paris)*
- 13h00 Déjeuner
- 14h15 « Learning multimodal word representations: visual grounding and Visual Question Answering » par Patrick Gallinari *(Université Paris 6)*
- 15h15 « Annotation sémantique d'un corpus pour les besoins agronomiques » par Catherine Roussey *(IRSTEA Clermont-Ferrand)*
- 15h35 « Traitement automatique des langues pour la recherche d'information et vice versa » par Vincent Claveau *(CNRS Rennes)*
- 17h00 « Entre annotation et recherche d'information sémantiques, les défis de l'accès à l'information juridique » par Adeline Nazarenko *(Université Paris 13)*
- 18h00 Clôture

Organisation

Cet événement est organisé par Brigitte Grau (CNRS Orsay) pour l'AFIA et Haifa Zargayouna (Université Paris 13) pour l'ARIA, avec le soutien du LIMSI et du LIPN.

Inscriptions

Les inscriptions se font sur https://ia-ri.sciencesconf.org/. L'inscription est gratuite mais obligatoire. Elle inclut la participation aux pause-café. Le repas sur site n'est garanti que pour les 50 premiers inscrits.

EGC&IA 2018

Cette troisième journée EXTRACTION ET GESTION DES CONNAISSANCES et INTELLIGENCE ARTIFICIELLE réunit les deux communautés autour du thème des « DONNEES PERSONNELLES, VIE PRIVEE ET ETHIQUE ». À l'ère de Big Data, chaque utilisateur de notre monde hyper-connecté laisse derrière lui une myriade de traces numériques tout en effectuant ses activités quotidiennes.

L'objectif de la journée est d'encourager les discussions et les recherches fondées sur des principes qui conduiront à l'avancement de l'analyse des données personnelles, du développement des services personnels, de la protection de la vie privée, du respect de l'éthique, de la protection des données et de l'évaluation des risques liés à la vie privée.

Date et Lieu

- Date 23 janvier 2018
- Lieu Auditorium de la MSH Paris Nord, 20 avenue George Sand 93210 Saint-Denis

Programme

- 13h00 Présentations de l'AFIA, par Yves Demazeau *(Président de l'AFIA) et* de EGC par Fabrice Guillet *(Président de EGC)*.
- 13h20 Présentation de la journée par Jérôme Azé (*Université de Montpellier)* et Thierry Charnois *(Université de Paris 13)*.
- 13h30 « Contexte juridique de la circulation internationale des données personnelles » par Annie Blandin *(Telecom Bretagne Brest)*.
- 14h30 Pause
- 15h00 « A declarative approach to linked data anonymization » par Angela Bonifati *(Université Lyon 1)*, Rémy Delanaux, Marie-Christine Rousset et Romuald Thion.
- 15h30 « Co-clustering for differentially private synthetic data generation » par Tarek Benkhelif *(Oranges Labs Meylan)*, Françoise Fessant, Fabrice Clérot and Guillaume Raschia.
- 16h00 « Système de gestion de l'obsolescence dans une base d'informations personnelles » par Salma Chaieb *(Université de Valenciennes et du Hainaut Cambrésis)*, Véronique Delcroix, Ali Benmrad et Emmanuelle Grislin.
- 16h30 Pause
- 17h00 « Traitement Automatique des Langues et Ethique » par Karën Fort *(Université Paris-Sorbonne)*.
- 18h00 Clôture

Organisation

Cet événement est organisé par Jérôme Azé pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Thierry Charnois pour l'Association Extraction et Gestion des Connaissances (EGC).

Inscriptions

L'atelier se tient dans le cadre de la conférence EGC 2018. Les inscriptions se font sur le site de la conférence (https://egc18.sciencesconf.org/resource/page/id/20).

E&IA 2018

Cette troisième journée ETHIQUE et INTELLIGENCE ARTIFICIELLE est organisée par le Comité Opérationnel d'Evaluation des Risques Légaux et Ethiques d'Inria (COERLE) et l'AFIA pour croiser les questionnements et les approches, et de contribuer à créer une communauté de réflexion sur ces sujets.

La journée est organisée en association avec le CERNA.

Cette journée est inscrite au sein de la Plate-Forme Intelligence Artificielle 2018.

Date et Lieu

- Date 02 Juillet 2018
- Lieu Université de Lorraine, Campus Scientifique, Bâtiment C, 54506 Vandoeuvre-lès-Nancy

Programme

- 08h30 « Amnestic Forgery » par Aldo Gangemi *(Université Paris Nord)*
- 09h30 Pause-café
- 09h45 Présentations de l'AFIA, par Yves Demazeau *(Président de l'AFIA)* et du COERLE de INRIA par Claude Kirchner *(Président du COERLE)*.
- 10h05 « Un éclairage sur le rapport IEEE Ethically Aligned Design » par Raja Chatila *(Université Paris 6)*
- 10h55 « Un éclairage sur les rapports CNIL - Comment permettre à l'homme de garder la main ? - et CERNA - Éthique de la recherche en apprentissage machine - par Claude Kirchner *(INRIA Grand-Est, Nancy)*
- 11h45 Déjeuner
- 13h30 « Éthique et traitement automatique des langues » par Maxime Amblard (*Université de Lorraine*) et Karën Fort (*Sorbonne Université*)
- 14h00 « Interprétabilité et validation d'applications métiers basées sur de l'IA statistique » par Christophe Denis *(EDF Miramas)*
- 14h30 « De l'IA mimétique à l'IA autonome en santé : une boîte à outil éthique pour évaluer le statut des utilisateurs humains dans la relation de soin » par Joël Colloc *(Université Le Havre Normandie)* et Relwendé Aristide Yameogo *(Université Le Havre Normandie)*
- 15h00 Pause-café
- 15h30 « Contributions de l'IA encorporée et de la robotique développementale aux questions d'éthique en robotique » par Alexandre Pitti *(Université de Cergy-Pontoise)*
- 16h00 « L'éthique peut-elle être mise en algorithme » par Nazim Fatès *(INRIA Grand-Est Nancy)*
- 16h30 Table Ronde « Jeunes Chercheurs en Ethique et IA » avec Fiona Berreby *(Université Paris 6)*, Azzedine Benabbou *(UTC Compiègne)*, et Relwendé Aristide Yameogo *(Université Le Havre Normandie)*
- 17h30 Clôture

Organisation

Cet événement est organisé par Grégory Bonnet pour l'AFIA et Pascal Guitton pour le COERLE de l'INRIA, en association avec Catherine Tessier pour le CERNA.

Inscriptions

Les inscriptions, incluant déjeuner et pauses-café, se font sur le site http://pfia2018.loria.fr/inscriptions/.

IM&IA 2018

Cette troisième journée INFORMATIQUE MEDICALE et INTELLIGENCE ARTIFICIELLE fait se rencontrer les deux disciplines autour du thème de la SANTE. Les thématiques de la journée concernent le traitement intelligent de données et des connaissances en santé, les représentations et les méthodes de l'IA.

Cette journée est inscrite au sein de la Plate-Forme Intelligence Artificielle 2018.

Date et Lieu

- Date 03 Juillet 2018
- Lieu Université de Lorraine, Campus Scientifique, Bâtiment C, 54506 Vandoeuvre-lès-Nancy

Programme

- 08h30 « Brain Machine Integration » par Zhongzhi Shi *(Académie Chinoise des Sciences Pékin)*
- 09h30 Pause-café
- 09h45 Présentations de l'AFIA, par Yves Demazeau *(Président de l'AFIA)* et de AIM par *Philippe* Cinquin *(Président de AIM)*.
- 10h00 « Définitions, réalisations et considérations éthiques » par Jean Charlet *(APHP Paris)*
- 10h45 « ROMEDI, une terminologie médicale française pour la détection des médicaments en texte libre » par Sébastien Cossin *(CHU Bordeaux)* et al
- 11h05 « Découverte d'associations entre évènements indésirables médicamenteux par les structures de patrons et les ontologies » par Gabin Personeni *(Université de Lorraine)* et al
- 11h25 « Artificial intelligence assistance for fetal head biometry : assessment of automated measurement software » par Gaëlle Ambroise Grandjean *(Université de Lorraine)* et al
- 11h45 Déjeuner
- 13h30 « De la difficulté d'évaluer les algorithmes de scorage automatique des données médicales temporelles : exemple de la polysomnographie » par Adrien Ugon *(ESIEE Paris)*
- 14h00 « Utilisation de cercles thérapeutiques pour l'affichage des recommandations de bonne pratique dans la prise en charge de patients » par Brigitte Séroussi *(Université Paris 11)* et al
- 14h30 « Design d'un algorithme d'IA en grande dimension pour prédire la réadmission à l'hôpital » par Simon Bussy *(Université Paris 6)* et al
- 15h00 Pause-café
- 15h30 « Pré-screening en cancérologie : automatisation à partir des entrepôts de données cliniques » par Jeremy Pasco *(Université de Tours)* et al
- 16h00 « MultiWAS : interactive and multimodal phenome-wide scan for data-driven cohort description » par Antoine Neuraz *(APHP Paris)* et al
- 16h30 « Intégration et synthèse visuelle de données multi-sources et hétérogènes de chimiothérapie » par Anne-Sophie Jannot *(APHP Paris)* et al
- 17h00 « Interprétation de bonnes pratiques de codification médicale par du raisonnement à partir de cas » par Michael Schnell *(Institut de la Santé du Luxembourg)* et al
- 17h30 Clôture

Organisation

Cet événement est organisé par Jean Charlet pour l'AFIA et Christine Aigret, Marc Cuggia, Brigitte Seroussi et Pascal Staccini pour AIM.

Inscriptions

Les inscriptions, incluant déjeuner et pauses-café, se font sur le site http://pfia2018.loria.fr/inscriptions/.

Edu&IA 2018

Cette première journée EDUCATION et INTELLIGENCE ARTIFICIELLE fait se rencontrer les deux domaines autour du thème ANALYSE D'APPRENTISSAGE.

L'objectif de cette journée est de rassembler les chercheurs en IA et ceux travaillant dans le domaine de l'éducation, des entreprises du secteur de l'éducation et des praticiens pour échanger autour de problématiques, de techniques et de concepts communs.

Cette journée est inscrite au sein de la Plate-Forme Intelligence Artificielle 2018.

Date et Lieu

- Date 04 Juillet 2018
- Lieu Université de Lorraine, Campus Scientifique, Bâtiment C, 54506 Vandœuvre-lès-Nancy

Programme

- 08h30 « The Automated-Reasoning Revolution » par Moshe Vardi *(Rice University Houston)*
- 09h30 Pause-café
- 09h45 Présentations de l'AFIA, par Yves Demazeau *(Président de l'AFIA) et* du DNE-Numérilab par Vanda Luengo *(Animatrice GTnum2 de la Direction du Numérique pour l'Education).*
- 10h00 « Tuteurs intelligents : boucler la boucle » par Yolaine Bourda *(Supélec Gif-sur-Yvette)*, C. Chaudet, B. Choffin, J. Parmentier, F. Popineau et J.-J. Vie.
- 10h25 « Un algorithme bandit manchot pour le choix de nouvelles situations d'apprentissage à l'intérieur d'un environnement virtuel » par Yannick Bourrier *(Sorbonne Université Paris)*, V. Luengo, F. Jambon, C. Garbay et J. Teigny.
- 10h50 « X5-GON : Vers l'utilisation de l'Intelligence Artificielle pour une meilleure utilisation des Ressources Éducatives Libres » par Colin de la Higuera *(Université de Nantes)*, H. Le Capitaine, W. Ben Romdhane, P. Leray et N. Hernandez.
- 11h15 « Manier l'intelligence artificielle sans coder : création de chatbots éducatifs » par Samuël Di Emidio *(Université de Mons)*, G. Temperman et B. De Lièvre.
- 12h00 Déjeuner
- 13h30 « Que peut apporter la fouille de données aux Learning Analytics ? » par Armelle Brun *(Université de Lorraine)*
- 14h15 « Using a Predictive Analytics Process to identify droppers in MOOCS » par Alya Itani *(Telecom Bretagne)*, L. Brisson, I. Rebai et S. Garlatti.
- 14h40 « Explorer les traces d'apprenants avec l'extraction d'épisodes séquentiels » par Béatrice Fuchs *(IAE Lyon)*
- 15h05 « SEDELA » par Jean-Marie Gilliot *(Telecom Bretagne)*, I. Rebai, S. Garlatti.
- 15h30 Pause-café
- 16h00 « Analyse de l'hybridation entre les communautés LAK, EDM et AIED » par Hugues Labarthe *(Sorbonne Université Paris)*, V. Luengo, F. Bouchet
- 16h30 Discussion et synthèse de la journée

Organisation

Cet événement est organisé par Marie Lefèvre pour l'AFIA et Vanda Luengo pour le DNE-Numérilab du MEN.

Inscriptions

Les inscriptions, incluant déjeuner et pauses-café, se font sur le site http://pfia2018.loria.fr/inscriptions/.

Robotique & IA 2018

Cette troisième journée ROBOTIQUE et INTELLIGENCE ARTIFICIELLE fait se rencontrer les deux domaines autour des thèmes APPRENTISSAGE POUR LA ROBOTIQUE HUMANOIDE le matin, et VEHICULES AUTONOMES l'après-midi.

L'objectif de la journée est de faire le point sur l'impact des avancées en IA sur la Robotique, et à l'inverse, de la Robotique sur les problématiques et résultats de l'IA

Cette journée est inscrite au sein de la Plate-Forme Intelligence Artificielle 2018.

Date et Lieu

- Date 05 Juillet 2018
- Lieu Université de Lorraine, Campus Scientifique, Bâtiment C, 54506 Vandœuvre-lès-Nancy

Programme

- 08h30 « One Robot for Every Task » par Daniela Rus *(MIT Boston)*
- 09h30 Pause-café
- 09h45 Présentations de l'AFIA, par Yves Demazeau *(Président de l'AFIA)* et du GdR Robotique par Philippe Martinet *(Directeur du GdR Robotique)*.
- 10h00 « Apprentissage ouvert en robotique, l'approche du projet DREAM » par Stéphane Doncieux *(Sorbonne Université Paris)*
- 10h30 « Génération de mouvements pour les robots humanoïdes : Optimisation, contrôle et mémoire » par Olivier Stasse *(CNRS Toulouse)*
- 11h00 « Jugement, prudence, sagacité, moralité : les différentes facettes d'un agent éthique illustrées par les retours d'expérience sur la voiture autonome » par Jean-Gabriel Ganascia *(Sorbonne Université Paris)*
- 11h30 « Autonomous cars navigation on roads opened to public traffic : How can infrastructure-based systems help ? » par Philippe Bonnifait *(Sorbonne Université Compiègne)*
- 12h00 Déjeuner
- 13h30 « Apprentissage par Renforcement Développemental » par Alain Dutech *(Inria Nancy)*
- 14h00 « Apprentissage de représentations d'état pour la robotique » par David Filliat *(ENSTA ParisTech)*
- 14h30 « Planification et Ordonnancement pour Systèmes Robotisés » par Christophe Guettier *(SAFRAN Paris)*
- 15h00 « Coopération dans les flottes de robots mobiles : enjeux et approches » par Olivier Simonin *(INSA Lyon)*
- 15h30 Pause-café
- 16h00 Contributions courtes par Roxana Agrigoroaie *(ENSTA ParisTech)* Mélanie Jouaiti *(Université de Lorraine)* Chuang Yu *(Paris Saclay)* Nicolas Pépin-Hermann *(BA Systèmes Rennes)* Moad Kissai *(ENSTA ParisTech)* et Pierre-Henri Orefice *(ENSTA ParisTech)*.
- 17h00 Discussion et clôture

Organisation

Cet événement est organisé par Serena Ivaldi pour l'AFIA et Alexandre Pitti pour le GdR Robotique du CNRS.

Inscriptions

Les inscriptions, incluant déjeuner et pauses-café, se font sur le site http://pfia2018.loria.fr/inscriptions/.

TAL & IA 2018

Cette quatrième journée TRAITEMENT AUTOMATIQUE DES LANGUES et INTELLIGENCE ARTIFICIELLE fait se rencontrer les deux domaines autour du thème RESEAUX PROFONDS et REPRESENTIONS CONTINUES DE MOTS.

Cette journée vise ainsi à faire le point sur les méthodes actuellement employées en IA & TAL, notamment les travaux en cours sur les réseaux profonds et représentations continues de mots, leurs limites et les recherches entreprises pour les pallier.

Cette journée est inscrite au sein de la Plate-Forme Intelligence Artificielle 2018.

Date et Lieu

- Date 06 Juillet 2018
- Lieu Université de Lorraine, Campus Scientifique, Bâtiment C, 54506 Vandœuvre-lès-Nancy

Programme

- 08h30 « 25 Years of Applied Ontology and Ontological Analysis » par Nicola Guarino *(CNR Trento)*.
- 09h30 Pause-café
- 09h45 Présentations de l'AFIA, par Yves Demazeau *(Président de l'AFIA)* et de *l'ATALA* par Patrick Paroubek *(Président de l'ATALA)*.
- 10h00 « Plongements lexicaux pour l'analyse discursive automatique » par Chloé Braud *(CNRS Nancy)*.
- 10h45 « Identification automatique de chaînes de coréférences : vers une analyse des erreurs pour mieux cibler l'apprentissage » par Frédéric Landragin *(CNRS Paris)* et Bruno Oberlé.
- 11h15 « Production endogène de règles déductives dans le réseau JDM » par Mathieu Lafourcade *(Université Montpellier 2)* et Alain Joubert.
- 11h45 Déjeuner
- 13h30 « Modèles de langue neuronaux à grand vocabulaire » par Alexandre Allauzen *(Université Paris-Sud Orsay)*.
- 14h15 « Comparaison de systèmes de traduction automatique, probabiliste et neuronal, par analyse d'erreurs » par Emmanuelle Esperança-Rodier *(Université Grenoble Alpes)* et Nicolas Becker.
- 15h00 Pause-café
- 15h30 « An Efficient Semantic Graph-Based Approach for Text Representation » par Ahmed Mabrouk *(ENGIE Lab Paris)*, Rim Hantach et Philippe Calvez
- 16h00 « Modélisation d'un contexte global d'étiquettes pour l'étiquetage de séquences dans les réseaux neuronaux récurrents » par Marco Dinarelli *(CNRS Paris)* et Loïc Grobol.
- 16h30 Clôture

Organisation

Cet événement est organisé par Pierre Zweigenbaum pour l'AFIA et Didier Schwab pour l'ATALA.

Inscriptions

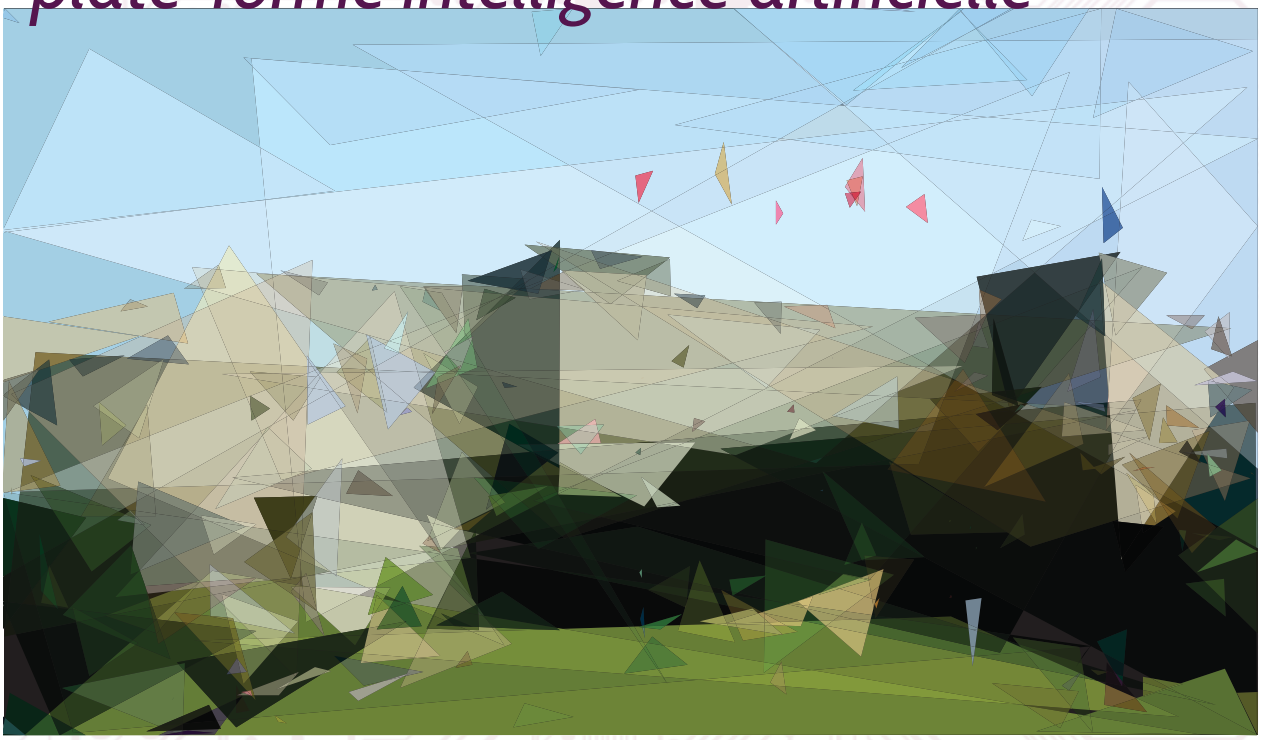
Les inscriptions, incluant déjeuner et pauses-café, se font sur le site http://pfia2018.loria.fr/inscriptions/.

APIA • IAF • IC • JFPDA • JFSMA • RJCIA

jeux informatisés & IA • éthique & IA • compétition d'IA • démos • FR@IJCAI

PFIA17

plate-forme intelligence artificielle



du 3 au 7 juillet 2017

Université de Caen Normandie – Campus 2

https://pfia2017.greyc.fr

APIA & CNIA & IC & JFPDA & RJCIA

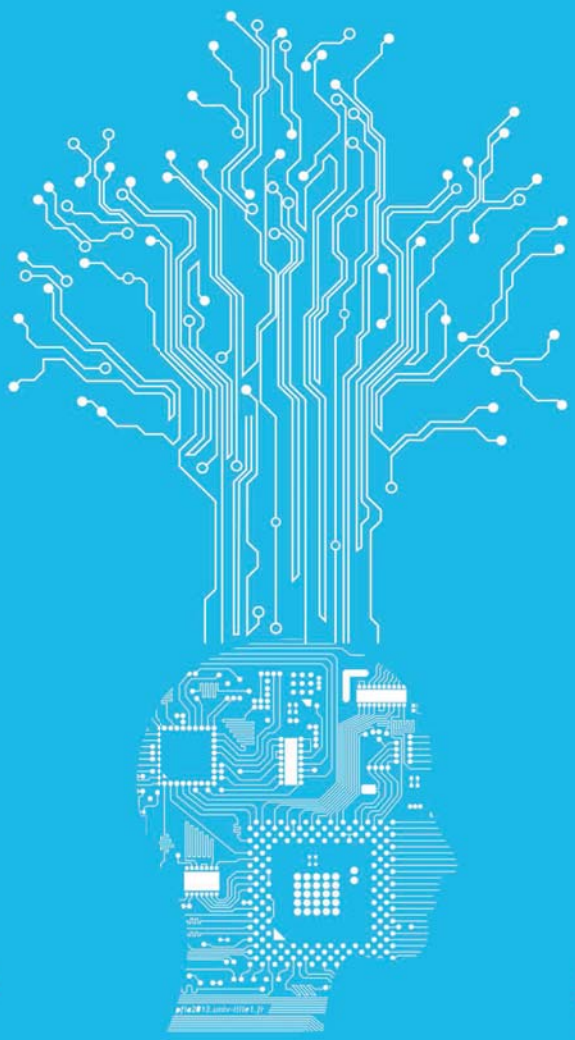
IA pour l'éducation IA & Santé TAL & IA Ethique & IA
Robotique & IA France@IJCAI2018

PFIA 2018

11e Plate-forme Intelligence Artificielle

2 au 6 juillet 2018 - Nancy

Campus Sciences - Université de Lorraine
Vandœuvre-lès-Nancy



PFIA2018.LORIA.FR

En partenariat avec

Berger Levrault

L'AVENIR EST AUX VALEURS SÛRES